



**SURVEYING &
SPATIAL SCIENCES
CONFERENCE**

Proceedings of the **Surveying and Spatial Sciences Conference** 21-25 November 2011, Wellington, New Zealand

ISBN: 978-0-478-11042-1



Land Surveying / Remote Sensing & Photogrammetry / Engineering & Mining / Hydrographic Surveying / Spatial Information & Cartography

Surveying and Spatial Sciences Conference 2011

Platinum sponsor:



Gold sponsors:



Australian Government
Geoscience Australia

Conference convenors:



Proceedings of the
SURVEYING & SPATIAL SCIENCES BIENNIAL CONFERENCE 2011
21-25 November 2011, Wellington, New Zealand

SSSC2011 Committee list

Steering Committee:

Mark Dyer (NZIS President)
Gypsy Bhalla (SSSI President)
Barry Davidson (NZIS National Manager)
Roger Buckley (SSSI CEO)
Steve Critchlow (SIBA NZ Chair)

Convenors:

Karl Majorhazi
Jeff Needham

Committee:

Chris Weir
Barbara Hock
Tony Standen
Andrew Clouston
Catherine O'Shaughnessey
Melissa Bethwaite
Fran Lovell
Dene Lynch
Colin McElwain

Conference Organisers:

Paardekooper and Associates

This work is licensed under a [Creative Commons Attribution 3.0 New Zealand License](https://creativecommons.org/licenses/by/3.0/).

Papers published in these proceedings were submitted by authors in electronic media. While the papers have been reviewed, authors are responsible for content and accuracy of their individual papers. The use of trade or firm names in this publication is for reader information and does not imply endorsement by the conference organisers.

Proceedings published by Scion, Rotorua, New Zealand

ISBN 978-0-478-11042-1

Preface

The theme of the Surveying and Spatial Sciences Conference 2011, Innovation in Action – Working Smarter, is certainly reflected by the papers within this conference proceedings. They form a solid basis for moving a number of fields forward.

Bob Williams sets the scene with a challenge on how the surveying and spatial sciences can continue to transform us. Multi-purpose data, delivered, used.

The topics in this proceeding include surveying advances, remote sensing, models, and data management and delivery. Web sourcing location descriptions bridge the continuum from precise measures to imprecise concepts of place. 3D cloud data adds comprehensiveness to the concept of precision. Environmental models maximise reuse of data and hence their value. From emergency response to land loss through earthquakes and coastal erosion, the diversity of topics touches on many surveying and spatial developments, ultimately improving how society functions.

Barbara Hock, Editor-in-Chief

Reviewers:

Thomas Adams, Scion

Luke Barry, Scion

Sharon Burrell, Scion

Andrew Dunningham, Scion

Duncan Harrison, Scion

Barbara Hock, Scion

Matt Amos, LINZ

Andrew Clouston, LINZ

John Dymond, Landcare Research

Colin McElwain, Cuttriss Consultants

Antoni (Tony) Moore, University of Otago

Formatting: Katie Payn, student, University of Canterbury

Proceedings production: John Smith, Scion



INDEX

Welcome to Te Ika a Maui OIESS 2020 R.J. (Bob) Williams	1
Using Volunteered Information to Map the Queensland Floods Kevin McDougall	13
Embedding GIS in the Departments of Health and Human Services Clare Brazenor, Miles da Costa	25
Building Better Hyperspectral Datasets: The Fundamental Role of Metadata Protocols in Hyperspectral Field Campaigns Barbara Rasaiah, Simon Jones, Chris Bellman, Tim Malthus	35
Geospatial enablement of the Māori Land Court's www.maorilandonline.govt.nz : A Case Study Tobias Luetticke	49
Starting to Talk about Place Stephan Winter, Allison Kealy, Matt Duckham, Abbas Rajabifard, Kai-Florian Richter, Tim Baldwin, Lesley Stirling, Lawrence Cavedon, Daniela Richter	63
Harvesting User-Generated Content for Semantic Spatial Information: The Case of Landmarks in OpenStreetMap Kai-Florian Richter, Stephan Winter	75
Spatially Enabling Risk for Management of Land and Property Katie Potts, Abbas Rajabifard, Ian Williamson, Rohan Bennett	87
Analysing the Effect of Drought on Net Primary Productivity of Tropical Rainforests in Queensland using MODIS Satellite Imagery Sipho S.T. Shiba, Armando Apan	97
Implementing the Phoenix Fire Spread Model for Operational Use Gillian Paterson, Derek Chong	111
Modelling Habitat Networks Using the Concept of Graph Theory Fatemeh Poodat, Colin Arrowsmith, Elizabeth Farmer	125
Towards a Stream Hierarchy and Classification System for Victorian Waterways Phillip Delaney, Joanne Poon, Zaffar Sadiq Mohamed-Ghouse, Derek Goodin, Kate Wilson, Darren McKinty	139
The Victorian Land Use Information System (VLUIS): A new method for creating land use data for Victoria, Australia Elizabeth Morse-McNabb	155
Issues related to the use of non-traditional, digital data sources for enhancing park management data Monique Elsley, William Cartwright	169
Understanding the relationship between spatial information, property markets and macroeconomic policy Nilofer Tambuwala, Rohan Bennett, Abbas Rajabifard, Ian Williamson	187

Review of Australian Land Use Mapping and Land Management Practice Govinda Prasad Baral, Kevin McDougall, Albert Chong	199
Developing Spatio-temporal Prediction Models for Arbovirus Activity in Northern Australia Based on Remotely Sensed Bioclimatic Variables Bernhard Klingseisen, Robert J. Corner, Mark Stevenson	211
Innovative applications of remotely sensed evapotranspiration and vegetation cover Kathryn Sheffield, Mohammad Abuzar, Des Whitfield, Andy McAllister, Mark O'Connell, Lexie McClymont	221
Disaster Change Detection Using Airborne LiDAR John Trinder, Mahmoud Salah	231
Evaluation of Photo Imaging Methods for Vegetation Condition Assessment Armando Apan, Govinda Baral, Ernest Dunwoody, Lucy Richardson, Kevin McDougall	243
Semi-Automated Colour Registration and Evaluation of Digital Photogrammetry and Terrestrial Laser Scanning Kwanthar Lim, Kwang-Ho Bae and David Belton	259
Compressing Images Using Contours Gabriel Scarmana	269
Standardizing Australian Semi-Urban land cover mapping through FAO Land Cover Classification System Kithsiri Perera, David Moore, Armando Apan, Kevin McDougall	279
A dense surface modelling technique for foot surface imaging Jasim Ahmed Ali AL-Baghdadi, Albert K. Chong, Kevin McDougall, Duaa Alshadli, Peter Milburn, Richard Newsham-West	295
Automated Matching of Segmented Point Clouds to As-built Plans David Belton, Brian Mooney, Tony Snow and Kwang-Ho Bae	303
Assessing Spatial Information Access, Use and Sharing for Catchment Management in Australia Dev Raj Paudyal, Kevin McDougall, Armando Apan	317
Shove over Gen Y: Gen Z is almost here Gita Pupedis, Chris Bellman	331
Mapping Cycling Pathways and Route Selection Using GIS and GPS Matthew Huntley, Xiaoye Liu, Kevin McDougall, Peter Gibbings	341
Wetland Mapping Using Remote Sensing Imagery and ModelMap Xuan Zhu	353
Characterising Heterogeneous Vegetated Surfaces Using Multiangular Satellite Data Geoff McCamley, Ian Grant, Simon Jones, Chris Bellman	365
Mapping Coastal Carbon Sinks in South Australia Matt Miles, Darcy Peters and Peter Fairweather	377
Preparing for the Invasion – Retreat from the coast Mick Strack	385

CORSnet-NSW: Improving Positioning Infrastructure for New South Wales Volker Janssen, Joel Haasdyk, Simon McElroy and Doug Kinlyside	395
How will all the new GNSS signals help RTK surveyors? Craig Roberts	411
Interpolation of the GNSS Wet Troposphere Delay Johnny Lo, Ahmed El-Mowafy	425
Schema Element Dependencies in a Federated Spatial Database System Xiaoying Wu, Jianhong (Cecilia) Xia, Geoff West, Bert Veenendaal, Lesley Arnold	439

INDEX by title

A dense surface modelling technique for foot surface imaging	295
Analysing the Effect of Drought on Net Primary Productivity of Tropical Rainforests in Queensland using MODIS Satellite Imagery	97
Assessing Spatial Information Access, Use and Sharing for Catchment Management in Australia	317
Automated Matching of Segmented Point Clouds to As-built Plans	303
Building Better Hyperspectral Datasets: The Fundamental Role of Metadata Protocols in Hyperspectral Field Campaigns	35
Characterising Heterogeneous Vegetated Surfaces Using Multiangular Satellite Data	365
Compressing Images Using Contours	269
CORSnet-NSW: Improving Positioning Infrastructure for New South Wales	395
Developing Spatio-temporal Prediction Models for Arbovirus Activity in Northern Australia Based on Remotely Sensed Bioclimatic Variables	211
Disaster Change Detection Using Airborne LiDAR	231
Embedding GIS in the Departments of Health and Human Services	25
Evaluation of Photo Imaging Methods for Vegetation Condition Assessment	243
Geospatial enablement of the Māori Land Court's www.maorilandonline.govt.nz : A Case Study	49
Harvesting User-Generated Content for Semantic Spatial Information: The Case of Landmarks in OpenStreetMap	75
How will all the new GNSS signals help RTK surveyors?	411
Implementing the Phoenix Fire Spread Model for Operational Use	111
Innovative applications of remotely sensed evapotranspiration and vegetation cover	221
Interpolation of the GNSS Wet Troposphere Delay	425
Issues related to the use of non-traditional, digital data sources for enhancing park management data	169
Mapping Coastal Carbon Sinks in South Australia	377
Mapping Cycling Pathways and Route Selection Using GIS and GPS	341
Modelling Habitat Networks Using the Concept of Graph Theory	125
Preparing for the Invasion – Retreat from the coast	385
Review of Australian Land Use Mapping and Land Management Practice	199
Schema Element Dependencies in a Federated Spatial Database System	439
Semi-Automated Colour Registration and Evaluation of Digital Photogrammetry and Terrestrial Laser Scanning	259
Shove over Gen Y: Gen Z is almost here	331
Spatially Enabling Risk for Management of Land and Property	87
Standardizing Australian Semi-Urban land cover mapping through FAO Land Cover Classification System	279
Starting to Talk about Place	63
The Victorian Land Use Information System (VLUIS): A new method for creating land use data for Victoria, Australia	155
Towards a Stream Hierarchy and Classification System for Victorian Waterways	139
Understanding the relationship between spatial information, property markets and macroeconomic policy	187
Using Volunteered Information to Map the Queensland Floods	13
Welcome to <i>Te Ika a Maui</i> OIESS 2020	1
Wetland Mapping Using Remote Sensing Imagery and ModelMap	353

INDEX by first author

AL-Baghdadi, Jasim Ahmed Ali	A dense surface modelling technique for foot surface imaging Jasim Ahmed Ali AL-Baghdadi, Albert K. Chong, Kevin McDougall, Duaa Alshadli, Peter Milburn, Richard Newsham-West	295
Apan, Armando	Evaluation of Photo Imaging Methods for Vegetation Condition Assessment Armando Apan, Govinda Baral, Ernest Dunwoody, Lucy Richardson, Kevin McDougall	243
Baral, Govinda Prasad	Review of Australian Land Use Mapping and Land Management Practice Govinda Prasad Baral, Kevin McDougall, Albert Chong	199
Belton, David	Automated Matching of Segmented Point Clouds to As-built Plans David Belton, Brian Mooney, Tony Snow and Kwang-Ho Bae	303
Brazenor, Clare	Embedding GIS in the Departments of Health and Human Services Clare Brazenor, Miles da Costa	25
Delaney, Phillip	Towards a Stream Hierarchy and Classification System for Victorian Waterways Phillip Delaney, Joanne Poon, Zaffar Sadiq Mohamed-Ghouse, Derek Goodin, Kate Wilson, Darren McKinty	139
Elsley, Monique	Issues related to the use of non-traditional, digital data sources for enhancing park management data Monique Elsley, William Cartwright	169
Huntley, Matthew	Mapping Cycling Pathways and Route Selection Using GIS and GPS Matthew Huntley, Xiaoye Liu, Kevin McDougall, Peter Gibbings	341
Janssen, Volker	CORSnet-NSW: Improving Positioning Infrastructure for New South Wales Volker Janssen, Joel Haasdyk, Simon McElroy and Doug Kinlyside	395
Klingseisen, Bernhard	Developing Spatio-temporal Prediction Models for Arbovirus Activity in Northern Australia Based on Remotely Sensed Bioclimatic Variables Bernhard Klingseisen, Robert J. Corner, Mark Stevenson	211
Lim, Kwanthar	Semi-Automated Colour Registration and Evaluation of Digital Photogrammetry and Terrestrial Laser Scanning Kwanthar Lim, Kwang-Ho Bae and David Belton	259
Lo, Johnny	Interpolation of the GNSS Wet Troposphere Delay Johnny Lo, Ahmed El-Mowafy	425
Luetticke, Tobias	Geospatial enablement of the Māori Land Court's www.maorilandonline.govt.nz : A Case Study Tobias Luetticke	49

McCamley, Geoff	Characterising Heterogeneous Vegetated Surfaces Using Multiangular Satellite Data Geoff McCamley, Ian Grant, Simon Jones, Chris Bellman	365
McDougall, Kevin	Using Volunteered Information to Map the Queensland Floods Kevin McDougall	13
Miles, Matt	Mapping Coastal Carbon Sinks in South Australia Matt Miles, Darcy Peters and Peter Fairweather	377
Morse-McNabb, Elizabeth	The Victorian Land Use Information System (VLUIS): A new method for creating land use data for Victoria, Australia Elizabeth Morse-McNabb	155
Paterson, Gillian	Implementing the Phoenix Fire Spread Model for Operational Use Gillian Paterson, Derek Chong	111
Paudyal, Dev Raj	Assessing Spatial Information Access, Use and Sharing for Catchment Management in Australia Dev Raj Paudyal, Kevin McDougall, Armando Apan	317
Perera, Kithsiri	Standardizing Australian Semi-Urban land cover mapping through FAO Land Cover Classification System Kithsiri Perera, David Moore, Armando Apan, Kevin McDougall	279
Poodat, Fatemeh	Modelling Habitat Networks Using the Concept of Graph Theory Fatemeh Poodat, Colin Arrowsmith, Elizabeth Farmer	125
Potts, Katie	Spatially Enabling Risk for Management of Land and Property Katie Potts, Abbas Rajabifard, Ian Williamson, Rohan Bennett	87
Pupedis, Gita	Shove over Gen Y: Gen Z is almost here Gita Pupedis, Chris Bellman	331
Rasaiah, Barbara	Building Better Hyperspectral Datasets: The Fundamental Role of Metadata Protocols in Hyperspectral Field Campaigns Barbara Rasaiah, Simon Jones, Chris Bellman, Tim Malthus	35
Richter, Kai-Florian	Harvesting User-Generated Content for Semantic Spatial Information: The Case of Landmarks in OpenStreetMap Kai-Florian Richter, Stephan Winter	75
Roberts, Craig	How will all the new GNSS signals help RTK surveyors? Craig Roberts	411
Scarmana, Gabriel	Compressing Images Using Contours Gabriel Scarmana	269
Sheffield, Kathryn	Innovative applications of remotely sensed evapotranspiration and vegetation cover Kathryn Sheffield, Mohammad Abuzar, Des Whitfield, Andy McAllister, Mark O'Connell, Lexie McClymont	221
Shiba, Sipho S.T.	Analysing the Effect of Drought on Net Primary Productivity of Tropical Rainforests in Queensland using MODIS Satellite Imagery Sipho S.T. Shiba, Armando Apan	97
Strack, Mick	Preparing for the Invasion – Retreat from the coast Mick Strack	385

Tambuwalla, Nilofer	Understanding the relationship between spatial information, property markets and macroeconomic policy Nilofer Tambuwalla, Rohan Bennett, Abbas Rajabifard, Ian Williamson	187
Trinder, John	Disaster Change Detection Using Airborne LiDAR John Trinder, Mahmoud Salah	231
Williams, R.J. (Bob)	Welcome to Te Ika a Maui OIESS 2020 R.J. (Bob) Williams	1
Winter, Stephan	Starting to Talk about Place Stephan Winter, Allison Kealy, Matt Duckham, Abbas Rajabifard, Kai-Florian Richter, Tim Baldwin, Lesley Stirling, Lawrence Cavedon, Daniela Richter	63
Wu, Xiaoying	Schema Element Dependencies in a Federated Spatial Database System Xiaoying Wu, Jianhong (Cecilia) Xia, Geoff West, Bert Veenendaal, Lesley Arnold	439
Zhu, Xuan	Wetland Mapping Using Remote Sensing Imagery and ModelMap/Xuan Zhu	353

Welcome to *Te Ika a Maui* OIESS 2020

Dr. R.J.Williams

Retired

Tura Beach, New South Wales, Australia
rosnbobw@bigpond.com.au

ABSTRACT

The beginning of this decade witnessed an array of natural and man-made events that seriously challenged many nations' abilities to predict, respond, recover and progress from the effects of tragic events. Iceland's Eyjafjallajökull volcanic eruption; Haiti's and Christchurch's earthquakes; Queensland's, New South Wales' and Victoria's floods; Western Australia's fires; Japan's earthquakes and tsunami along with Poland's Tu-154 air crash and the Shen Neng 1 grounding on the Great Barrier Reef have all stressed nations' emergency response capabilities.

The year is now 2020 and most nations are still struggling to respond any better to incidents than they were at the beginning of the decade. The exception is the *Te Ika a Maui* Oceania Infrastructure and Environmental Support System (OIESS).

The *Te Ika a Maui* OIESS evolved throughout this decade to lead the world in sophisticated monitoring, strategic and operational planning, commanding and controlling a range of manned and autonomous vehicles and assets, and providing intelligence and informed advice to key authorities.

This paper takes us on an odyssey; reviewing past initiatives and predicting future capabilities. The paper shows how this capability came to be, here in *Aotearoa*, and why those 'over the ditch' lagged behind!

KEYWORDS: Cartography; Geospatial Information Infrastructure; Natural Disaster Information Integration; Decision Support; Future Navigation

1 PROLOGUE

A new era in the way we plan and manage our infrastructure and our environment has arrived.

It is a balmy 20 degrees Celsius this evening, Wednesday 18 November 2020, at Pukenui, New Zealand. I am enjoying a pleasant conversation and a glass of New Zealand white wine with Professor Michael Goodchild outside AEPCOTAT's Conference Centre. AEPCOTAT stands for Aotearoa Experimental Prototype Community of Today and Tomorrow¹. Adjoining AEPCOTAT is the visionary system the *Te Ika a Maui* Oceania Infrastructure and Environmental Support System (OIESS). It is the OIESS that will enable a new era and AEPCOTAT that will prototype some of the new concepts initiated in the laboratories.

I recall a memorable discussion with Mike on Thursday 15 September 2005 at the SSC2005 Spatial Science Institute Biennial Conference dinner in the National Gallery of Victoria. We started our discussion by talking about Mike's opening address to the conference titled "What Does Google Earth Mean for the Spatial Sciences?" Mike went on to describe the elegance of the

¹ Adapted from Walt Disney World's EPCOT theme park. EPCOT is one of 4 Theme Parks at Walt Disney World Resort and is divided into Future World and World Showcase. **Future World** is full of sensational attractions as well as inspiring entertainment and shows, all of which *focus on technological advancements, innovation and wonder*. In **Future World**, *ideas become reality*.

mathematics model that underpinned Google Earth: the tessellation of an octahedron. Whilst that discussion was interesting, what followed was more so. Mike indicated that he was a contributor to (then) Vice-President Al Gore's 1998 paper titled "The Digital Earth: Understanding our Planet in the 21st Century"². Mike³ used the following quote from Gore's paper in his presentation to highlight Al Gore's Digital Earth vision:

"Imagine, for example, a young child going to a Digital Earth exhibit. She sees Earth as it appears from space. Using a data glove, she zooms in, using higher and higher levels of resolution, to see continents, then regions, countries, cities, and finally individual houses, trees, and other natural and man-made objects. Having found an area of the planet she is interested in exploring, she takes the equivalent of a 'magic carpet ride' through a 3-D visualization of the terrain".

Gore's 1998 paper, according to Mike, was the 'political driver' to take, what some would view as revolutionary, the next evolution in mapping - *a digital earth*.

During the evening we chatted about the *Digital Earth* and the evolution of our digital capability, contemporary challenges and future concepts. I mentioned that many issues are difficult for senior managers to understand so it seemed appropriate to use a 'story-telling' approach. I have continued to use a 'story-telling' approach at conferences since and, so, now onto 'future' stories.

So now the 2020 conference has commenced and the theme is concerned with Managing Events and Incidents in a Complex environment.

2 MANAGING EVENTS AND INCIDENTS IN A COMPLEX ENVIRONMENT

By 2020 the Oceania has become a truly complex environment:

- International trade and associated infrastructure costs have become contemporary challenges driven by very competitive markets in the region. The efficient management of the infrastructure and the protection of the infrastructure are vitally important.
- Terrorism remains a serious concern in the South-East Asia – Oceania region; concerns that include the safety of transportation. National security is a high priority for the nation's government.
- Environmental incidents seem to have become far more extreme and more common. The past decade has witnessed devastating events that have contributed to loss of life and high financial costs.

Managing our nation, managing our environment, managing our natural resources, managing our infrastructure and managing our profession are fundamental requirements⁴. It is now time to embrace a sophisticated decision support system that facilitates natural disaster information integration; commands and controls incident response to national security tasks; interacts with future navigation systems; and achieves outcomes via an advanced Geospatial Information Infrastructure.

The *Te Ika a Maui* Oceania Infrastructure and Environmental Support System (Figure 1) has been designed to address those requirements. Furthermore, the *Te Ika a Maui* OIESS has a capability to record and replay its operations and the conference participants of 2020 can now witness two recent events.

² Given at the California Science Center, Los Angeles, California on January 31, 1998.

³ Goodchild was Chair of the National Research Council's Mapping Science Committee from 1997 to 1999.

⁴ These 'requirements' are enduring. These headings were the themes of the Mapping Sciences '96 Conference of the Mapping Sciences Institute, Australia, Canberra, September, 1996.

1. Last month Air New Zealand flew its latest 'green flight'; twelve years after its first **ASPIRE-GREEN** flight⁵. The flight was to fly from Los Angeles direct to Brisbane. Because of the airline's ability to adopt new air traffic management and navigation concepts it has become internationally competitive on 'long-haul' flights. Unexpectedly, despite many measures taken over the preceding two decades, a terrorist has taken control of the cockpit - renewing memories of the New York 911 incident. Instantly, OIESS was notified. The safety and security status of the incident were such that key National authorities were notified immediately and the aircraft was configured to its autonomous mode. The plane's future navigation system automatically calculated a new flight path to Auckland's secure runway and flew itself and landed gently with precision.
2. After a long period of dormancy the Egmont Volcano has erupted as an explosive ash eruption. Fortunately the eruption has not been 'major' but it has put the OIESS on alert as the event has caused a number of serious bush fires. The rainforest property of Pukeiti, located between Mount Taranaki (Egmont) and Plymouth has been isolated by burning trees crossing roads providing access from the north and the south. The lives of staff and visitors to the internationally recognized garden and plant collection were in danger. OIESS dispatched its new autonomous fire vehicle to clear roads and prepare an exit route.



Figure 1: An illustration of Te Ika a Maui Oceania Infrastructure and Environmental Support System⁶

⁵ Air New Zealand's first ASPIRE (Asia & South Pacific Initiative to Reduce Emissions) flight was from Auckland to San Francisco, departing Auckland 19.30, Friday 12 September 2008. This flight was the first of several demonstration flights undertaken by the ASPIRE partners to demonstrate the capability of air navigation services and airline fuel optimization initiatives to reduce aviation emissions in the Asia and South Pacific region. The flight departed New Zealand's domestic airspace, entering New Zealand's Oceanic airspace at 200 nautical miles out from Auckland. The cruise phase of flight took the Air New Zealand B777 aircraft North-East across the Flight Information Region (FIR) administered by Airways New Zealand, across the international dateline, into the Oakland FIR administered by the Federal Aviation Administration from Fremont, California by Oakland Air Route Traffic Control Center. Towards the end of the flight the B777 entered domestic United States of America airspace approximately 225 miles offshore before commencing its descent for landing and arriving at San Francisco Airport. <http://www.aspire-green.com/>

⁶ Illustration by Wendy Gorton for *Dr Bob* in 2004

So, briefly, what are some of the capabilities from a spatial perspective?

Future air navigation and security. Over the past two decades there has been a number of navigation and surveillance initiatives including:

- FANS (Future Air Navigation System). Air Traffic Control's ability to monitor aircraft was being rapidly outpaced by the growth of flight as a mode of travel. So, in an effort to improve aviation communication, navigation, surveillance, and air traffic management a new integrated system were created. FANS first initiated in 1983.
 - Air New Zealand approached the Boeing Company in 1993 and requested that Boeing support the development of a FANS capability for the 747-400 airplane.
- Automatic Dependent Surveillance-Broadcast (ADS-B). ADS-B will be replacing radar as the primary surveillance method for controlling aircraft worldwide. The ADS-B system can provide traffic and government generated graphical weather information. ADS-B enhances safety by making an aircraft visible to ATC and to other appropriately equipped ADS-B aircraft. ADS-B data can be recorded and downloaded for post flight analysis. ADS-B provides the data infrastructure for inexpensive flight tracking, planning and dispatch.
 - In 2002, Australia became the first country with full, nationwide ADS-B coverage, though only above Flight Level 3000FT (FL300).
- Local Area Augmentation System (LAAS). LAAS is an all-weather aircraft landing system based on real-time differential correction of the GPS signal. Local reference receivers located around the airport send data to a central location at the airport. This data is used to formulate a correction message.
- The Amazon Surveillance System or SIVAM (Portuguese: *Sistema de Vigilância da Amazônia*), is a complex surveillance system used for monitoring the "legal Amazon area". SIVAM assists in curbing the trafficking of illegal narcotics and in curbing illegal logging or burning of the forest. The system utilizes a mixture of fixed and mobile ground radar, as well as airborne surveillance.
- The Aeronautical Information Exchange Model (AIXM). AIXM is designed to enable the management and distribution of Aeronautical Information Services (AIS) data in digital format. AIXM is based on Geography Markup Language (GML) and is one of the GML Application Schemas which is applicable for the Aeronautical domain. One main goal of AIXM is to enable 'digital NOTAM'. In this concept, the traditional free text information contained in NOTAM messages is replaced with structured information, which is suitable for automated computer processing.
 - It is being developed by the US Federal Aviation Administration (FAA) and the European Organisation for the Safety of Air Navigation (EUROCONTROL).

Combined, these established initiatives provide insights on how to address the management of air infrastructure from both spatial and temporal perspectives; on how to describe and represent four dimensional information types; and on how to integrate different geographic information types into command and decision support and navigation systems. OIESS has adopted features described above.

Of course, air navigation and safety is only a part of the overall scope of the OIESS activities. The challenge to address far more ranging topics is to examine abilities to predict, respond, recover and progress from the effects of tragic events. Examples include Iceland's Eyjafjallajökull and Chile's Puyehue volcanic eruptions; Haiti's and Christchurch's earthquakes; Queensland's, New South Wales and Victoria's floods; Western Australia's fires; Japan's earthquakes and tsunami along with Poland's Tu-154 air crash and the Shen Neng 1 grounding on the Great Barrier Reef.

Every feature and every event are spatially and temporally related and if location and relationships are not understood or known then visions and goals won't be achieved.

All features and activities form components of the domain of geographic features. The challenges (though this past decade) have been to model geographic information and to develop processes focusing on this vision.

3 A FEASIBILITY STUDY

– AN ENVIRONMENTAL AND GEOGRAPHIC INFORMATION INFRASTRUCTURE

Geographic information is characterised by its fundamental spatial and temporal nature. Geographic information includes the description of geographical features, infrastructure and environmental aspects of our world. These features include political and administrative entities; population, demography and urban areas; the road, rail, air, sea infrastructure; facilities, communications and resources; physiographic, hydrographic, oceanographic, and meteorological features and state; and the natural and cultivated landscape (Figure 2). The description then involves the modelling, representation and location of these features to place them into a reference system thereby enabling us to *better understand our region and the world*.



Figure 2: The domain of geographic features

The first step leading to today's *Te Ika a Maui* OIESS occurred in early 2012 when a group of interested professionals from New Zealand and Australia gathered at a workshop in Wellington with the goal of producing a **Strategic Plan for the Future Infrastructure and Environmental Information Infrastructure for Oceania**. The group concluded that a better way to represent the domain of geographic features was essential.

So, in addition to addressing future emergency management capabilities for existing incidents such as fires, floods, cyclones, and so on, the workshop agreed to scope future concepts (beyond ten years), as well as agreeing to expand the scope of threats to include natural disasters, pandemic disease, illegal immigration, illegal fishing, people smuggling, environmental degradation, narcotics and transnational crime.

The workshop agreed that the **Environmental and Geographic Information Infrastructure** is an enabling infrastructure that meets the needs of a community for geographic information. The groups agreed that infrastructure is the collection of people, doctrine, policies, architecture, standards, and technologies necessary to create, maintain, and utilise a shared geographic framework.

The infrastructure:

- Establishes a framework for acquiring, producing, managing, and disseminating geographic information;
- Provides the supporting services needed to ensure information content meets future application needs, is easily accessible, and can readily be applied to support operational information requirements, including autonomous operations; and
- Ensures the supporting infrastructure components (including policy, doctrine, training, and organisation) are in place to optimise the use of the geographic information, products, and services provided.

During the scoping activities the Action Working Group determined that it would be appropriate to implement the overall concept in two ways:

1. Design and develop an Oceania Infrastructure and Environmental Support System (OIESS). This identifies that the focus of the area of coverage would be the Oceania region⁷. This identifies that Oceania infrastructure means air, maritime and land transportation and communication systems associated with the future concepts (above). This means that the Oceania environment includes the lands, oceans, atmosphere, weather and climate, and other natural phenomena such as tectonic plate movements, earthquakes, volcanoes, and so on.
2. Design and develop Aotearoa Experimental Prototype Community of Today and Tomorrow (AEPOTAT). This development involved the construction of a multifunctional *City of the Future*⁸ which would present new ideas for consideration by various industry groups. The prototype community would serve as a centre for cultural and technological change in the Oceania region.

4 CAPABILITY DEVELOPMENT - OIESS

The Action Working Group noted that the implementation process needed to acknowledge some key components:

1. **Capability.** Capability was defined as the **capacity** to achieve a specific role or function; in a nominated operational area; within a specified degree of notice; and to maintain that effect for a given period of time.
2. **Command support infrastructure.** OIESS was defined as a command support system managing surveillance, intelligence, and resources assets as shown in Figure 3.
3. **Geographic (or Geospatial) Information Infrastructure** (described above). The overarching OIESS's Geospatial Infrastructure Infrastructure is shown diagrammatically in Figure 4. The OIESS forms an application module of the infrastructure.

As the OIESS is an operational command support system there is reliance on cooperation and collaboration with a diversity of agencies and organizations. Therefore, a key to the success of the OIESS was formal policy and associated agreements and understandings. This endorsement came in 2016 when Prime Minister 'Ritchie' addressed a SSSI forum at Wellington.



**PM 'RITCHIE' AT THE
REDEVELOPED JADE
STADIUM 2016**

The Hon PM Ritchie stated that his government has "acknowledged that a geographic information infrastructure provides information about the world and is vital in supporting the development of the nation and the region". PM Ritchie went on to say the "Te Ika a Maui Oceania Infrastructure and Environmental Support System would lead the world in sophisticated monitoring, strategic and operational planning, commanding and controlling a range of manned and autonomous vehicles and assets, and providing intelligence and informed advice to key authorities."

WELLINGTON JULY 2016

⁷ In this concept Oceania refers to the South West Pacific and its island, New Zealand, Australia and Papua New Guinea.

⁸ A Multi-Function Polis (MFP) was first proposed by the Japanese Ministry of International Trade and Industry in 1987. The intention was to design and build a City of the Future north of Adelaide, South Australia. After a decade of debate its official demise by the South Australian Government occurred in August 1997.

Formal policy and doctrine⁹, treaties, agreements, memorandums of understanding, etc lend themselves to being reported on and presented using innovative techniques. The *Te Ika a Maui* Oceania Infrastructure and Environmental Support System became the first system of its type in the world to use an avatar¹⁰, known as 'Dr Bob', for this task (shown on the right side of Figure 1).

During the past half decade, OIESS sponsored and supported a range of technological advances in other agencies. For example, the New Zealand Fire Service acquired an autonomous vehicle¹¹ capable of clearing burning debris from roads using a precise road three dimensional road network in its database plus a configuration of a range of sensors. This development required that Land Information New Zealand (LINZ) re-digitise road networks in areas prone to bushfires using precision stereophotogrammetric techniques.



Figure 3: Infrastructure and Environmental Support Infrastructure

This past decade has been amazing - virtual advisers; autonomous vehicles on the ground, in the air, on and under the water; four dimensional models of the natural environment; and so on. This situation, however, could not be achieved without similar advances in the education sector. The traditional scientific disciplines of geodesy and cartography have been revitalized and have now become the key areas of study. In addition, local universities have initiated new multi-

⁹ Authoritative implementation of policy

¹⁰ Australia's Defence Science and Technology Organisation (DSTO) developed a Future Operations Centre Analysis Laboratory (FOCAL) in 2005. FOCAL used 'virtual advisers'
http://www.dsto.defence.gov.au/publications/3654/5959FOCAL_Fact_Sheet.pdf

¹¹ The US DARPA (Defense Advanced Research Projects Agency) conducted a series of 'challenges' in 2004, 2005 and 2007 with the aim of developing autonomous vehicles
http://en.wikipedia.org/wiki/DARPA_Grand_Challenge

discipline courses such as occurred in the 1970s and 1980s when, at the Canberra College of Advanced Education for example, computer science (including artificial intelligence) were coupled with cartography courses.

Such has been the success that the next phase – the **2030 Virtual World** – has already been under discussion (Figure 5)!

By 2030 we should have reached an era where we have truly intelligent systems. The basic dictionary meaning of the word intelligence refers to having knowledge, understanding and awareness. This suggests that systems may have the ability to learn or understand from experience. They might have the ability to acquire and retain knowledge. They should have the ability to respond quickly and successfully to a new situation.

Spatial and temporal reasoning will be based on knowledge of the full spectrum of environmental and infrastructure information at various levels of resolution and relevance. Geospatial information will include administrative and legislative information, demographic information, the fundamental infrastructure (ports and harbours, road and transportation systems and the airspace), assets (facilities, industry, and commerce), natural resources, and the environment including the atmosphere and oceans.

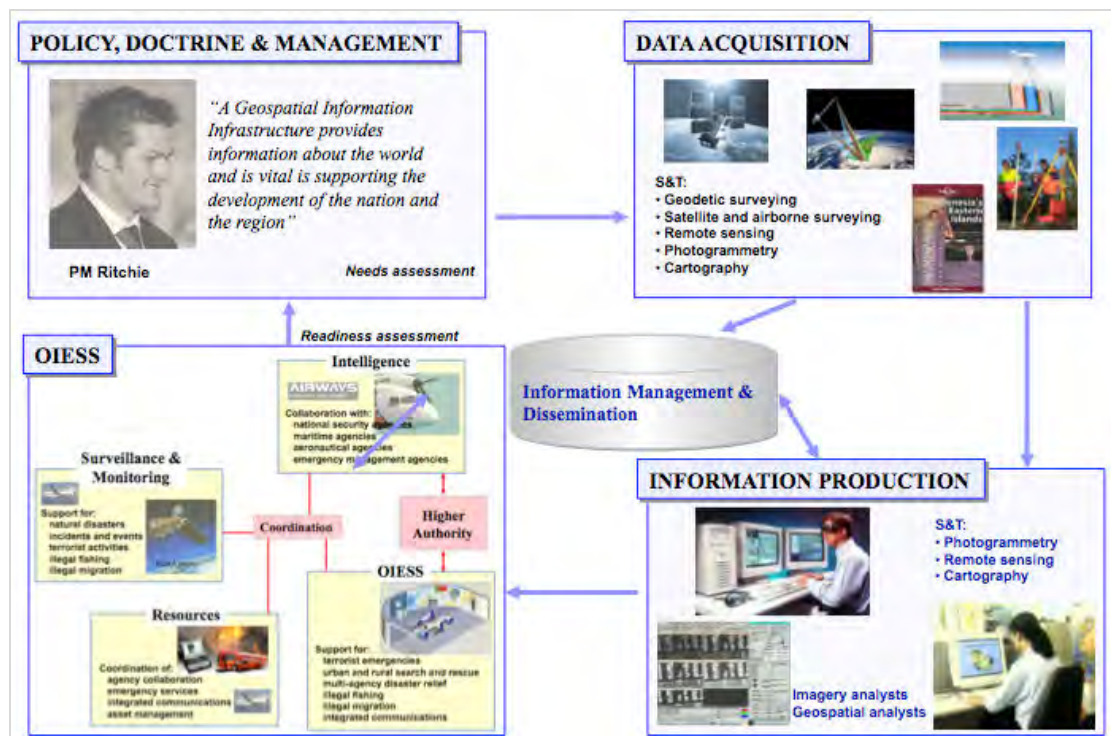


Figure 4: Geographic (or Geospatial) Information Infrastructure

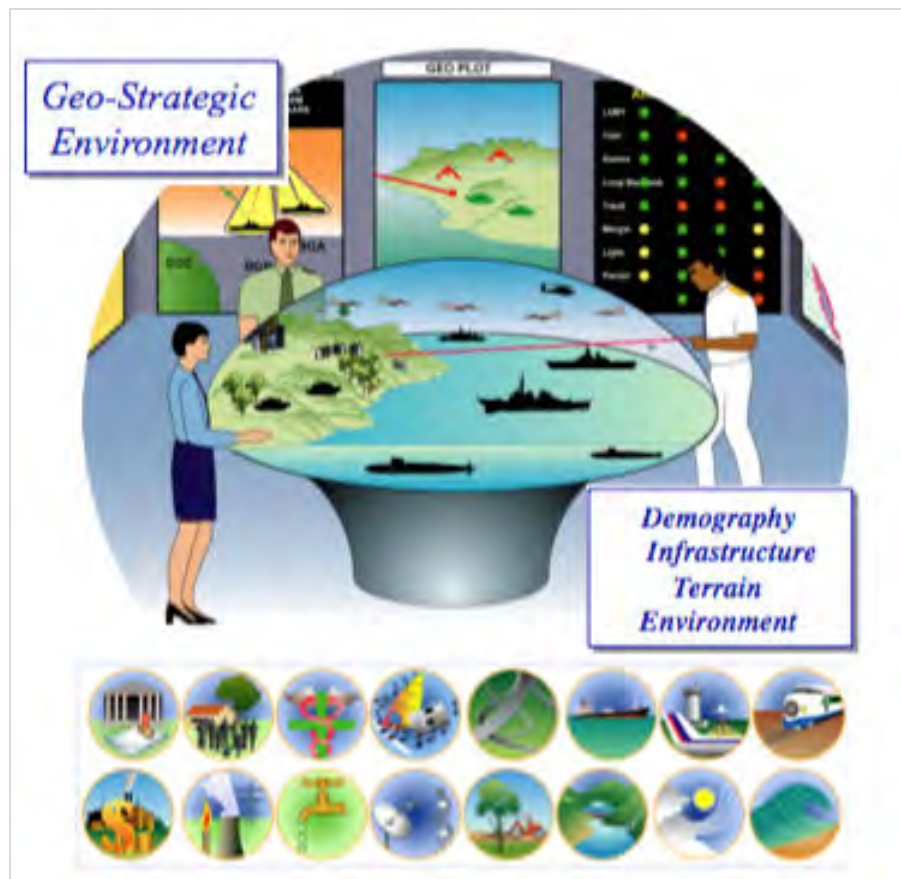


Figure 5: 2030 Virtual World – A National Security / Defence profile

There is probably little doubt that 2030 systems will include immersion characteristics whereby users are surrounded by virtual reality. 2030 systems will probably be scenario-driven with users being given real-time options along with costs and risks.

An exciting new era in geodesy and cartography has begun!

5 EPILOGUE

Is OIESS possible? Capabilities for *Te Ika A Maui* OIESS 2020 already exist either as proven capabilities or as capabilities that have been subject to testing as for autonomous vehicles by US DARPA.

So, it is not technology that is the problem but it will be the administrative issues that continue to be difficult to overcome.

I personally believe that the 2020 vision is achievable. I have been on an **odyssey** for the past four decades and summarize my contribution below.

Dr Bob's ODYSSEY¹²

- 1976 Poster display **AUTOMAP – Some applications of the data base**, Second Australian Cartographic Conference 1976 Adelaide
- 1980 “An overview of a cartographic mapping package”, **Cartography**, Volume 11 No. 3, The Australian Institute of Cartographers
- 1980 “Automated cartography with navigational applications”, **Proceedings – 4th Australian Cartographic Conference**, The Australian Institute of Cartographers, Hobart
- 1984 “Who or what is DES?”, **Cartography**, Volume 13 No. 3, The Australian Institute of Cartographers
- 1985 “Enquiry systems for the interrogation of infrastructure”, **Proceedings – Auto-Carto 7**, American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, Falls Church, VA
- 1985 **Enquiry systems for the interrogation of infrastructure in areas of large geographic extent**, Unpublished Masters Thesis, University of Wisconsin – Madison, WI
- 1986 “Automated cartography: the next development”, **Proceedings – 6th Australian Cartographic Conference**, The Australian Institute of Cartographers, Melbourne
- 1987 “Evolution in cartography: data intelligence”, **Cartography**, Vol 16 No. 2, The Australian Institute of Cartographers
- 1988 “Analysis of the road transportation network”, **Proceedings – 7th Australian Cartographic Conference**, Volume 2, The Australian Institute of Cartographers, Sydney
- 1989 “Geographic information: aspects of phenomenology and cognition”, **Proceedings – Auto-Carto 9**, American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, Falls Church, VA
- 1989 **Analysis of Geographic Information: A Cognitive Approach**, A thesis submitted for the degree of Doctor of Philosophy, The University of New South Wales, Australian Defence Force Academy, Canberra

¹² I conducted a workshop on GIS and Automated Cartography at the Australian Army's School of Military Survey in 1991. During my talk I was talking about spatial relations and noted that “*a tessellation should be infinitely recursively decomposable into similar patterns of smaller size*”. At that point a Sergeant stood up and said ‘**Doctor Bob!**’ – said with the inflection by Fozzie Bear on *The Muppets* TV show when addressing Dr Bob in the medical parody skit “Veterinarians’ Hospital”. The nickname seems to have stuck and old associates still call me Dr Bob.

- 1990 “Phenomenological structuring of geographic information”, **Technical papers – 8th Australian Cartographic Conference**, The Australian Institute of Cartographers, Darwin

- 1990 “Innovation in cartographic communication”, **Technical papers – 8th Australian Cartographic Conference**, The Australian Institute of Cartographers, Darwin

- 1993 “Digital geographic data exchange standards and products: Descriptions, comparisons and opinions”, **Cartography**, Volume 22 No. 1, The Australian Institute of Cartographers

- 1994 “DGIWG and DIGEST”, **Technical papers of Mapping Sciences ’94**, The Australian Institute of Cartographers, Gold Coast

- 1996 “Cartography: A new era has begun”, **Cartography**, Volume 25 No. 1, The Mapping Science Institute, Australia

- 1996 “Cartography and cartographers enter a new era”, **Technical papers of Mapping Sciences ’96**, The Mapping Sciences Institute Australia, Canberra

- 1996 “EGICS: A Defence Environment and Geographic Information Capability Study”, **Technical papers of Mapping Sciences ’96**, The Mapping Sciences Institute Australia, Canberra

- 2000 “42: *Videre Parare Est*”, **Keynote Address Beyond 2000**, Conference proceedings Mapping Science Institute Australia, Sydney

- 2002 “Geographic Intelligence: The key to information superiority”, Presented at Joint Conference of AURISA and Institution of Surveyors Australia, Adelaide

- 2005 “Geospatial Intelligence and a Geospatial Information Infrastructure: Contemporary Defence and Security Challenges”, **Proceedings of SSC2005: Spatial intelligence, innovation and praxis**, Spatial Sciences Institute Australia, Melbourne

- 2006 “Geospatial Intelligence: Know your world through story telling”, **Keynote Address Geo Cart 2006**, National Cartographic Conference, Auckland

- 2007 “Geospatial Intelligence: Know your world through story telling”, **Proceedings of SSC2007**, Spatial Sciences Institute Australia, Hobart

- 2007 “Where on earth are you? The incident on the Shatt al Arab waterway and other border incidents”, Presented at Australian Command and Staff College, Canberra

The summary above shows a timeline of initiatives that are no closer to be implemented than they were when first presented. So, why is this so?

The birth of Australia is often referred to as 'federation'. This is because the Constitution created a 'federal' system of government. Under a federal system, powers are divided between a central government and several regional governments. In Australia, power was divided between the then Commonwealth Government and the governments of the six colonies, which were renamed 'states' by the Constitution.

Surveying and mapping responsibilities were split between the Commonwealth and the States and although there have been initiatives, such as ANZLIC, the reality is that capability development as I have described are difficult to achieve.

The challenge is to move beyond these limitations to develop a truly advanced system in a specialized experimental facility; a system that can provide effective decision support and operational capabilities needed by contemporary incidents and events related to the management of our nations; their environments; their infrastructures; and the security of their peoples.

So, *Te Ika A Maui* OIESS 2020 may well be the first truly advanced system developed to address contemporary events of the 21st century!

Using Volunteered Information to Map the Queensland Floods

Kevin McDougall

Faculty of Engineering and Surveying, University of Southern Queensland
West Street, Toowoomba, Queensland, 4350, Australia.
mcdougak@usq.edu.au

ABSTRACT

The recent flood events and tropical cyclones in Queensland have dramatically impacted on people's lives across the State. Damage in excess of five billion dollars has been reported and the cost to the economy continues to rise. Although most of the flooding has receded, the clean-up and re-building will continue for years. The Queensland floods were characterised by the unprecedented use of social media to report events as they happened and was used very effectively by the main stream media. Social media networks such as Twitter and Facebook, not only informed people of the events as they unfolded, they have now also provided a historical archive for use in future planning and mapping. Although the Commonwealth and State governments and the private sector did a magnificent job in mapping the flood events where possible, a number of gaps still exist. This paper discusses the use of volunteered geographic information such as photographs and videos to assist in mapping the flood extents in regions where there was little or no mapping available. Through the integration of volunteered information with existing geographic information, hydrological data and local knowledge, flood extents can be re-constructed and hence mapped.

KEYWORDS: Spatial Information, Mapping, Floods, Volunteered Geographic Information

1 INTRODUCTION

During December 2010 to February 2011, the State of Queensland experienced a series of damaging floods which caused billions of dollars in damage and the loss of over 20 lives. Major flooding was experienced at over 30 cities, towns and rural communities over southern and western Queensland including significant inundation of agricultural crops and mining communities. Consistent rain during the Australian spring resulted in many of the large catchments becoming heavily saturated and the larger storage reservoirs and dams reaching capacity. These conditions were further exacerbated by the presence of a number of tropical cyclones which in addition to heavy rainfall result in significant property and landscape damage due to cyclonic winds.

As the varying flood events unfolded social media and crowd sourced geographic information played an important role in keeping people informed, especially as official channels of communication began to fail or were placed under extreme load. The government's management of the Queensland floods and especially the role of the community in their assistance were widely applauded. Information and communication technologies played a critical role during the disaster and its management via the conventional communication channels such as radio, television and newspapers but also through third party social media networks such as Twitter and Facebook. People who had never signed up to Facebook and Twitter started doing so and the Australian Broadcasting Commission (ABC) radio launched a link to crowd sourced reports on flooding.

This paper reviews the various forms of volunteered and shared information that occurred throughout the Queensland floods and their impacts. The potential of volunteered geographic

information for post-disaster assessment including damage assessment, planning and official flood lines is examined.

2 FLOODS AND FLOOD MAPPING

Most of the major floods in Queensland occur in summer or early autumn due to tropical cyclones or intense monsoonal depressions. These systems are capable of producing excessive quantities of rainfall in short periods of time. For example at Bellenden Ker in North Queensland, tropical cyclone “Peter” caused 1,947mm of rain in January 1979 during a 48 hour period. In 1999 cyclone “Rona” produced 1,870mm in 48 hours at the same location (Bureau of Meteorology 2011a).

Prior to 1860 three major floods were reported for the Brisbane/Ipswich regions, with the January 1841 flood having the highest recorded level of 8.43m at Brisbane (Bureau of Meteorology 2011a). A further five major floods inundated Brisbane and Ipswich between 1885 and 1900. In this time period, the Brisbane River peaked at 8.3m, and the Bremer River at 24.5m – its highest recorded level (Centre for the Government of Queensland 2010). The Bremer River experienced an additional nine major floods between 1900 and 1972. It was not until 1974 however, that Brisbane and Bremer Rivers flooded to 5.45m and 20.7m respectively – the highest levels since 1893 (Bureau of Meteorology 2010).

The Queensland flood warning network derives its data from a series of rainfall and river height stations (Bureau of Meteorology 2011b). There are two types of rainfall station in use by the Bureau of Meteorology (BoM): Floodwarn and Daily Reporting. The Floodwarn rainfall stations are designed specifically for flood warning purposes. They are either manual or automatic, and report every 25 or 50mm, and every 1mm of rainfall respectively. Daily reporting rainfall stations consist of manual and automatic stations that report the rainfall received in a 24 hour period to 9am each day (Bureau of Meteorology 2011b). The Floodwarn river height stations have both manual and automatic varieties which report river levels whenever the water reaches a threshold height, and at regular intervals thereafter. The Bureau of Meteorology’s Flood Warning Centre receives the data provided by these stations, and uses it in hydraulic models to produce river height predictions. In the event of an expected flood, the Flood Warning Centre issues warnings to radio stations, Councils, emergency services and various other agencies involved in flood response activities.

There has been a significant amount of research done towards the creation of flood models, and associated topics. Much of the work between 1999 and 2005 focused on creating models that were tested in rural areas (Bates & De Roo 2000, Bates & Horritt 2001, Ervine & Macleod 1999). A number of these models were later utilised to predict flood inundation levels in urban areas (Bates & De Roo 2000, Yu & Lane 2006). A 1D model measures flood levels in the channel, whereas a 2D model measures flood depth for the extent of the floodplain.

A limit to raster-based flood models is the resolution of cells used in the model – if they are too small the computational requirements became restrictive (Haider *et al.* 2003). Yu and Lane (2006) investigated the effect of model cell size for models applied to urban areas, and concluded that even small variations in model resolution have significant effects on inundation extent. Accordingly, as processing power increases, using progressively smaller cell sizes will be a viable option. The accuracy of any flood model is dependent on the range of input data and the closeness of the model to the true behaviour of the flood water.

Mapping of the actual flood extents is often the best method of calibration of hydraulic models and allows models to be improved for future predictive purposes. The primary goal of flood mapping is to identify areas that are flooded or not flooded. This process consists of two steps – (1) determining wet/dry areas before and during a flood event, and (2) comparing these areas to determine which areas were flooded.

Three main data sources are used to map flood extents: optical data, radar data, and topographic and river gauge data (Wang 2002). Optical data include aerial photographs and satellite data such as from a Landsat Thematic Mapper (TM) sensor. With different reflectance responses of dry and wet/water surfaces, aerial photographs and TM data can easily distinguish between surfaces. Wang (2002) also concluded that using TM data for flood extent mapping is: 1)

Reliable and accurate; 2) Simply applied: georeference two TM images, identify wet/dry areas, and compare before/after imagery; and 3) Efficient and cost-effective.

However, there are limitations to using many satellite sensors. As satellites have a fixed orbit pattern, their revisit time (the time taken between the subsequent collection of data from the same location) may mean data is collected long after a flood has receded. The limited spatial resolution of satellite data may also be too coarse for identifying small flooded areas, particularly in vegetated, commercial, or residential areas. Additionally, the many sensors do not penetrate vegetation well, so flooding may not be reliably detected under the canopy (Wang 2002). Finally, both satellite imagery and aerial photography should be collected during the day and will not penetrate cloud cover.

The same basic principle to determine flood extent i.e. detection and comparison of wet/dry surfaces before and during a flood applies also to extent mapping when using radar data. The key advantage in using synthetic aperture radar (SAR) data over optical data is the ability of radar microwave to penetrate cloud cover and forest canopies (depending on wavelength). Because current SAR sensors are satellite-mounted, this system suffers from the same revisit time limitation as optical sensors.

Finally, using topography DEMs and river gauge data is perhaps the simplest of the three methods. It involves getting river levels before a flood, and during its peak for each gauge, and then flooding a DEM – once with the pre-flood levels, and once with peak levels. The inundated areas can then be compared to determine existing bodies of water, and flood extent (Wang 2002). Advantages of using this data include: 1) Data is reliable and accurate, 2) Methodology is simple, efficient, and economical, and 3) The data is easily updated. Its limitations include: only being able to map areas that have flood gauges, and it is sensitive to the accuracy of the input DEM.

This paper examines some of the information that was volunteered by citizens during the flood events in Queensland as part of the social networking and media activities. The imagery taken during the events was geocoded and used to determine the possible extents of the flooding by generating a flood level DEM. The benefits of volunteered information and the utility of these new data sources are discussed.

3 VOLUNTEERED INFORMATION/SOCIAL NETWORKING DURING THE FLOODS

Goodchild (2007) defines volunteered geographic information (VGI) as spatial information collected voluntarily by private citizens. Geo-tagged images submitted by individuals to the web may therefore be considered VGI. Goodchild outlines some popular examples of VGI, including: Wikimapia <<http://wikimapia.org>>, Flickr <<http://www.flickr.com/>>, and OpenStreetMap <<http://www.openstreetmap.org/>>. Wikimapia lets anyone with an internet connection select an area of the Earth's surface, and provide it with a description. Flickr allows users to upload photos and tag them with a latitude and longitude. OpenStreetMap is 'an editable map of the whole world, which is being built largely from scratch, and released with an open content license' (Openstreetmap 2011).

Social networking also played a major role in keeping people informed during the January 2011 flood. Ushahidi is a non-profit technology company that specialises in developing free and open source software for information collection, visualisation and interactive mapping (Ushahidi 2011). Crowdmap is an online interactive mapping service, based on the Ushahidi platform (Crowdmap 2011). It offers the ability to collect information from cell phones, email and the web, aggregate that information into a single platform, and visualise it on a map and timeline. The Australian Broadcasting Corporation launched QLD FLOOD CRISIS MAP – a crowdmap of the Queensland floods in January 2011 (ABC 2011). This crowdmap allowed individuals to send flood-related information via email, text message, Twitter, or via the website itself (Australian Broadcasting Commission 2011). This information was then available to anyone with an internet connection. The Courier Mail also provided a similar service, though only allowing people to submit photos, via email (Courier Mail 2011).

The social networking service Twitter <www.twitter.com> allowed people to post and receive short text based updates about the flood in real time. Photos and videos were also able to be attached to these updates. Similarly, the website Facebook <www.facebook.com> allowed groups

such as the Queensland Police Service to provide flood information updates to anyone who browsed to their Facebook page. Finally, YouTube <www.youtube.com> provided a forum for people to connect and inform through the use of user-generated and contributed videos.

At the peak of the Queensland floods there were between fourteen and sixteen thousand tweets per hour on the 'qldfloods hashtag' which was used to coordinate the conversation around the flood event itself. These peaked at around the time Brisbane and the surrounding areas began to become inundated. Agencies and organisations alongside members of the community began using the Twitter platform as a place to distribute 'raw' footage and information, but then began to trust and 'follow' particular accounts. Some of the most dramatic flooding occurred in Toowoomba and the Lockyer Valley on 10 January 2011 during a flash flood event that claimed a number of lives. The flood waters from the Brisbane catchment moved progressively towards the coast and the cities of Ipswich and Brisbane which peaked around two to three days after the flash flood events.

The response to the Queensland floods by both the full levels of government and the community were widely applauded and recognized as being above and beyond their respective call of duty. State and local government staff in particular worked long hours under difficult conditions to, firstly, meet the critical emergency response needs, and then to provide critical information to enable the re-building exercises to get underway. Information and mapping on the extents of the various floods across Queensland have been pivotal in prioritizing resources, distributing emergency relief and clarifying the inevitable insurance issues.

Under international disaster agreements, the Australian and Queensland governments were able to access a variety of mapping resources including satellite imagery during and after the floods. This information was utilized together with high resolution imagery from providers such as Nearmap to rapidly generate flood extent maps. Nearmap, in particular, flew missions at times near the peak of the floods in Brisbane and other regional areas to produce very high resolution imagery of the actual floods. This imagery was available to the public with hours of the mission and being used by the community to assist in the emergency efforts (Figure 1).

The Queensland Government utilized the Nearmap imagery to begin the process of mapping the flood extents and making them available. Under International Emergency Agreements the Queensland Government also had access to a range of other data including satellite imagery from commercial and government agencies around the world. A special agency called the Queensland Reconstruction Authority launched an interactive map (Figure 2) which detailed the areas which were flooded or inundated. This was a valuable source of information for individuals, community organizations, governments and private sector organizations such as insurance firms.

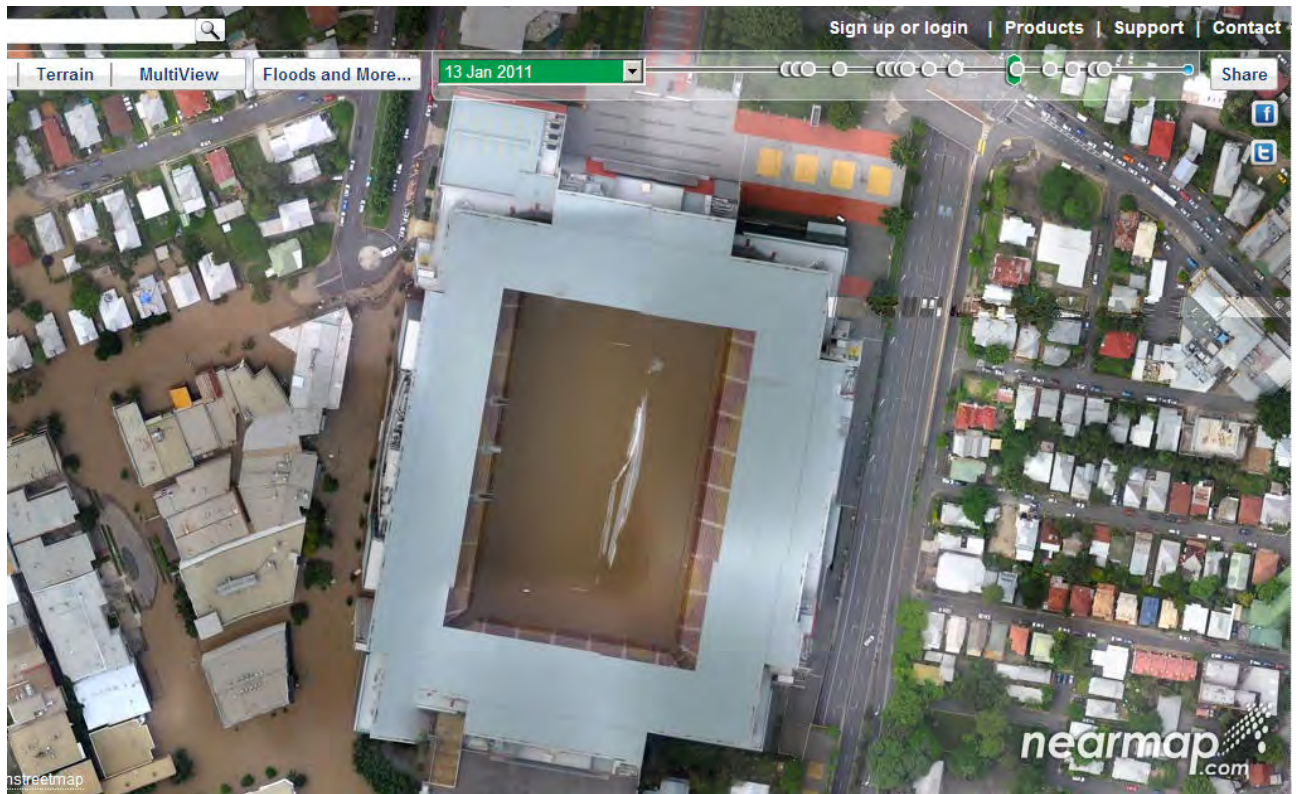


Figure 1: Nearmap imagery showing Suncorp Stadium and surrounds (Source: Nearmap 2011)

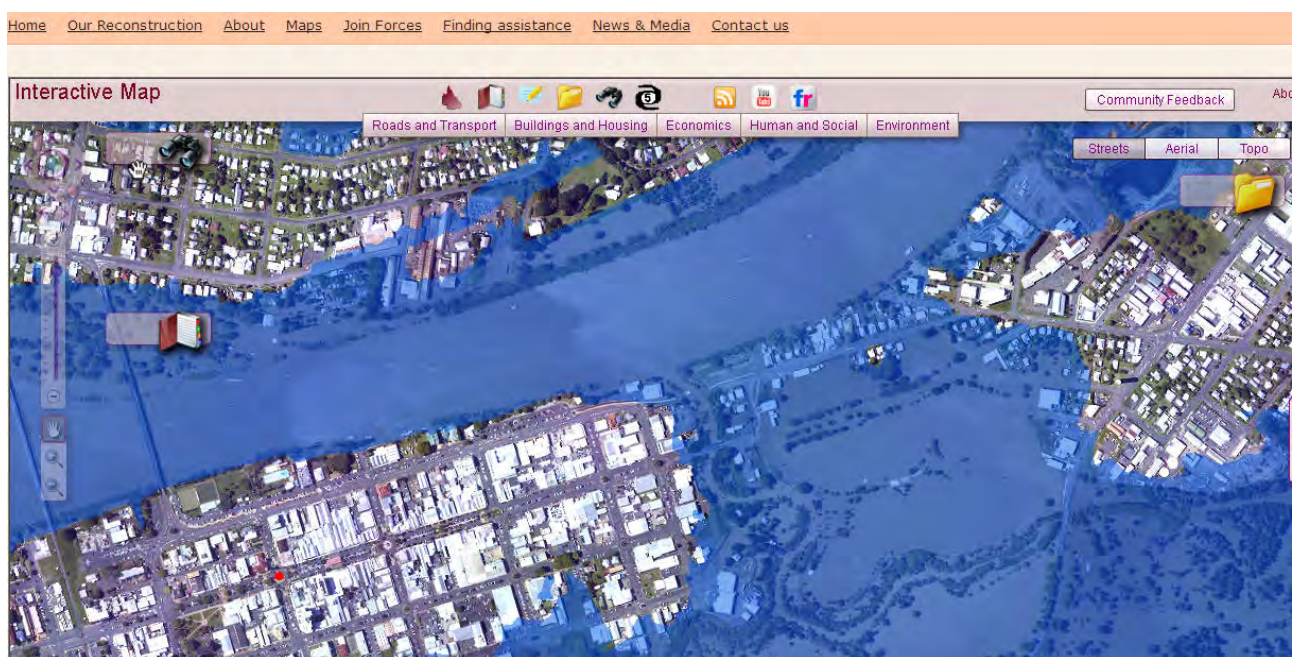


Figure 2: Queensland Reconstruction Authority Interactive Map (Queensland Government 2011)

4 USING VGI TO MAP THE FLOOD EXTENTS

The Australian Broadcasting Corporation launched QLD FLOOD CRISIS MAP – a crowdmap of the Queensland floods based on the Ushahidi platform in January 2011 (Figure 3). This crowdmap allowed individuals to send flood-related information via email, text message, Twitter, or via the website itself. This information was then available to anyone with an internet connection. This service proved to be very popular and the servers struggled at times during the crisis to keep up with the demand.



Figure 3: Ushahidi Crowdmapped the Queensland Floods

Photography and imagery of the floods across different regions were posted on sites such as Flickr which were linked at each location through the crowdmap. Individuals had the opportunity to add comments and additional information regarding the context of the images. The posting time is also time stamped by the system. These images provide an excellent historic and current record of the flood events and features in the imagery can easily be used to reference flood heights at a particular time (Figure 4).

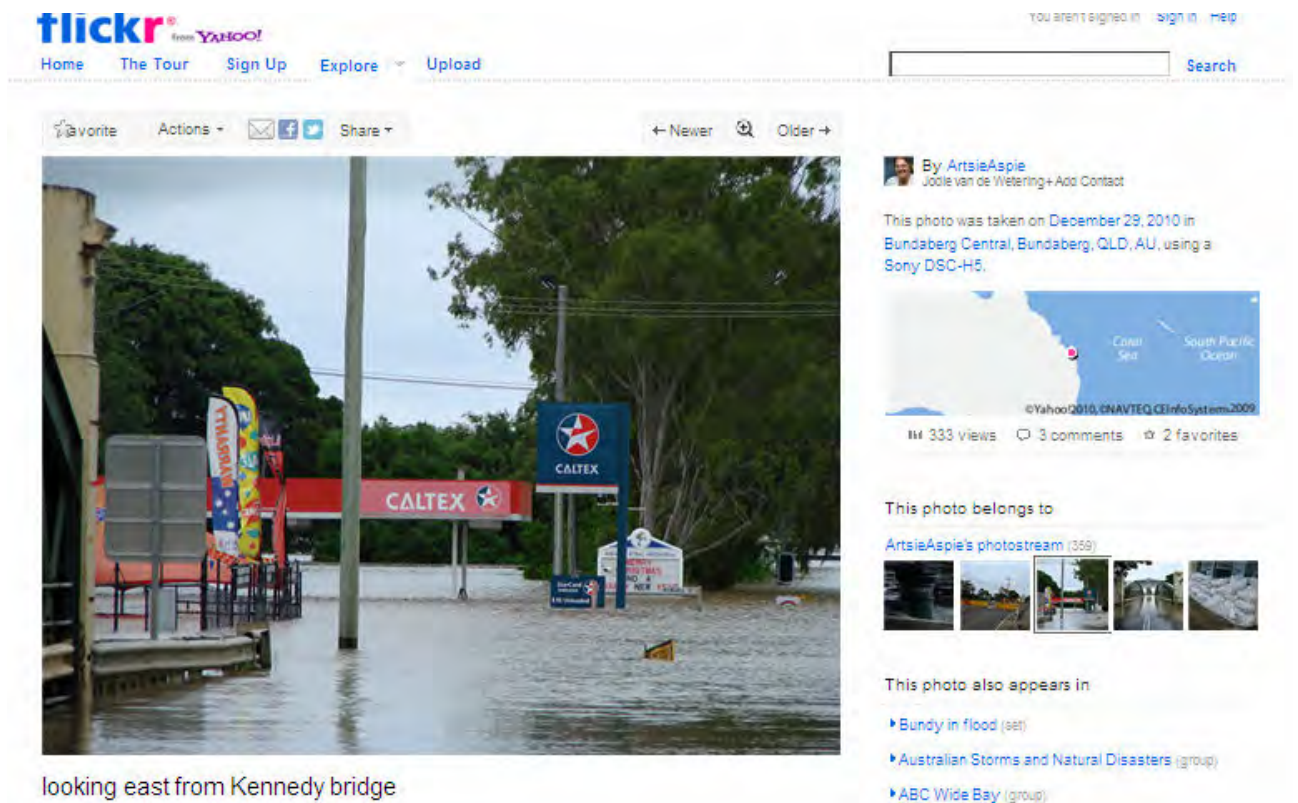


Figure 4: Example of flood photography around Bundaberg in December 2010.

Once individual imagery was collected then flood heights at a particular point could be identified and measured. This required field visits to verify the exact locations and to accurately locate the flood point with respect to horizontal and vertical position. In the case of this study RTK GPS was employed to capture 35 points from a variety of locations around the Brisbane flood areas. This data was then utilised to generate a DEM surface of the flood in the locality which represented the actual flood surface.

The DEM also included a number of actual river gauge stations where they were available to improve the quality of the data along the river. The river gauges are linked to the Australian Height Datum (AHD) and the peak heights can be incorporated into the DEM. Figure 5 provides an example of the plotted gauge data at various stations linked to rainfall.

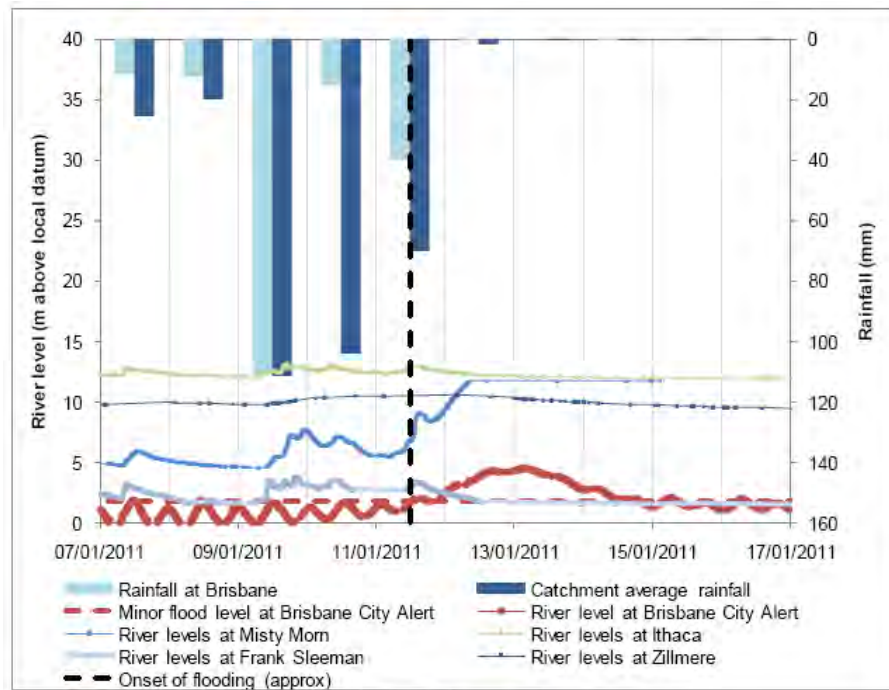


Figure 5: River gauge and rainfall plots along the Brisbane River

Once the DEM of the flood surface has been calculated it can then be intersected with the DEM of the existing or natural surface. The intersection of these two surfaces provides a reasonable estimate of the flood extents over a local area. Figure 6 illustrates the points which were measured from the identified imagery and the final flood surface that was calculated from the two DEMs.



Figure 6: Example of the computed flood extents from volunteered information

5 DISCUSSION

The use of social media provided the opportunity for people from a wide range of varying backgrounds to participate and contribute to the dissemination of information throughout the Queensland flood events. The benefits of the current technology were immediately obvious and within minutes of the event people were sending emails, photos and videos to their friends to update them about the evolving crisis. Twitter and Facebook sites facilitated the wider dissemination of the crisis to others within their network including the mainstream media. Members of the media obtained their early information from the emails, tweets and Facebook postings from friends and colleagues. The media and emergency services quickly identified the power of this resource and began to establish channels to support and build their communication and information collection.

The various channels of communication provided a near real time coverage of the event that was rich in information including continuous commentary, voice, photographs and video. Early in the event the information was accurate and often breath-taking, putting all of us in the position of the observer. However, after the early stages of the event, a number of spurious postings began to appear including duplications of photography and misinformation regarding the flooding events.

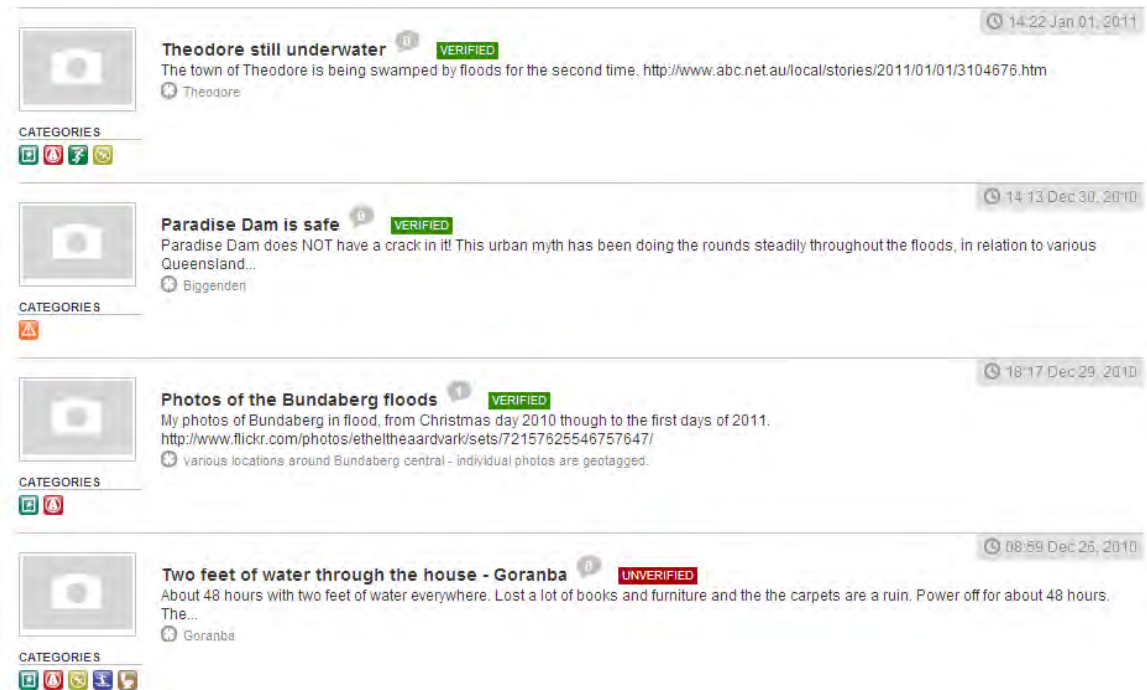


Figure 7: Moderation of the ABC Crowdmap

Verifying information is an important element of crowd mapping if you want the users to continue to have confidence in the information. This has always been one of the key issues with volunteered information and there are a number of mechanisms that can be used to improve the veracity of the information. Trust is an important commodity within these environments and a trusted source, just like a media source, is well respected by the community of users. Just like with Wikipedia, the entries are available for all to scrutinise and to edit, and so it becomes a self moderating community.

In the case of the crowdmap users could vote up and down on the reports as they came through which improved the veracity of the information. Information that was not challenged or came from a trusted sourced was marked as verified whilst new reports or distrusted sources are identified as unverified (Figure 7). However, the community of users must still rely on a degree of common sense with these sites and preferably have some local knowledge to validate the reports.

The techniques for the reconstruction of flood extents are very dependent on the amount and quality of the imagery that is available. Most imagery was taken from various ground based vantage points and provides a good indication of the behaviour of the flood waters at that particular location. In other cases, media reports in the form of television footage may be available to analyse from an aerial vantage point. This provides the opportunity to directly map the flood boundary at a particular point in time using a combination of ground truthing and digitising.

6 CONCLUSIONS

In conclusion, the use of social media has now added another dimension to volunteering information and its value is undeniable in respect to its immediacy and depth of information. The imagery and related data that is available through crowdmaps and sites such as Flickr provide a ready resource to begin mapping flood extents in the absence of other aerial or satellite data. These maps can provide valuable data for the improvement in hydrological modelling of flood events and planning for future emergencies. However, the harnessing of this information including its veracity and validation still remains a challenge for those who need accurate and reliable information.

7 ACKNOWLEDGEMENTS

The author wishes to acknowledge the contribution of Phil Temple-Watts in providing data and examples of outputs.

REFERENCES

- Australian Broadcasting Commission, 2011. ABC QLD Flood Crisis Map, <<http://queenslandfloods.crowdmap.com/reports/>>, (accessed 20 February 2011).
- Bates, P. and De Roo, A., 2000. A simple raster-based model for flood inundation simulation, *Journal of Hydrology*, vol. 236, no. 1-2, pp. 54-77.
- Bates, P. and Horritt, M., 2001. Predicting floodplain inundation: raster-based modelling versus the finite-element approach, *Hydrological Processes*, vol. 15, no. 5, pp. 825-42.
- Bureau of Meteorology, 2010. Flood Warning System for the Bremer River to Ipswich, <http://www.bom.gov.au/hydro/flood/qld/brochures/brisbane_bremer/bremer.shtml>, (accessed 15 August 2011).
- Bureau of Meteorology, 2011a. Queensland Flood Summary Pre-1860, <http://www.bom.gov.au/hydro/flood/qld/fld_history/floodsum_pre_1860.shtml>, (accessed 15 August 2011).
- Bureau of Meteorology, 2011b. Description of Queensland Flood Warning Network Details, <<http://www.bom.gov.au/hydro/flood/qld/networks/description.shtml>>, (accessed 10 August 2011).
- Centre for the Government of Queensland, 2010. Queensland Places, <Centre for the Government of Queensland n.d., Ipswich | Queensland Places, viewed 18/05 2011, <<http://queenslandplaces.com.au/ipswich>>.>, (accessed 15 August 2011).
- Courier Mail, 2011. The Brisbane River Interactive Flood Maps, <<http://brisfloodguide.indev.com.au/>>, (accessed 15 February 2011).
- Crowdmap, 2011. Crowdmap Website, <<http://crowdmap.com/>>, (accessed 15 May 2011).
- Ervine, A. and Macleod, A., 1999. Modelling a river channel with distant floodbanks, *ICE - Water Maritime and Energy*, vol. 136, no. 1, pp. 21-33.
- Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography, *GeoJournal*, no. 69, pp. 211-21.

- Haider, S., Paquier, A., Morel, R. and Champagne, J.-Y., 2003. Urban flood modelling using computational fluid dynamics, *Proceedings of the ICE - Water and Maritime Engineering*, vol. 156, no. 2, pp. 129-35.
- Openstreetmap, 2011. Openstreetmap Website, (accessed 10 August 2011).
- Queensland Government, 2011. Reconstruction Queensland Interactive Map, <<http://www.qldreconstruction.org.au/maps/interactive-map>>, (accessed 15 March 2011).
- Ushahidi, 2011. Ushahidi Mission Website, <<http://www.ushahidi.com/about-us>>, (accessed 20 May 2011).
- Ushahidi Website, 2011. Ushahidi Mission, <<http://www.ushahidi.com/about-us>>, (accessed 20 May 2011).
- Wang, Y., 2002. Mapping Extent of Floods: What We Have Learned and How We Can Do Better, *Natural Hazards Review*, vol. 68, no. 3, pp. 68-73.
- Yu, D. and Lane, S., 2006. Urban fluvial flood modelling using a two-dimensional diffusion-wave treatment, part 1: mesh resolution effects, *Hydrological Processes*, vol. 20, no. 7, pp. 1541-65.

Embedding GIS in the Departments of Health and Human Services

Clare Brazenor and Miles da Costa

Department of Health (Victoria) and Department of Human Services (Victoria)
50 Lonsdale Street, Melbourne, Australia
clare.brazenor@health.vic.gov.au - miles.da_costa@dhs.vic.gov.au

ABSTRACT

The Victorian departments of Health (DH) and Human Services (DHS) have delivered to their organisations an enterprise GIS that provides a platform for spatial information and quality control processes to become embedded within departmental program areas. The system is a shared service between the Departments and has been designed to facilitate the needs of a divergent user base, both in ability and geography whilst maintaining a centralised management structure and data model. It provides a centre for GIS excellence and a single source of truth for spatial information.

Over the past year the GIS team has faced many challenges engaging program areas that have had little exposure to GIS products and capabilities, particularly around data quality, information security and requirements gathering. This paper describes the approach taken by DH/DHS to meet these challenges, outlines the current system and provides an overview on how the GIS presents a spatial view of business information.

KEYWORDS: Data Quality, Enterprise GIS, Web Mapping, Spatial Information.

1 INTRODUCTION

The Department of Human Services (DHS) and the Department of Health (DH) in the State of Victoria, Australia, encompass a broad range of government responsibilities. The DHS is committed to making a positive difference for Victorians most in need through the provision of housing and community services. The main portfolios include:

- (i) child protection - statutory responsibilities such as child protection and youth justice,
- (ii) disability - the provision of services and support for Victorians with intellectual, physical, sensory disabilities and,
- (iii) housing - the provision of public and social housing to improve and extend the range of affordable housing options available in the community.

(Department of Human Services 2011)

The DH is focused on achieving the best health and wellbeing for all Victorians. This is accomplished through planning, policy development, funding and regulation of health service providers and activities which promote and protect Victorians' health. The DH has three main portfolios:

- (i) health - the provision of health care services through the public hospital system, community health services, ambulance services and dental services and health promotion and related preventative services, education and regulation,
- (ii) mental health - the provision of public mental health service system and psychiatric rehabilitation and support services together with a range of alcohol and drug prevention and treatment services and,

- (iii) aged care - the provision of public sector residential and rehabilitation care together with support in the home.

(Department of Health 2011)

The Departments also have key emergency management coordination roles¹ for the State making it critically important to know where all facilities and clients are located from a single source of truth repository and to support whole-of-government data interoperability requirements during an emergency event.

Geographical Information Systems (GIS) has been utilised in the Departments for service planning purposes for many years. The Departments had rich geographical knowledge residing in stand-alone, project-specific textual and spatial systems. This geographical knowledge was not universally shared, nor easily updated or uniformly represented. The stand-alone configurations and work practices only precipitated internal silos of data and work practices and ultimately hindered an integrated response to service delivery and planning (for the 'right services, in the right place at the right time'). To garner high-level Executive support from the Departments, a key GIS Strategy document was developed. This GIS Strategy document identified the shortcomings of the current use of GIS and the opportunities for the Departments with a corporate approach to GIS. The support from management ensured that systems were not duplicated and solutions remained aligned with well established departmental IT policies and priorities.

2 THE CORPORATE GIS ENVIRONMENT

The shortcomings of the stand-alone approach to GIS in the Departments presented problems specifically for the management of software licences. The Department's workforce is located in offices throughout the state which presented logistical challenges. The centralised management and administration of software licences is a core component of the corporate GIS environment, removing the need to configure specific desktops. The removal of data silos and the creation of a central (single point of truth) spatial database also forms the foundation for the upgrade of the corporate GIS environment. The centralised spatial databases allow for spatial data standards (standard data projections, inclusion of metadata and the administration of data for current and authoritative sources) to be implemented.

The innovation in GIS for the Departments has been in the upgrade to the corporate GIS environment that provides:

- (i) a single source of truth for spatial data through centralised databases,
- (ii) centralised servers for the hosting of desktop and online mapping applications,
- (iii) standards for data management and,
- (iv) integration of business data within the corporate GIS environment.

2.1 Principles of the Corporate GIS Environment

The principles driving the upgrade design and ongoing administration of the corporate GIS aim to increase the value of departmental data by enabling spatial analysis and visualisation, by the principle "create once and use many times" in centralised databases. A strategic approach was taken with the work divided into manageable and achievable stages. Stage One (completed in October 2010) involved the upgrade to a centralised GIS infrastructure to support desktop GIS software and online mapping capabilities. Stage Two (completed in August 2011) involved the integration of business data views through the Department's Corporate Reporting Toolset (CRT) with the GIS environment as implemented in Stage One.

¹ As per the *Emergency Management Act 1986* and Part 4 of the *Emergency Management Manual Victoria* (Offices of the Emergency Services Commission 2011)

2.2 Stage One – Key Considerations in Upgrading the Corporate GIS Environment

The Stage One GIS upgrade aimed to build, configure and implement the core GIS infrastructure required to support a diverse and growing need for GIS capabilities which includes advanced mapping and analysis via ArcGIS and supports easy to use and highly accessible online mapping applications.

With users working from offices located throughout the state it was important that the GIS environment be accessible and reliable, and for performance not to be impacted by distance or networking issues. To achieve this, the Stage One GIS environment leverages virtual servers and the Department's CITRIX application setup to deliver ArcGIS capabilities. This allows users to login to any computer on the Departments' intranet and access not only GIS applications but also the centralised GIS databases, a suite of commonly used layers and GIS resources. This centralisation and high accessibility of applications and standards for spatial data together with additional GIS resources has fostered collaboration across the GIS community in the Departments.

A key consideration when upgrading the corporate GIS environment was data security. The Departments cover a broad area of work and in many program areas data is sensitive and requires appropriate infrastructure rules. As part of the Department's IT environment there is a dedicated data security review. The classification of data dictated much of the environmental settings such as server locations and also workflows with the access to data controlled via user permissions. Ensuring data security was paramount and guided much of the system configuration.

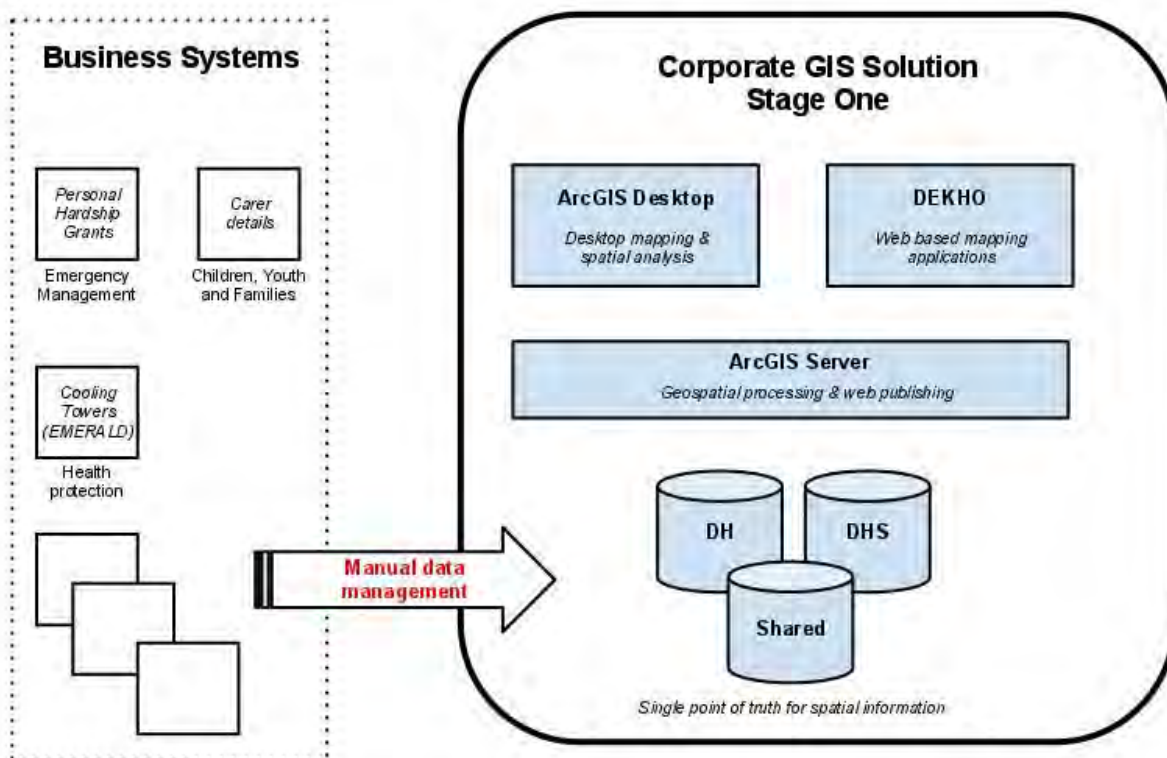


Figure 1 : High level diagram of the DHS and DH corporate GIS solution

2.3 Benefits of Stage One

The benefits of centralising GIS were immediately evident. Within a few weeks of rolling out the system there were a significant number of new GIS users wanting access and training across the department. Online mapping applications were used frequently and supported in a secure setting. In-house administration of the system was essential in the system configuration and has resulted in the responsive deployment of online mapping applications to program areas. This responsiveness to the GIS needs of program areas and the in-house deployment and support has significantly saved time and costs. It has also fostered continual conversations with the program

areas as changes and requests are made quickly, thereby providing the program areas with a sense of ownership and empowerment.

There was considerable streamlining of licence management with users being provided with a suite of additional tools and extensions that were previously not available to them. The spatial databases provide a single point of truth for spatial data and are accompanied with appropriate spatial data standards (including metadata, standard naming conventions and standard data projections). The databases also provide for user-based permissions to be applied to data items with access restrictions. This upgrade of the corporate GIS infrastructure provided a standard platform for the integration of business data views with GIS capabilities.

3 STAGE TWO – BUSINESS INTEGRATION

Stage One of the GIS project provided a firm foundation for spatial analysis and reporting. However the system was constrained by a silo-based data management process. It required the GIS Administrators to manually extract, transform and load information into the system. This process was open to potential errors, information delays and security risks.

Stage Two of the Corporate GIS has provided a method by which:

- (i) Line of Business applications can be integrated into the GIS,
- (ii) data can be obtained from an authoritative source,
- (iii) the program area has control over the data views that can be made available in the GIS,
- (iv) datasets can be added/removed from the process and,
- (v) the integration complies with the departmental infrastructure security model.

The current integration covers three key program areas; the Emergency Management Branch RIEMS (Response, Incident and Event Management System) application, the cooling tower inspection register EMERALD and child protection carer locations from the Children Youth and Families CRIS (Client Relationship Information System) application.

3.1 Identifying the Authoritative Source

The first challenge faced with the implementation of Stage Two was identifying the authoritative source for business data and educating the program areas about the appropriate use and limitations of their data. The datasets required were all similar in nature, included business specific information and included spatial information captured in varying formats. Some program areas had multiple repositories for the same data. In discussions, an often expressed opinion was that one system was better than the other. However there was never any empirical evidence that could confirm this opinion. The view of the GIS project team was that locations should be sourced from the primary system of entry whenever possible and we were able to win over the program areas to this point of view.

3.2 Definition of the Business Data View

With the source system identified, the next task was to define the business data views. We found the best approach was to create a mock-up of their data in a web mapping application and hold work shopping sessions where we could discuss:

- (i) what filters if any should be available for the data,
- (ii) which attribute fields should be visible,
- (iii) what symbology should be applied and,
- (iv) what sort of questions users are likely to ask of the data (and therefore what attributes are required).

The workshops provided an excellent method for the program areas to begin to grasp what was possible with a spatial view of their data and engaged their interest.

3.3 Integration Method and Security

With the program source system and business data views of the datasets defined, the next task was to determine the integration method and data security requirements. The infrastructure security model was a large factor in the design of the integration solution. The following constraints were identified:

- (i) the GIS was capable of storing only 'In Confidence' level information (some program areas require Highly Protected status),
- (ii) real time data updates were not necessarily possible and,
- (iii) direct connections between the GIS and source systems were tightly controlled.

To overcome these constraints it was decided that the system would leverage an existing departmental application, the CRT (see figures 2 and 3). The CRT provides a consistent and integrated foundation for departmental reporting and analysis. Secure data exchange processes with the source systems were not required to be built. It already had fine-grained security levels around an individual's access to information. In essence this meant that the GIS need not hold highly protected information (and therefore be capable of providing the required layers of security); it merely requires the GIS to house a spatial object and a linking mechanism to the CRT. Users can run reports from within the GIS that call the CRT reports available to them (and vice-versa).

Given that the source systems were enterprise databases (a mixture of Oracle and SQL Server) and that the destination was SQL Server 2008, SQL Server Integration Services (SSIS) packages were considered the best integration option between the GIS and the CRT. When an SSIS package executes, the data from the CRT is extracted to a holding database and is classified into two types, 'Core' and 'Non-Core'. Core data is the attributes that determine the location (latitude and longitude) and those attributes that are required for the GIS layers (they become the spatial datasets). SQL Server 2008 also allowed us to take advantage of the inbuilt geometry functions to calculate object-in-object relationships automatically (for example; what local government area does this client reside in). Non-Core data is additional information that is business specific and exists as tables in the holding database. These Non-Core items are controlled by the program area and they can add/remove attributes with no effect on the GIS. The Core and Non-Core layers are linked via a unique key and take advantage of the Query Layer functionality available in ArcGIS 10. If the program area decides to change the information shown in these query layers, there is no need to alter database schemas. The GIS Administrators can simply change the SQL statement in the layer.

The whole process is managed via a GIS integration utility, a custom application that can be used to add/remove datasets in the SSIS packages, control the running of the packages (packages can be run as required or set to run on a schedule), and provide monitoring of the Extract, Transform, Load (ETL) process.

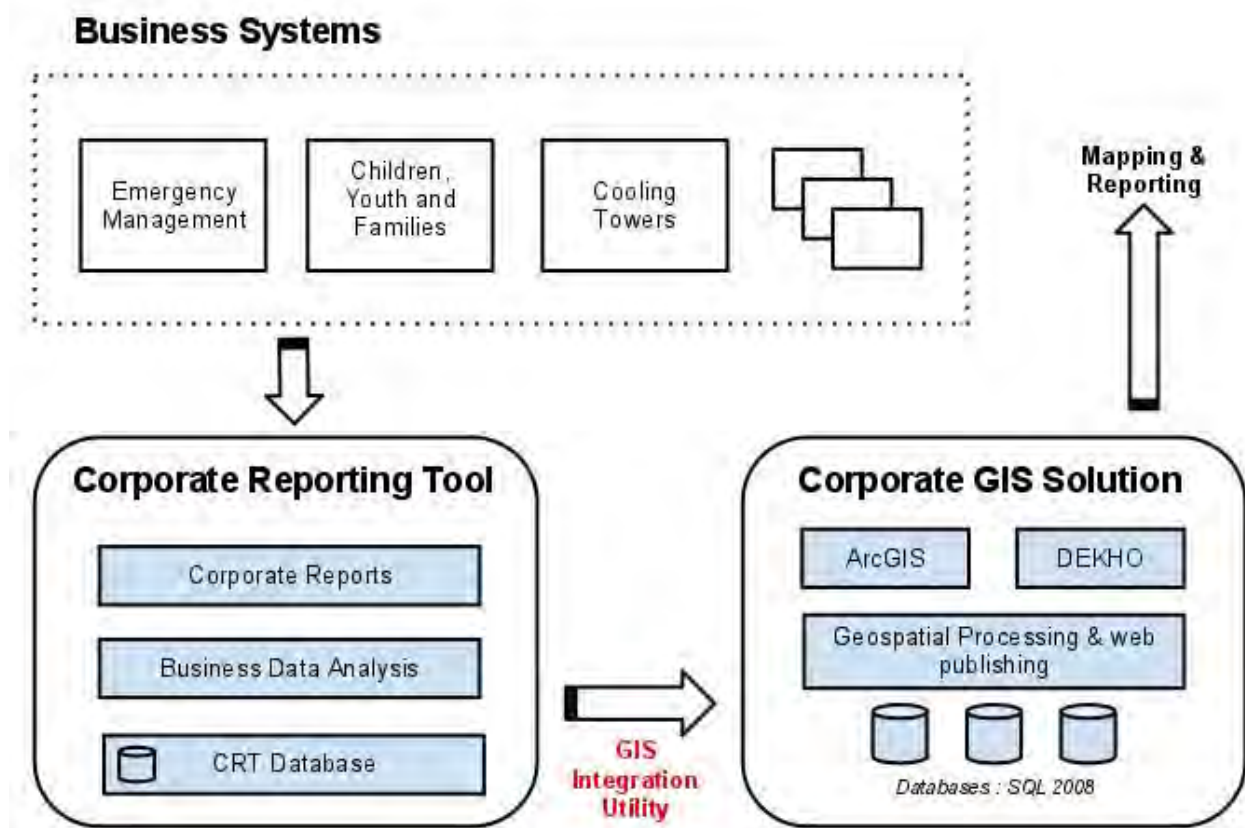


Figure 2: High level diagram of the DHS and DH corporate GIS solution with the CRT integration tool

4 RESULTS AND LESSONS LEARNT

The results from Stage One and Stage Two of the corporate GIS solution have been numerous and diverse. For the first time, many program areas are seeing their program-related data in a spatial context. It is within this context that program areas gain a true understanding of the value of their data, the need to adhere to data standards and to strive to continuously improve the quality and completeness of their datasets.

The key lessons learnt in the deployment of an integrated GIS solution for the Departments has been the identification of source data, the need to actively engage with internal stakeholders through demonstrations of their own data, to understand the security requirements of data items and the limitations of the infrastructure.

By keeping the design, development and deployment in-house we have been able to provide to our program areas a system that is flexible and responsive. The ability to manage the integration of spatial and non-spatial (from the CRT) datasets with the internally developed GIS integration utility has ensured the GIS office can add/remove new datasets without being reliant on external resources. Exploiting in-house skills and knowledge has ultimately saved money and time. Most importantly, it has kept knowledge in-house for continual administration and enhancements of the system.

The training of staff in the use of the desktop solution and the online mapping solutions has also been undertaken from in-house resources. This has been beneficial as training can be developed to suit a specific program need and training sessions can be undertaken at any time given the demand.

The key benefit for the deployment of an integrated GIS solution for the Departments has been the spatial awareness by many program areas. This in turn has added value to their datasets and allowed for an integrated use of programmatic data across the Departments to support an informed approach to planning and operations.

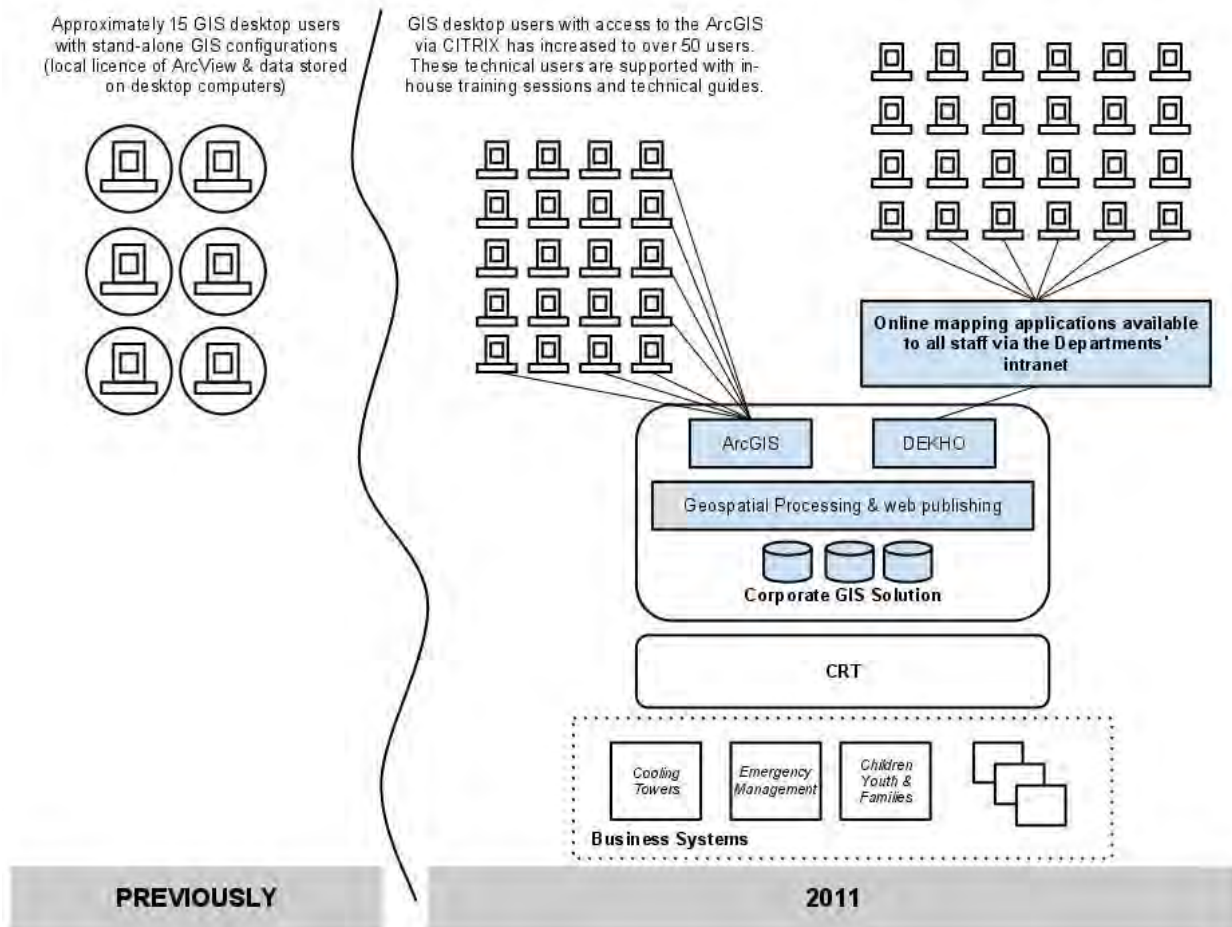


Figure 3 : The stand-alone approach to GIS in the Departments pre 2010 and the current centralized approach to GIS in the Departments.

5 CONCLUSION

The evolution of the corporate GIS within the Department of Human Services and Department of Health continues. Stage Two builds on the solid foundation of the initial deployment and has seen the successful integration of business system data with the GIS. Program areas now comprehend the value of having a spatial context to their information.

The keys to the success of Stage Two of corporate GIS have been in exploiting internal capabilities to design, develop and deploy the new capability, and the value to the Departments is just starting to be realised through the use of GIS to provide an integrated approach to planning and operations.

REFERENCES

Department of Health 2011, http://www.health.vic.gov.au/about_us/

Department of Human Services 2011, <http://www.dhs.vic.gov.au/about-the-department/our-organisation>

Offices of the Emergency Services Commission 2011,
<http://www.justice.vic.gov.au/emanuals/EMMV/default.htm>

Appendix

1. Detailed Design of the GIS Integration Utility

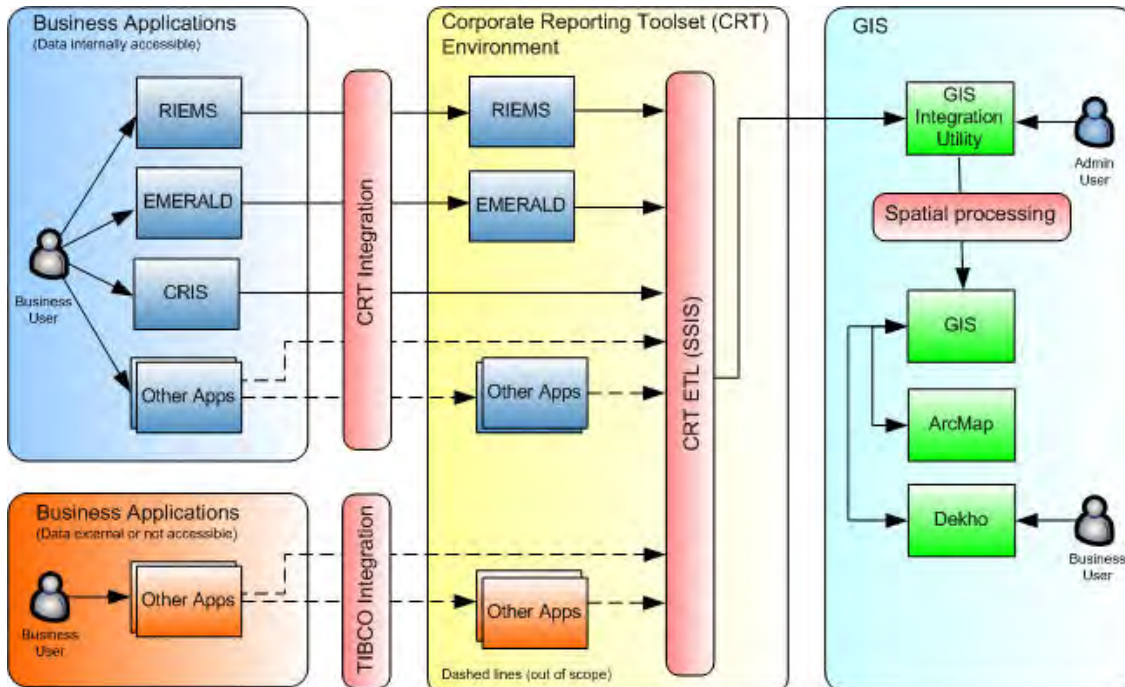


Figure 4: GIS Integration is achieved via the CRT. The CRT is a preexisting suite of integrated and consistent products. Data for the GIS is extracted via SSIS packages into a holding database. A series of processes are triggered to separate the data into core and non-core views, as well the calculation of object in object relationships. After processing, the core data is loaded into the GIS (ArcSDE on SQL Server 2008) databases.

Building Better Hyperspectral Datasets: The Fundamental Role of Metadata Protocols in Hyperspectral Field Campaigns

Barbara Rasaiah, Simon Jones, Chris Bellman

RMIT University
Melbourne, Australia

barbara.rasaiah@rmit.edu.au, simon.jones@rmit.edu.au, chris.bellman@rmit.edu.au

Tim Malthus

CSIRO Land and Water
Canberra, Australia
tim.malthus@csiro.au

ABSTRACT

Hyperspectral data exchange is becoming increasingly prolific in the international remote sensing community for cataloguing and mining data from a variety of sensors. Coupled with the growing volume of hyperspectral data being disseminated is the need for a standard method for successfully integrating a broad range of campaigns and data formats. Metadata is an important component in the cataloguing and analysis of such datasets, because of its central role in the quality and reliability of hyperspectral data and the products derived from it. Metadata affects all tiers of data processing, from the recording of *in situ* measurements, to instrument calibration and data validation. The metadata definitions must also provide the flexibility for users to both create reliable datasets on an ad-hoc basis, and to export them to datawarehouses. Field spectroscopy can be viewed as the primary platform for hyperspectral research, and consequently the metadata protocols defined at this stage have profound and long term effects on their associated datasets. Currently no standardized methodology for collecting *in situ* spectroscopy data or metadata protocols exists.

This paper presents the progress of an ongoing international expert panel survey investigating metadata protocols in field spectroscopy. Examined are issues such as field measurement methods, instrument calibration, and data representativeness and their expression as metadata entities. Panel recommendations include metadata protocols that are critical to all campaigns, and those that are restricted to individual campaigns, ranging from vegetation, mineral exploration, marine, and pseudo invariant targets. The survey is part of a doctoral research project to investigate approaches to a coordinated evolution of hyperspectral metadata protocols, field spectroscopy methods and data exchange standards within the hyperspectral remote sensing community. This paper also proposes a way forward for adapting and enhancing current geospatial metadata standards to the unique requirements of field spectroscopy.

Keywords: *Databases, Datawarehouse, Hyperspectral, Metadata, Field Spectroscopy, Validation, Calibration, Data Quality*

1 INTRODUCTION

Metadata is a vital component in the quality, validity, and reliability of hyperspectral datasets and the products derived from them. Hyperspectral datasets derived *in situ* are all sensitive to the presence and quality of their associated metadata. Hyperspectral metadata can describe the properties of the target being viewed, illumination and viewing geometry, sensor calibration, and environmental conditions -- all of which are influencing factors that affect standardized measurements (Pfitzner *et al.*, 2006). It is therefore not sufficient to rely upon the radiometric data alone to exploit the potential for accurate and consistent environmental modelling and long-term data legacy. In the context of remote sensing, field spectroscopy acts as the fundamental stage for primary research and operational applications (Milton, 1987). Both the imaging conditions and instrument itself can introduce systematic and random errors on recorded radiance, target discriminability and contrast, prompting the need for ancillary information including conditions of observation and field techniques (Duggin, 1985). Metadata can be an effective tool in describing and quantifying these errors, and potentially mitigating them. The time invested in metadata collection is outweighed by its benefits in reducing system bias and variability (Pfitzner *et al.*, 2006). To address such concerns, a standardized methodology for defining and storing metadata must be closely aligned to *in situ* data collection practices. Currently no such methodology for collecting *in situ* spectroscopy data or metadata protocols exist.

Creating a protocol for *in situ* metadata collection requires a definition of the metadata fields required, and their prioritization in a field campaign. Opinions on this matter vary among the expert user groups throughout the international remote sensing community. HYRESSA (HYperspectral REmote sensing in Europe - specific Support Actions) has identified highly individual preferences among hyperspectral users in Europe, with each application group (vegetation, atmosphere, land, water) queried assigning varying importance to spectral, geometric, radiometric, and temporal parameters, and placing unique emphasis on spectral calibration quality and preferred observation time (Nieke *et al.*, 2007). Other weaknesses in hyperspectral data collection and sharing have been identified by users in the European remote sensing community and include a lack of quality assurance and calibration information for sensors; no real capability to define accuracy or validation for data processing; a lack of agreed standards in data processing, and the need for more transparency on calibration processes (Reusen *et al.*, 2007). User needs for quality assurance must be formally identified, in tandem with a standardized protocol for hyperspectral metadata storage and data exchange. Such an activity would require a determination of the metadata requirements for field spectroscopy for the full range of hyperspectral campaigns, standardizing field spectroscopy protocols for accuracy and consistency, and establishing file exchange protocols allowing flexibility to capture hyperspectral metadata and enabling fusion with other remote sensing products (Rasaiah *et al.*, 2011).

Even when calibration and instrument setup activities (instrument calibration, white referencing, viewing geometry, etc.) are executed under controlled conditions, subsequent sample measurements of the same phenomena can still vary from operator to operator. An experiment involving fifteen spectroscopy laboratories using the same instrument, targets, and a uniform instrument calibration protocol showed variations in output reflectance measure, suggesting strong implications for the accuracy of spectral matching and quality of spectral data when building and querying spectral libraries (Jung *et al.*, 2010). Detection of pseudo invariant targets is no more immune than others to the dependency of documentable and repeatable protocols for accurate and reliable results. This was demonstrated in highly controlled laboratory-based sampling for concrete quality determination (Brook and Ben-Dor, 2011) and in soil reflectance measures, where variances were shown to be affected

by the operator, thereby demonstrating the need to minimize uncontrolled spectral variation by standardizing measurement protocols (Ben-Dor, 2011).

More complications for metadata standardization are introduced in specialized field campaigns in the natural environment using heterogeneous targets. For example, a customized underwater spectrometer system was developed at the University of Queensland, tailored specifically to coral reef ecology, and the ecology and physiology of animal colour vision; it is accompanied by protocols for recording metadata *in situ*. These protocols are interdependent with the challenges of radiometric data collection underwater as they are designed to simultaneously ensure the requisite operator safety (Roelfsema *et al.*, 2006). In the context of international data sharing of substratum and benthic spectral data, the establishment of standards for the capture, storage, and use of spectral signature files with associated metadata is required due to the effect of environmental factors in shallow water environments on the derived data (Dekker *et al.*, 2010).

These concerns prompt the need for creating a mechanism to determine the metadata fields that are critical for valid and reliable field spectroscopy datasets. Such a mechanism must establish the required metadata with enough integrity to generate datasets for long-term cataloguing and data warehousing across a range of campaigns. Ideally it should also be easily accessible to an international audience with a broad spectrum of expertise. In this paper, the detailed requirements for such metadata are investigated.

2 ASKING THE EXPERTS

An online and duplicate hardcopy survey was launched in early 2011 in the form of a user-needs analysis for field spectroscopy metadata. The purpose of the survey, still open to the remote sensing community¹, is to determine, based on the input of an expert panel, the metadata fields that are critical for creating valid and reliable field spectroscopy datasets, with enough integrity to generate datasets for long-term cataloguing and data warehousing across a range of campaigns. The metadata fields are closely associated with common *ad hoc* field spectroscopy protocols practiced by remote sensing communities around the world. The audience is an international panel of scientists with expertise in *in situ* hyperspectral remote sensing, who were asked to respond on an anonymous basis. Each participant assesses the criticality of several categories of metadata fields, and can propose additional metadata fields that they believe could enhance the quality of a hyperspectral dataset generated in the field. Open-ended comments are available throughout the survey for further input in each metadata category.

Respondents have the option of participating in the categories of their choice, and are also asked to nominate themselves as experts in one or more areas of field spectroscopy ('forest and woodland' / 'agriculture' / 'other specified vegetation' / 'marine' / 'estuarine' / 'mineral exploration' / 'soils' / 'snow' / 'urban environments' / 'other specified'). This self-nomination of area of expertise does not in any way limit the categories available to each participant, and primarily serves the purpose of informing correlative analysis between a participant's area of expertise and their assessment of metadata criticality. Table 2.1 shows the metadata categories, divided between general campaign metadata and specific campaign metadata (approximately 200 fields in total). Metadata fields presented in the survey can be given one and only one ranking:

¹ This web-based survey is open to hyperspectral remote sensing scientists with expertise in *in situ* campaigns. It remains open at the time of the conference. Please contact Barbara Rasaiah directly at barbara.rasaiah@rmit.edu.au if you wish to participate.

- ‘critical’ (required metadata field for a field spectroscopy campaign; without this data the validity and integrity of the associated spectroscopy data is fundamentally compromised)
- ‘useful’ (not required, but enhances the overall value of the campaign)
- ‘not useful now but has legacy potential’ (not directly relevant to the associated field spectroscopy data but potentially has use in a related hyperspectral product)
- ‘not applicable’ (this metadata is not relevant to this campaign)

Table 2.1 Categories of metadata fields in the survey

Generic campaign metadata	Campaign-specific metadata
♦ instrument	♦ vegetation
♦ reference standards	♦ woodland and forest
♦ calibration	♦ agriculture
♦ hyperspectral signal properties	♦ soil
♦ illumination information	♦ mineral exploration
♦ viewing geometry	♦ snow
♦ environment information	♦ urban environments
♦ atmospheric conditions	♦ marine and estuarine
♦ general project information	♦ underwater substratum targets
♦ location information	
♦ general target and sampling information	

Survey participants are given the definitions of each metadata ranking before responding in each category. Figure 2.1 shows a screenshot of the instrument calibration category, with associated metadata fields and their proposed rankings.

Figure 2.1 Metadata fields for ‘Calibration’ category in the survey

	Critical	Useful	Not useful now, but has legacy potential	N/A
Date	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Irradiance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Radiance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dark noise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Signal to Noise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Linearity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stray light	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Calibration data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Traceability (e.g. Yes, No)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Standard (e.g. NIST, NPL)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

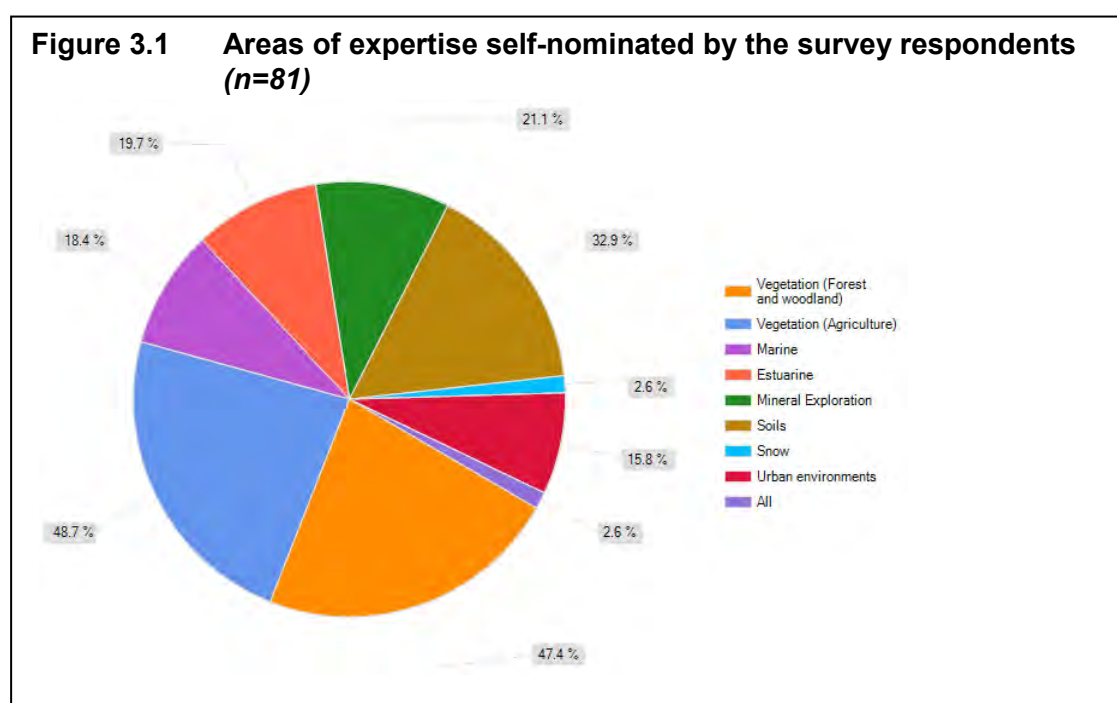
Ascertaining the metadata fields that are critical and useful for creating reliable datasets, and which ones, in the opinion of the expert panel, are negligible or even irrelevant, will enable confident decision-making for the inclusion of metadata fields in a given campaign; it will also influence the method in which metadata can be

defined, stored, flagged for quality, shared, and archived. Trends for consensus and variances in responses, both among expert groups and between them, will help inform future protocol design for metadata collection across a range of campaigns.

3 SURVEY RESULTS & DISCUSSION

The survey results as of August 2011² show a total number of 81 participants, with many comments provided by the respondents in each category of expertise and their opinions on metadata and *in situ* protocols. Preliminary albeit limited analysis is possible for some of the categories with comprehensive analysis only possible when the survey has closed. Final decisions about the criticality of metadata fields must be reserved for the conclusion of the survey.

Figure 3.1 illustrates the areas of expertise identified by the survey respondents. Areas of spectroscopy research beyond this scope, as stated by the respondents, include atmospheric study, calibration and validation activities for airborne sensors, and wetlands and peatlands research.



An example of the variety of answers between groups can be found in the responses provided by the marine, estuarine, and soils groups on illumination metadata. Table 3.1 shows the percentage of responses by the three groups for six fields in the illumination category. Some soils experts rated several fields (optical measure of ambient conditions, bulb intensity, light spectrum, single beam/multi-beam) as 'not applicable' ('N/A') that were rated as 'critical' and 'useful' by their marine and estuarine counterparts. However, the fields 'source of illumination' and 'beam coverage' showed more agreement among the three groups in being designated as 'critical' or 'useful' more often than the other metadata fields in the illumination category. Causes for the discrepancy in responses may be numerous

² This web-based survey is open to hyperspectral remote sensing scientists with expertise in *in situ* campaigns. It remains open at the time of the conference. Please contact Barbara Rasaiah directly at barbara.rasaiah@rmit.edu.au if you wish to participate.

and include the purpose for which the data is being collected (calibration/validation, environmental modelling, archiving in a spectral library, etc), the instruments being used by each group, the kind of data being recorded (radiance/irradiance/reflectance) the novelty of the campaign for which the data is being collected, such as collecting new signatures for a coral reef ecosystem library (Dekker *et al.*, 2010), and its fitness for use in the applications the respondents collect the data for. A more thorough statistical analysis at the conclusion of the survey may highlight more trends and patterns. Further speculation on the cause for this pattern of responses would require an investigation into these factors and the additional comments provided by the participants.

Table 3.1 Criticality ranking for illumination metadata fields from Marine, Estuarine, and Soils expert groups as a percentage of responses ($n=38$)

		Marine	Estuarine	Soils
Source of illumination (eg. sun, lamp)				
	Critical	91.7	93.3	90.5
	Useful	8.3	6.7	9.5
	Not useful now, but has legacy potential	0.0	0.0	0.0
	N/A	0.0	0.0	0.0
Optical measure of ambient conditions (direct, diffuse)				
	Critical	58.3	60.0	55.0
	Useful	41.7	40.0	35.0
	Not useful now, but has legacy potential	0.0	0.0	5.0
	N/A	0.0	0.0	5.0
Bulb Intensity				
	Critical	36.4	33.3	42.9
	Useful	54.5	53.3	42.9
	Not useful now, but has legacy potential	9.1	13.3	9.5
	N/A	0.0	0.0	4.7
Light Spectrum				
	Critical	45.5	53.3	45.0
	Useful	45.5	40.0	40.0
	Not useful now, but has legacy potential	9.1	6.7	5.0
	N/A	0.0	0.0	10.0
Single beam / multi-beam				
	Critical	27.3	26.7	55.0
	Useful	72.7	73.3	30.0
	Not useful now, but has legacy potential	0.0	0.0	5.0
	N/A	0.0	0.0	10.0
Beam coverage (as a degree measure)				
	Critical	18.2	20.0	45.0
	Useful	81.8	80.0	40.0
	Not useful now, but has legacy potential	0.0	0.0	5.0
	N/A	0.0	0.0	0.0

When designing a protocol based in the input of the respondents, it is important to specify what determines the threshold for a metadata field to be considered critical and therefore incorporated into a metadata schema. The difficulty in obtaining an easy answer becomes apparent when viewing some of the responses for the viewing geometry category for 47 vegetation experts (Figure 3.2) For some metadata fields ('distance from target', 'area of target in field of view'), there is a higher proportion of the group designating these fields as critical. However, the opinion becomes more nebulous on a metadata field such as 'distance from ground/background', where the responses are divided almost equally between 'critical' and 'useful'. By comparison, the answers given by this group for environment information metadata (Figure 3.3) show much less consensus. For example, the responses for 'ambient temperature' are assigned 'useful' nearly as often as 'not useful now but has legacy potential'. The low degree of 'critical' response also prompts for future analysis to ascertain how often a metadata field needs to be identified as critical before it is included as mandatory in a protocol.

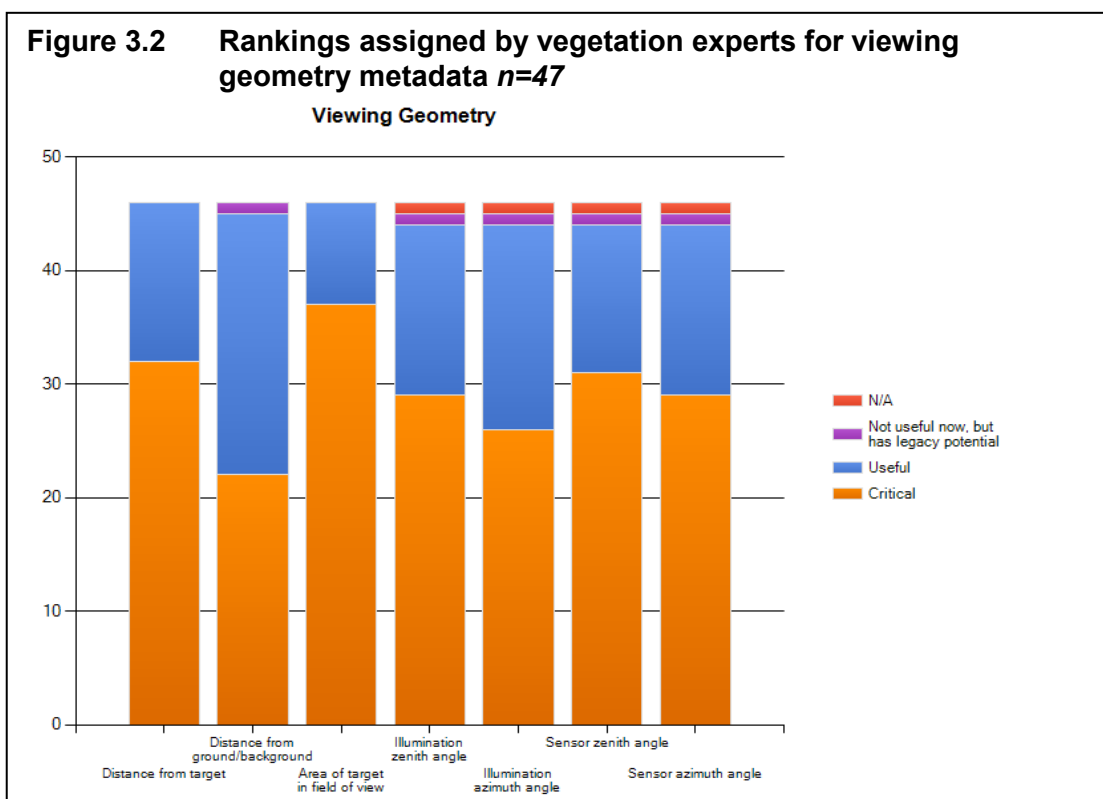
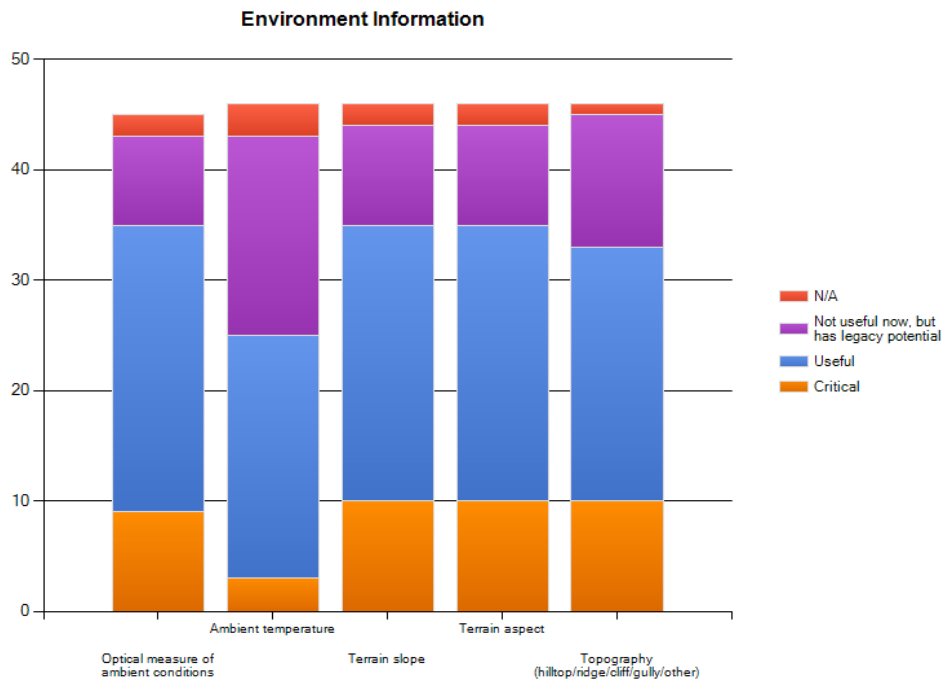


Figure 3.3 Rankings assigned by vegetation experts for environment information metadata $n=47$



A detailed quantitative analysis after the close of this survey will aid in determining how a metadata protocol can be designed, including the statistical trends of consensus and least consensus both among groups and between them. The analysis will also address the question of how much flexibility must be allowed in a protocol to ensure a quality standard while permitting the operator in the field to adapt the schema *ad hoc* to allow novel research to take place. Some of the questions that the survey will help answer:

1. Which metadata categories show most consensus?
2. Which metadata categories show least consensus?
3. Are there consensus trends among and between groups?
4. Which metadata fields are identified as critical/useful/potential/NA most often?
5. In which categories is there most agreement on criticality?
6. How can we determine that a metadata field can be designated as “critical”?
7. Are there metadata fields that are difficult to designate as critical?
8. Are there other trends across all respondents and individual groups?

Protocol design cannot be informed solely by the quantitative data, and the comments and suggests can provide additional recommendations. Some of the suggestions and comments from the participants thus far include:

- “the context of inquiry must be specific enough to address the variety of type of radiometric data (reflectance, radiance, irradiance, transmission, etc.) and the purpose of the measurements (field survey, algorithm development)”

- “regardless [of] the applications of the field spectroscopy, metadata should contain sufficient information for users 1) to repeat the sampling (or in the least to imagine the measurements and its surrounding condition), 2) to cite and pinpoint the dataset for the reference, and 3) to explore the data as much flexible as possible, even beyond its original purpose”
- “depending on the campaign and available budget and instrumentation different [metadata] points become critical and other[s] useful or negligible”
- “there's a need for an integrated 'quality flag' so that people can rapidly assess whether to utilise the data or not”

At the closure of this survey, these responses and recommendations can be used to build a framework for metadata protocols that support the requirements of the remote sensing community for efficient storage and sharing of reliable and high quality *in situ* metadata.

4 FIELD SPECTROSCOPY METADATA SHARING AND INTEROPERABILITY

The survey results provide a platform for investigating methods for adapting and enhancing current geospatial metadata standards to the unique requirements of field spectroscopy. Incorporating the recommendations and comments provided by the expert panel into a metadata protocol that ensure long-term data integrity has a relationship with the methods chosen for both encoding and sharing the metadata. Data encoding standards affect the manner in which the hyperspectral metadata fields are defined and stored *in situ*. With the emergence of online data archives and the proliferation of web-interfacing applications, it is necessary to use a metadata encoding format that is sufficiently robust to capture the required data elements and provide hierarchical information, while being popular, compatible with spectroradiometric software (ViewSpecPro, SpecWin, SVC), and with the capability for efficient distribution across wide-area networks for integration into external archives.

The potential for finding some solutions to the issues raised earlier by the survey respondents, specifically those referring to radiometric data type, dataset referencing and flexibility, and quality flagging may lie in the metadata modelling techniques currently being explored around the world. OGC (Open Geospatial Consortium) and INSPIRE (Infrastructure for Spatial Information in the European Community) have both adopted architecture and data interoperability protocols for geospatial metadata based on EN ISO 19115 and EN ISO 19119. Although providing general guidelines, neither of these explicitly address the metadata requirements of hyperspectral field collection techniques, or the ontologies and data dependences required to model the complex interrelationships among the observed phenomena as data and metadata entities. Dependencies include the influence of environmental phenomena such as wind speed and cloud cover on the recorded spectrometer signal, or user-controlled viewing conditions such as sensor orientation and height above the target.

The US LTER (Long Term Ecological Research) network, composed of 26 global research sites extending over a breadth of biomes and ecosystems, demonstrated the use of the Ecological Metadata Language (EML) in improving the flexibility of metadata for archiving processes (LTER, 2011). More specific metadata schema for vegetation observations have been proposed (Pfitzner *et al.*, 2006 and Hüni *et al.*, 2007) but these are mostly on an *ad hoc* basis and do not yet encompass the full spectrum of common field spectroscopy campaigns.

An Extensible Markup Language (XML)-based exchange format for spectroradiometric metadata is proposed because it facilitates searching and selection, it is human and machine readable, platform independent, convertible to other formats and allows quick assessment of suitability for other research products (Malthus and Shironola, 2009). The XML format can be easily accommodated in a variety of data archiving schema and software, including spectral libraries, databases, and datawarehouses. The strengths and weaknesses of data encoding format becomes a valuable debate when designing an *in situ* hyperspectral metadata protocol, because both the representation of the metadata in a digital format and its subsequent compatibility with schema such as the datawarehousing model for large-scale archiving, sharing, and mining of hyperspectral metadata need to be considered (Rasaiah *et al.*, 2011). XML is self-descriptive with extensibility features (Mahboubi and Darmont, 2010), facilitating integration with multi-dimensional remote sensing data sets. One of its greatest strengths is platform independence, and a framework for XML-based data interchange is espoused in the Common Warehouse Metamodel, which includes XML Metadata Interchange (XMI) standards for datawarehouses (Mangisengi *et al.*, 2001 and Torlone, 2009). Water ML 2.0, now an OGC specification, is an example of an XML schema that could be modified and adopted for hyperspectral *in situ* data; it is used for encoding hydrological observation and measurement data, and accommodates quality flagging by incorporating metadata (contaminated sample, holding time exceeded, etc.) that affects the data's fitness for use in future applications (Terhorst, 2009).

Metadata definitions must also provide the flexibility for users to create reliable datasets on an *ad hoc* basis, and to export their data to enterprise-scale databases and data warehouses for other users. Coupled with this is the need to preserve the integrity and quality of the original dataset and safeguard it from data corruption and information loss as it cascades through the extraction, transformation and loading processes in a datawarehouse.

Lack of standardization in this area is sufficiently common in the remote sensing community that there are ongoing efforts to provide a solution. NASA's SPG (Earth Science Data Systems Standards Process Group) is investigating candidate standards for data access protocols and interoperability between OGC (Open Geospatial Consortium) Catalog Services for the Web and Web Coverage Services protocols and is working on development of a NASA Earth Science Missions Data System reference architecture (Ullman and Enloe, 2010). The IEEE GRSS (Geoscience and Remote Sensing Society) Data Archiving and Distribution Technical Committee is currently exploring data archiving, developing data availability methods and online discovery functionality for distributed datasets, as well as standards for information and systems (Rochon *et al.*, 2010). OGC has also adopted the Sensor Observation Service standard and Sensor Model Language for modelling the interface between *in situ* sensors, dynamic remote sensors, and sensor networks, and for enabling data retrieval either in real-time or through data archives.

5 CONCLUSION

There is an established and persistent need in the remote sensing community for a coordinated evolution of hyperspectral metadata protocols, field spectroscopy methods and data exchange standards. One solution lies in formally identifying user needs for creating high quality, reliable metadata, and determining a method that enables sharing of the data for long-term archiving and fusion with other datasets, as is proposed in this research.

Utilising the expert panel input in the field spectroscopy metadata survey will provide building blocks for *in situ* metadata protocols that are robust enough to meet the requirements of a range of hyperspectral campaigns, while ensuring accuracy

and consistency. Determining thresholds for criticality is an important step in designing a schema that meets the needs of both the creators of the data and the users of the data.

Much potential exists for adapting and improving current geospatial metadata standards for the unique requirements of the hyperspectral remote sensing community. A standardized protocol for *in situ* hyperspectral metadata storage and data exchange remains the best option for the management of hyperspectral data and metadata, and can be accomplished by aligning data sharing and mining within the remote sensing community with datawarehousing practices. A spirit of collaboration and innovation can bring great benefits to international efforts for providing the data sharing capabilities and quality control tracking for the hyperspectral remote sensing community.

If you wish to participate in this survey, please contact Barbara Rasaiah directly at barbara.rasaiah@rmit.edu.au.

ACKNOWLEDGEMENTS

- Anonymous survey participants who generously devoted their time and expertise
- Kind-hearted survey participant who mailed their survey overseas to my office in Australia

Others who have contributed invaluable input to this paper:

- A. Riaza, *Instituto Geologico y Minero de España, Spain*
- Adrian Schembri, *RMIT University, Australia*
- Alexander Goetz, *ASD Inc., USA*
- Andy Hueni, *RSL, University of Zurich, Switzerland*
- Anna Brook, *Remote Sensing Laboratory, Tel-Aviv University, Israel*
- Anthony Bedford, *RMIT University, Australia*
- Arnold G Dekker, *CSIRO Land & Water, Australia*
- Chariton Kalaitzidis, *Mediterranean Agronomic Institute of Chania, Greece*
- Christian Götze, *Martin-Luther-Universität Halle-Wittenberg, Germany*
- David J. Williams, *Research Physical Scientist, US EPA Environmental Sciences Division, USA*
- Chris Roelfsema, *Centre for Spatial Environmental Research (CSER), University of Queensland, Australia*
- Gita Pupedis, *RMIT University, Australia*
- Giuseppe Ciraolo, *Area Ingegneria Idraulica e Ambientale, Dipartimento di Ingegneria Civile, Ambientale e Aerospaziale, University of Palermo, Italy*
- H.J. van der woerd, *Institute for Environmental Studies, Vrije Universiteit De Boelelaan, Netherlands*
- James Cleverley, *CSIRO Earth Science and Resource Engineering, Australia*
- Janet Anstee, *Environmental Earth Observation Group, CSIRO Land and Water, Australia*
- Jon Huntington, *CSIRO Earth Science and Resource Engineering, Australia*
- Jorge Buzzi Marcos, *Geology Survey of Spain (IGME), Spain*
- Juan Pablo Guerschman, *Environmental Earth Observation, CSIRO Land and Water, Australia*
- Kiril Manevski, *Mediterranean Agronomic Institute of Chania, Greece*
- Lammert Kooistra, *Wageningen University, Netherlands*
- Loredana Pompilio, *Dipartimento di Scienze della Terra, Università degli Studi di Parma, Italy*
- Marcos Jimenez Michavila, *Remote Sensing Laboratory, INTA, Spain*
- Martin Bachmann, *German Remote Sensing Data Center (DFD) Land Applications (LA) Applied Spectroscopy Workgroup, German Aerospace Center (DLR), Germany*
- Michael Schaepman, *University of Zurich, Switzerland*
- Peter D. Hunter, *Biological and Environmental Sciences, School of Natural Sciences, University of Stirling, United Kingdom*

- Peter Gege, *Remote Sensing Technology Institute, German Aerospace Center (DLR), Germany*
- Petr Lukeš, *Global Change Research Centre, Academy of Sciences of the Czech Republic, Czech Republic*
- Reno Choi, *Geography and Environment, University of Southampton, United Kingdom*
- Roland Doerffer, *Helmholtz Zentrum Geesthacht, Germany*
- Sabine Chabrillat, *Remote Sensing section, GFZ German Research Center for Geosciences, Germany*
- Steven M. de Jong, *Utrecht University, Netherlands*
- Uta Heiden, *German Remote Sensing Data Center (DFD) Land Applications (LA), German Aerospace Center (DLR), Germany*
- Uwe Rascher, *Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Forschungszentrum Jülich GmbH, Germany*
- Vittorio Brando, *CSIRO Land & Water, Australia*

REFERENCES

- A. Brook and E. Ben-Dor. "Reflectance spectroscopy as a tool to assess the quality of concrete in situ", *Journal of Civil Engineering and Construction Technology*, Volume 2:8, 169-188, 2011.
- A. G. Dekker, V. E. Brando, J. M. Anstee, A. J. Botha, Y. J. Park, P. Daniel., T.J.M. Malthus, S. R. Phinn., C. M. Roelfsema, I.A. Leiper, S. Fyfe, "A Comparison of Spectral Measurement Methods for Substratum And Benthic Features in Seagrass and Coral Reef Environments" *Proceedings of ASD and IEEE GRS; Art, Science and Applications of Reflectance Spectroscopy Symposium*, Vol. II, 15pp, Boulder, CO, 2010.
- A. Hüni , J. Nieke, J. Schopfer, M. Kneubühler and K. I. Itten, "Metadata of Spectral Data Collections" *Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy*. Belgium, April 2007.
- A. Jung, C. Götze, and C. Gläßer, "White Reference Tour 2009. A round-robin test for better spectral libraries", *DGPF Tagungsband*, 433-439, 2010.
- A. Terhorst. "WaterML 2.0.", *GEOSS Sensor Web Workshop Tsukuba, Japan*, 21-22 May 2009.
https://portal.opengeospatial.org/files/?artifact_id=34678 (accessed August 02, 2011)
- B. Rasaiah, T.J. Malthus, S.D. Jones, C. Bellman, "The Role of Hyperspectral Metadata In Hypspectral Data Exchange and Warehousing", *Proceedings of the 7th EARSeL Workshop*, Edinburgh, Scotland, 2011 (*in press*) .
- C. Roelfsema, J. Marshall, E. Hochberg, S. Phinn, A. Goldizen, and K. Joyce, "Underwater Spectrometer System 2006 (UWSS04)", University of Southern Queensland, 2006, viewed August 01 2011, <http://ww2.gpem.uq.edu.au/CRSSIS/publications/UW%20Spec%20Manual%2029August06.pdf>.
- E. B. Dor, "Variation and Stability of Soil Reflectance Measurements With Different ASD Spectrometers Under Different Conditions", *Proceedings of ASD and IEEE GRS; Art, Science and Applications of Reflectance Spectroscopy Symposium*, Vol. II, 7pp, Boulder, CO, 2010.
- E.J. Milton, "Principles of Field Spectroscopy", *International Journal of Remote Sensing*, 8:12, pp1807-1827, 1987.

G. L. Rochon, H. Ramapriyan, E. F. Stocker, R. Duerr, R. Rank, S. Nativi, "Advances in Spatial Data Infrastructure, Acquisition, Analysis, Archiving & Dissemination", *IEEE International Conference on Geoscience and Remote Sensing Symposium*, 1980-1983, July 2010.

H. Mahboubi and J. Darmont, "Optimization in XML Data Warehouses", pp 232-253 *E-Strategies for Resource Management Systems: Planning and Implementation*, E. Alkhalifa (Ed.), University of Bahrain, 2010.

I. Reusen et al., "Towards an improved access to hyperspectral data across Europe", *ISIS meeting*, Hilo, 2007.

J. Nieke et al., "User-driven requirements of the European Hyperspectral Remote Sensing Community", *HYRESSA Workshop*, Davos, Switzerland, March 2007.

K. Pfitzner, A. Bollhöfer, and G. Carr, "A Standard Design for Collecting Vegetation Reference Spectra: Implementation and Implications for Data Sharing" *Spatial Science*, 52:2, 79-92, December 2006.

LTER. "Validating Metadata at the VCR/LTER". <http://databits.lternet.edu/spring-2011/validating-metadata-vcrlter> (accessed August 02, 2011)

M.J. Duggin, "Factors limiting the discrimination and quantification of terrestrial features using remotely sensed radiance", *International Journal of Remote Sensing*, 6: 1, 3-27, 1985.

O. Mangisengi, J. Huber, C. Hawel, W. Essmayr, "A Framework for Supporting Interoperability of Data warehouse Islands Using XML", *Lecture Notes in Computer Science*, 2001 *Data Warehousing and Knowledge Discovery*, Volume 2114, 328-338, 2001.

R. Torlone, "Encyclopedia of Database Systems", Part 9, p. 1560-1564, Springer Science + Business Media, LLC, 2009.

R. Ullman and Y. Enloe, "NASA's Standards Process for Earth Science Data Systems", *IEEE International Conference on Geoscience and Remote Sensing Symposium*, 4232 - 4235, July 2010.

T. Malthus and A. Shirinola, "An XML-based format of exchange of spectroradiometry data", *EARSeL Imaging Spectroscopy SIG*, Tel Aviv, March 2009.

T. U. Kampe, "Radiometric Calibration Concept of Imaging Spectrometers for a Long-term Ecological Remote Sensing Project", *Proceedings of ASD and IEEE GRS; Art, Science and Applications of Reflectance Spectroscopy Symposium*, Vol. III, 11pp, Boulder, CO, 2010.

Geospatial enablement of the Māori Land Court's www.maorilandonline.govt.nz: A Case Study

Tobias Luetticke

Ministry of Justice

Information and Communication Technology Services

The Vogel Centre

19 Aitken Street

SX10088

WELLINGTON

New Zealand

tobias.luetticke@justice.govt.nz

ABSTRACT

In order to improve services to public Māori land owners and internal staff the Māori Land Court initiated a project to geospatially enable its www.maorilandonline.govt.nz website. As authoritative custodian of all Māori land ownership information it has an interest in making this information available and easily accessible to all interested parties. The implementation project redeveloped the previous system and delivered a state of the art website that also includes a map to visualise Māori land blocks. It provides a wealth of associated information and context data.

Changing from a purely text based user interface to one that emphasises the visual gives a natural and intuitive delivery of the relevant information. This represents a paradigm shift and a significant improvement in terms of usability and access to Māori land information. By enriching the ownership data held by the Māori Land Court with cadastral information from LINZ and superimposing it over satellite imagery and street maps the value provided to users is greatly increased.

Positive user testimony from individuals, iwi, and Māori Land Court staff across the country demonstrates this leap in applicability and usefulness of the content made available. Interested parties can now easily check what interests they have in land titles and gather all relevant information about the particular piece of land. The geospatial component enables them to make better informed decisions about how to leverage the economic potential of the land.

The system makes heavy use of Open Source geospatial components and of GoogleMaps as the base image provider. To a large extent the implementation is an example of assembling pre-existing technology building blocks rather than pursuing a bespoke development approach or purchasing a proprietary commercial system. This is possible due to the inherent support of the Open Source products for open standards and open interfaces. The implementation of the new system proves that these Open Source components are reliable and mature enough to allow development and operation of a mapping application and public facing website that supports users 24 hours a day.

KEYWORDS: open source GIS, usability, commodity, public access, Google Maps, open standards, property

1 INTRODUCTION

As of today, approximately 5.5 per cent of New Zealand, equating to roughly 1.4 million hectares, represents Māori freehold or customary land (Judge Isaac W.W. 2011). This land is estimated to be worth in excess of \$6 billion (Te Puni Kōkiri, 2009). Despite the fact that it represents a significant asset with immense economic potential, its utilisation is currently

substantially hampered by structural constraints and lack of information among the owners. The Māori Land Court (MLC), as custodian of the land records, could not provide its services as efficiently as desired and the Māori land owners could not be provided with the relevant and comprehensive set of information that would have enabled them to much more actively make use of their properties.

1.1 The Māori Land Court and Māori land in New Zealand

The Māori Land Court acts as Registry of all Māori land. It is the custodian of over 2.3 million ownership interests in 27,137 land blocks (titles). It is a Court of Record and as such it has jurisdiction to hear matters relating to Māori land. “In all dealings the MLC recognises the unique relationship between Māori people and their land. The records form a substantial part of the whakapapa of all Māori.” (Chief Judge Isaac W.W. 2011). The court also assists Māori land owners to “promote the retention, use, development and control” (Māori Land Court website) of their land. It does that, among other things, by providing information and advice on land records and processing applications and enquiries. These services directly support the Te Ture Whenua Māori Act 1993 and the principles it embodies: the promotion of land retention and utilisation (Te Ture Whenua Māori Act 1993).

Despite these requirements, the over 200 Māori Land Court staff, including Judges, Court, Advisory and Registry staff, Information Officers, Principal Liaison Officers and Research Counsellors were poorly equipped to appropriately fulfil their mandate and support Māori land owners to actively exercise their ownership rights and utilise their land. The challenges the court and all Māori land owners are facing are quite unique (Chief Judge Isaac W.W. 2011):

- A special law relating to succession of ownership applies (Law Commission, 2001): Land is to be retained between blood relatives or under authority of the same hapū.
- There is an average of 85 owners per title ranging from a single owner or up to 629 owners per title.
- Often different interests are held by the same person under multiple names.
- Anecdotal evidence indicates that half of the interests are held by people who are deceased.
- Some owners do not know they are land owners.
- A number of owners do not have a known address.

Even though the 1.4 million hectares of land represent enormous economic assets, the threshold to activate this potential is very high. In September 2009 only 39 per cent of this land had a management structure associated with it, so much of it is likely to have been without the active management required to realise the full potential of its utilisation. No mortgages are granted unless a management structure is present. Even with a management structure in place, multiply owned land blocks are difficult to manage because agreement is hard to get. These factors make it extremely difficult for owners to manage or utilise their land.

In addition to these impediments, the lack of knowledge about the actual land further restricts the realisation of the potential of Māori land. An estimated 30 per cent of the land, for instance, has no means of access and is land locked. This is, however, not obvious from ownership records and not necessarily known by the owners.

In order to remedy this situation, the Māori Land Geographic Information System (MLGIS) project was initiated at the Ministry of Justice in March 2010. The planned Geographic Information System was intended to substantially lower the impediments for owners to unlock the potential of the land. It would be used by the Māori Land Court to provide better and more efficient services, and also by the Māori land owners themselves.

1.2 The Māori Land Geographic Information System Project

The Māori Land Geographic Information System project's goal was to establish a geographic information system highlighting Māori-owned land superimposed over a base map and providing comprehensive ownership information. The GIS was also intended to replace the existing application underpinning www.maorilandonline.govt.nz, that did not support viewing of Māori

freehold and customary land comprehensively and geospatially. Its interface was purely text based and very simplistic. It provided inadequate support for all stakeholder groups – those internal to Justice as well as the wider public including iwi, trusts and special interest groups.

The new system that went live in March 2011 was designed to make relevant information intuitively accessible and digestible. The intention was to unlock the existing authoritative information, enhance it further with cadastre information from Land Information New Zealand (LINZ), and enrich it with a map to visualise the land. Māori land blocks shapes were constructed from LINZ land parcels polygons based on their one-to-many relationship using a database procedure. The result was a seamless and comprehensive view of the authoritative information the Māori Land Court holds, embedded into relevant context.

The browser-based application, exposed to the general public, now allows all users to interactively navigate the map and see their land in context. It visualises key characteristics of the land in an easy to understand and natural way, but at the same time also offers the relevant facts, including address, a list of owners, land area, title, governance structures, land administrators, memorial schedules, trusts and minute book references, just to name a few.

The visual representation clearly shows the desired land blocks and their surrounding land, together with boundaries, adjacent land parcels, topography, land cover and access ways (Figures 1 and 2 give examples of some of these features).



Figure 1: Māori land blocks superimposed over a satellite image



Figure 2: The same land block superimposed over Google's street map

Prerequisite for the MLGIS project was the completion of the Māori Freehold Land Registration Project (Beehive website, 2010), which established a formal and transparent record of Māori land ownership and title that was not available previously.

The new www.maorilandonline.govt.nz website and the underpinning MLGIS application now answer these fundamental questions:

- Where is the land?
- Who are the owners of the land?
- What are the geographic features of the land?

The project has made it easier for activities such as exploring the best use of the land.

2 IMPLEMENTATION APPROACH

The predecessor project to the MLGIS initiative went to market to find an implementation partner for the introduction of the GIS. The vendor community, however, declined to respond to the Request for Tender (RFT). Reasons included too tight timelines, too little funding and lack of experience with the Open Source GIS technology that the RFT wanted to see considered. As a consequence, the project proposed to proceed with an in-house development heavily drawing on Open Source GIS components. Any solution including a vendor product would have involved another tender process, a time and cost prohibitive option. Requirements were cut back to allow the project to focus on a browser based system without any authoring or sophisticated GIS analysis capability.

While the value proposition question was sufficiently answered by the business case (Griggs, 2010), one question remained: Could an in-house development be successful at all? Given the lack of GIS experience at the Ministry of Justice, and in light of the unofficial vendor feedback that emphasised the immaturity of Open Source GIS technology, there were serious doubts among the Justice governance bodies. A pragmatic and efficient approach had to be identified that also allowed mitigation of the risk associated with the proposed in-house development. The answer that led to funding approval comprised two main elements:

1. Develop a prototype to prove validity of the architecture and the proposed Open Source technology as well as to explore whether the complexity would be manageable.

2. Adopt an agile (Beedle et al., 2001) software development approach and in particular the SCRUM methodology (Scrum Alliance), to establish a continuous feedback loop with all stakeholders and ensure full transparency of progress and issues at all times.

In the context of fiscal constraint the Ministry did not want to, and could not afford to, risk money on an approach with an uncertain outcome. Any development had to be innovative and efficient to ensure successful delivery within a reasonably short timeframe.

2.1 Prototyping the solution

A simple prototype was built by the GIS solution architect working part time on its implementation. It consisted of the following technology components:

- Oracle with its Locator component as spatial database.
- Apache Tomcat as light-weight application server and reference implementation of major Enterprise Java APIs.
- GeoServer as spatial middleware and map server, running on Tomcat. GeoServer is also the reference implementation of the OGC standards Web Map Service (WMS), Web Feature Service (WFS) and Web Coverage Service (WCS) (Open Geospatial Consortium website).
- OpenLayers as browser based mapping framework for rich mapping user interfaces.
- Sample datasets for polygon and point data (Survey Land Districts Titles, Territorial Local Authority Boundaries and Historical Sites), acquired from Koordinates.
- GoogleMaps as base map provider for satellite images, terrain and street maps.

The primary goal of building the prototype was to gain an understanding of the learning curve and complexity involved in developing an application with the above technology components. The prototyping showed that a simple mapping application based on the listed products predominantly required assembly and configuration of existing building blocks rather than a large degree of bespoke development. The prototype itself demonstrated the overlay of point and vector features over various GoogleMaps base maps. The web application included a navigable map with three layers that could be turned on and off. Data was retrieved from the Oracle database, image tiles rendered by GeoServer, exposed as standards compliant Web Map Service and consumed by OpenLayers to display the actual map and navigation controls.

While one of the likely challenges, the integration of MLIS land block data sets with LINZ land parcel data to construct the Māori Land Block layer, remained unaddressed at this stage, prototyping did confirm the following:

1. The chosen Open Source geospatial components GeoServer and OpenLayers are fit for purpose. A simple web mapping application with GoogleMaps and a few additional layers can be developed with minimal effort and configuration, deploying the components in an out-of-the-box fashion.
2. Specialised GIS expertise is unlikely to be required. Good JavaEE and JavaScript knowledge and broad development experience is more important.
3. Low level implementation can be time consuming without experience in the relevant products. Being able to tap into external expertise, when needed, is desirable.
4. Documentation in the traditional form (manuals, text books or articles) on OpenLayers and GeoServer is scarce. Utilising OpenLayers/GeoServer mailing lists, however, provided quick responses and reasonably efficient problem resolution.

The prototype application was used to demonstrate to all stakeholders what a user interface could look like and provided users with a hands-on understanding of what the planned GIS would entail. It was instrumental to getting buy-in from the project sponsor and relevant governance bodies. It effectively paved the way for funding approval for the actual development.

In summary, the prototyping achieved two main outcomes. Firstly, the proposed architecture including the technology stack was assessed as robust and fit for purpose. Secondly, it built confidence in that in-house development was feasible.

2.2 Software development process

Despite the fact that prototyping the solution proved the feasibility of an in-house implementation, there were still challenges to overcome. Even though product experience with OpenLayers and GeoServer was not considered critical, local spatial professional solution provider eSpatial (eSpatial website) was brought onboard as risk mitigation to exercise high-level quality assurance and validate any architectural decisions.

The actual project team consisted of a project manager, a business analyst (part time), a solution architect (part time), 2 developers and a tester. The project team adopted SCRUM as its agile software development methodology. The key practical implication for the project was the use of an iterative approach that started with a simplistic yet working core that evolved over time. Instead of spending a significant amount of time in upfront design, each iteration planned, designed and implemented new business features. Each iteration ended with a fully functional application. Breaking the overall application into smaller manageable chunks that were tackled individually had a number of advantages. Short iterations:

- allow adjustments to be made quickly and easily without a lot of rework,
- mean changes in priority or scope can be easily accommodated,
- ensure planning and design stays in touch with reality,
- end with a functional application which means that progress is demonstrable and tangible,
- provide risk mitigation as no risk can build up and become unmanageable, and
- allow issues to be addressed and resolved immediately.

The continuous involvement of a business representative who provided guidance and had the authority to make decisions was pivotal to the success of the project. It allowed the implementation to progress smoothly and gave the business constant insight and confidence in the development. Integral to the continuous evaluation approach was full transparency between members of the project team, including the business representative. This allowed timely assessment of the robustness of any assumptions made and of development progress. A consequence of this was the natural and ongoing application of a high degree of rigour and discipline without incurring a noticeable amount of overhead or bureaucracy.

3 TECHNOLOGY SOLUTION

3.1 Technology and system architecture

First and foremost, any solution has to conform to the Ministry's Information System Strategic Plan and Enterprise Architecture. As such, it has to adhere to the standard technology platform and implement open standards and interfaces. Beyond this, spatial applications at the Ministry are guided by the NZ Geospatial Strategy (New Zealand Geospatial Office, 2007).

To mitigate the risk posed by the lack of GIS experience at Justice, the solution architecture had to support two goals. Firstly, deliver the simplest solution that is fit for purpose. Secondly, maximise re-use of existing and proven components as well as available knowledge and expertise.

Additionally, the new MLGIS system was also designed according to these architectural principles:

1. Ensure extensibility by using loosely coupled components that can be individually enhanced or exchanged when required.
2. Allow the seamless addition of new layers, different imagery, authentication and fine-grained access control.
3. Support the ability to evolve to future and more sophisticated GIS activities across the enterprise.
4. Promote interoperability through open standards and interfaces (for example as defined by the Open Geospatial Consortium). This allows integration of the existing tools (for instance ArcGIS Desktop) and inclusion of other layers from external data providers.

5. Achieve usability, achieved by the re-use of user interface components the majority of users are already familiar with. This includes the browser itself as well as GoogleMaps base maps and a GoogleMaps-like navigation and interface style.

Consequently, the major building blocks within the MLGIS application represent well known, well-established and widely adopted technology. Figure 3 provides a comprehensive overview of the technologies used.

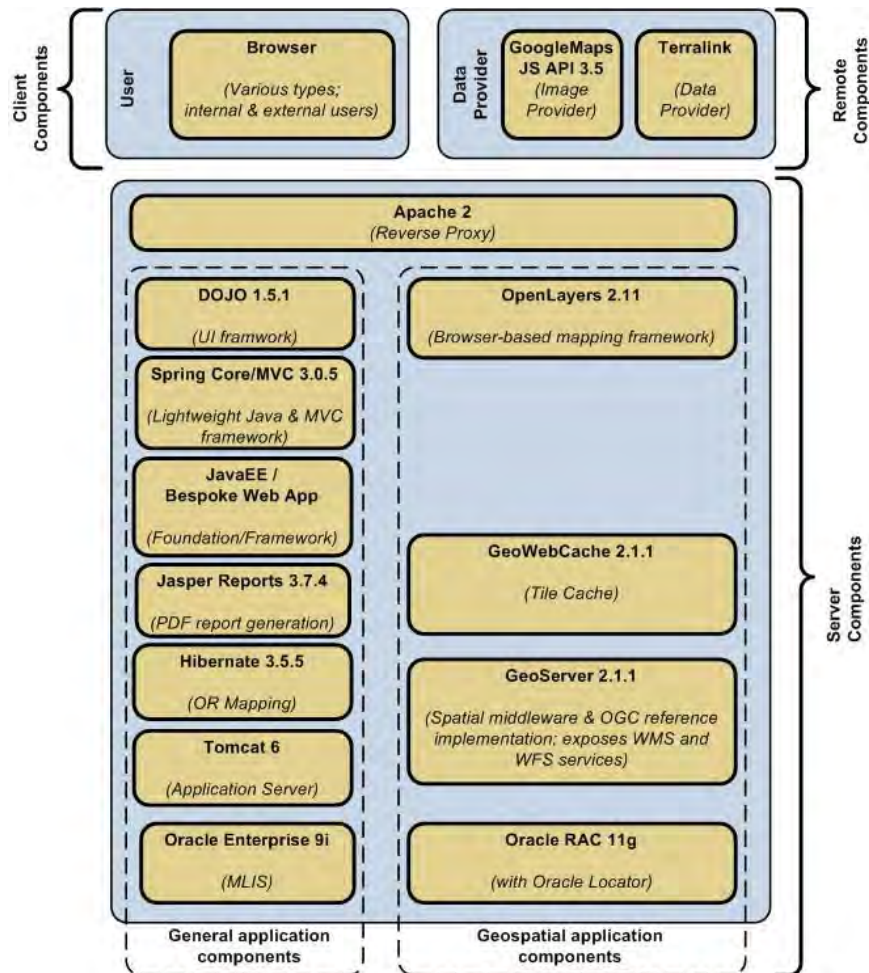


Figure 3: Technology components used in the implementation of the Māori Land Geographic Information System (MLGIS)

All technology components, with exception of GeoServer/GeoWebCache and OpenLayers, were already in use at the Ministry. They were well understood and existing skills were available. Figure 4 depicts a high level representation of the overall system architecture.

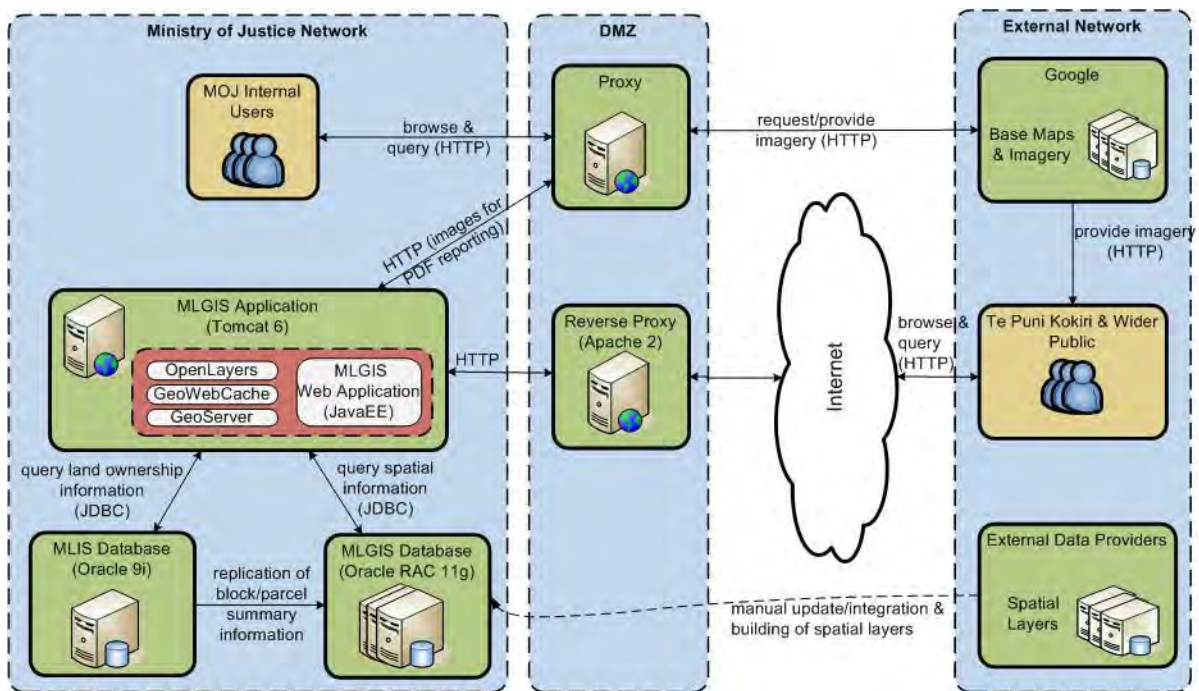


Figure 4: High level system architecture of the Māori Land Geographic Information System, showing major components and their interactions.

3.2 Performance as key requirement

In addition to fulfilling all functional needs, performance was a key requirement. The application had to be responsive in order to guarantee a smooth user experience and to become accepted by internal staff and the wider public alike. The following architectural decisions were applied to support this performance objective:

- Database level integration.
MLIS block definitions are used to construct land block shapes from LINZ land parcel polygons on a monthly basis, when a new data feed is received from Terralink. The layer is pre-calculated; nothing is assembled dynamically. This promotes speedy responses.
- Bespoke functionality and the degree of processing done at runtime are kept to a minimum. Interaction with the map interface is underpinned by OpenLayers interacting with GeoServer which in turn retrieves data from the Oracle database. No intermediaries slow the request/response process down.
- Image tiles are pre-computed and cached to support fast rendering of the map and to reduce image tile generation load on GeoServer.

The runtime environment for hosting the MLGIS application is a single virtual machine running SuSE Linux Enterprise Server 11. It is allocated 8 GB of RAM, and two 2.5 GHz virtual CPUs. Tomcat application server is allowed 4 GB of heap memory. The underlying hardware is a HP BL460c G1 blade server. The actual image cache consumes, on average, less than 50 GB of storage for 3 layers. It is purged on a monthly basis when a new data feed from Terralink is integrated. The layers themselves, and supporting data like spatial indexes, require less than 5 GB of database space.

Based on this configuration and almost five months of operation, the website has so far robustly supported the following load without any issues:

- At peak time 8 hits per second incurring a traffic volume of 25.5 GB/month (excluding base imagery from Google, which is directly routed from Google to clients).
- On average: 3 hits per second, resulting into 21 GB of traffic per month.

Figure 5 illustrates the average daily distribution of requests.

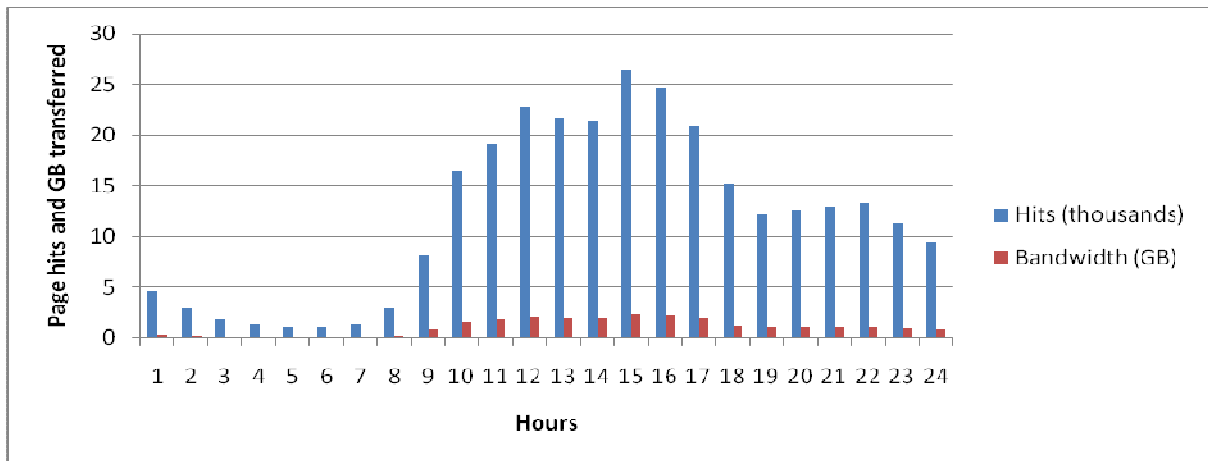


Figure 5: Average monthly distribution of page hits for www.maorilandonline.govt.nz across the day.

In order to minimise the tile cache storage requirements and avoid caching of rarely visited tiles, a smart cache strategy was adopted. The simplistic approach of caching all image tiles for all 21 zoom levels and all layers would have resulted into an enormous storage requirement in the region of many Terabytes. Application of the following two strategies helped to mitigate the actual storage requirement:

1. Combine Māori Land block layer with MLC District Boundary layer (using a GeoServer layer group).
This effectively removed a full layer and its associated storage requirement. Another benefit is increased performance as the runtime overlay of these two layers is now unnecessary.
2. Selective caching.
Not all zoom levels are fully cached. Instead, the cache is seeded (pre-populated) only for the most popular zoom levels. Instead of caching the entire layer, only tiles within the bounding box representing New Zealand are stored.

3.3 Technology challenges

Overall, all building blocks proved to be reliable and mature. Tomcat, GoogleMaps and Oracle caused no issues. Challenges the project team did have to deal with revolved around:

1. Cross-browser compatibility
2. Lack of depth in documentation
3. Immature peripheral features in GeoServer and OpenLayers

Overall, cross-browser compatibility, a core requirement for the public facing website, represented the main technology challenge. Government web standards require testing against A-grade browsers (The Department of Internal Affairs, 2009), which include Internet Explorer versions 6 to 9, Firefox v3 to v5 and Chrome. Achieving a consistent, robust and smooth user experience across all major browsers and multiple recent versions required a substantial amount of testing. Each browser had its peculiarities, with Internet Explorer showing particularly poor performance in rendering complex polygons. Internet Explorer and Firefox are the two approved browsers for the Ministry of the Justice with Internet Explorer 6 as the default browser in the Ministry's standard operating desktop environment. As such, both needed to be properly supported.

While the GoogleMaps API is thoroughly documented, GeoServer and OpenLayers documentation covers only the essentials well. Detailed and current documentation addressing more advanced features is not consistently available. Little is offered in terms of tutorials or textbook style explanations that go beyond the basics. At the time of development no books and very few useful online resources were offered that covered more than the first steps. The project

team found, however that the GeoServer and OpenLayers mailing lists are quite active with useful discussions and usually fast responses to any questions posted.

In terms of maturity, some features like GeoServer's *app-schema* extension to provide *complex feature* functionality and OpenLayer's drawing tools were not as robust as expected. The only real issue, however, was OpenLayers' inability to properly overlay the Māori land block layer for Chatham Island polygons, which was tied to the fact that the land blocks were across the International Date Line. The analysis performed by project staff resulted in a patch for OpenLayers that fixed the problem. The patch was submitted to the OpenLayers team and eventually included into the OpenLayers code base where it became part of a subsequent release.

4 OUTCOME: FULFILLING THE CUSTOMER NEED

The www.maorilandonline.govt.nz website, together with its underpinning MLGIS application, provides services to a diverse customer base that includes:

- Māori Land Court staff,
- Other Ministry of Justice business units, including Office of Treaty Settlement and Waitangi Tribunal,
- Other government organisations, for instance Te Puni Kōkiri,
- Māori land owners,
- Trusts,
- Iwi,
- Private law community.

While the use cases vary among that customer base, they share a common interest in Māori land and the desire to acquire as much relevant information as possible to support their specific goals. Regardless of what those goals are, the geospatial enablement of the [maorilandonline](http://www.maorilandonline.govt.nz) website and its data represents a major leap in terms of accessibility to the information and ease of use. The additional benefits the website now offers are highlighted in Figure 6. This shows the details page of a land block together with associated information on the land parcels the block comprises (shown in the info bubble on the map).

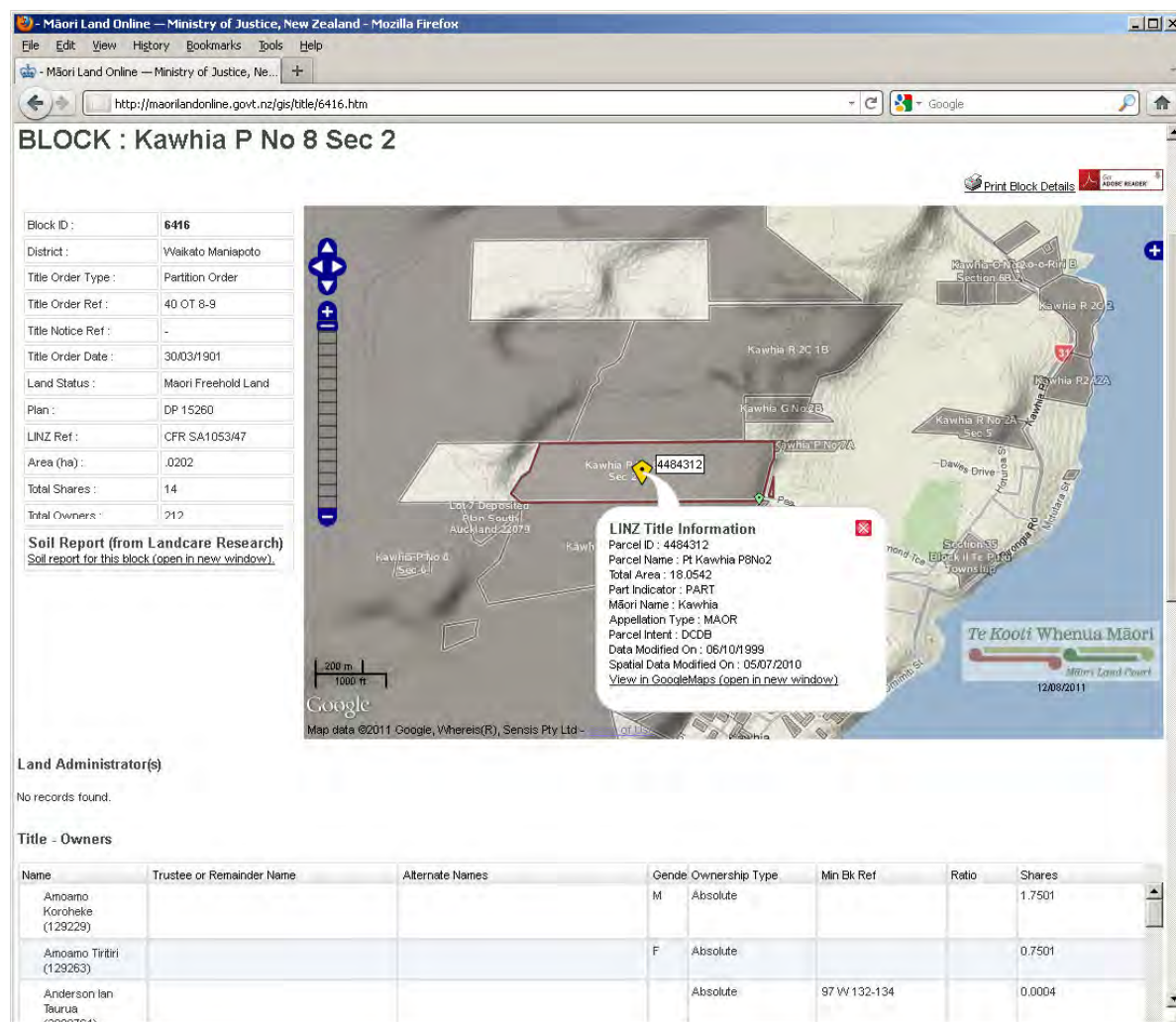


Figure 6: Māori land block with title and LINZ land parcel information

The key benefits that apply to all users are:

- Access to ownership data that was previously almost inaccessible is now quick and easy. Information at your fingertips rather than information obtained through visits to Māori Land Court offices is a radical paradigm shift for this data. Interested parties can quickly check whether they own a piece of land or what their specific interests are, simply by searching by owner name, trust name, block name or LINZ reference.
- A comprehensive set of information is available: this combination of data from the Māori Land Information System with LINZ cadastre means an immense increase in usefulness.
- The visualisation of location information on an easy to understand and easy to navigate map is intuitive and efficient. Browser-based delivery with GoogleMaps means the relevant content can be directly accessed by the target audience without requiring any installation of software or training. The map also offers a wealth of valuable context: it shows access ways, adjacent land parcels, roads, waterways, and elevation levels. This is helpful for assessing the economic potential of the land.
- The application fully supports Te Reo Māori, appropriate to the users.

Benefits specific to the Māori Land Court are:

- Data quality can be improved as faulty data sets are easily identified using the map to visually inspect ownership records. Since the system's availability, almost 6 per cent of records, equating to 1938 changes, have been amended based on findings made possible by the new visual interface.

- In at least one instance a remote site visit became unnecessary as the Judge consulted the system to make an informed decision thus saving travelling cost for all parties involved.
- The Māori Land Court can provide better quality of service through more accurate, more detailed and more comprehensive responses to enquiries.
- The expectation is for fewer adjournments and re-hearings due to better common understanding of the matter heard in court. More comprehensive and unambiguous information will allow more robust decision making.
- Citizen self-service will lead to fewer enquiries, reducing the manual effort for MLC.
- Establishment of an authoritative and current information base has promoted Justice-internal re-use and has already allowed support of Waitangi Tribunal and Office of Treaty Settlement with tailored data extracts in a number of cases. This enables a more accurate and more efficient delivery of services in those groups.

5 OPERATION AND SUPPORT

The Ministry of Justice is already managing a heterogeneous and complex IT environment resulting from the merger between the Department for Courts and the former Ministry of Justice. Due to the number of different technologies and a mix of bespoke and off-the-shelf products with greatly varying technical dependencies, any new software products introduced have to undergo rigorous scrutiny. The screening assesses how well they fit the Enterprise Architecture and the standard technology blueprints. This ensures any products brought into the environment can be properly supported and managed. Support includes monitoring, backup, and helpdesk as well as second and third level support. It covers rollout of patches and new versions. Browser-based web applications are preferred as this avoids installation on the users' desktops and any adverse impact on other desktop applications.

As a web application, MLGIS fits these criteria objectives well. The underpinning application servers and databases are already managed by the Ministry's Enterprise Systems Team. Releases follow a standard automated procedure and well defined change management process.

Integration of the new application into the Ministry's portfolio was seamless. No standards had to be compromised and no exemptions had to be made.

Since the chosen JavaEE development technology fully aligns with the Ministry's standard platform, handover to the normal business and third level support teams did not involve any extra steps or require any training. The vendor relationship with eSpatial as external service provider continues to be actively maintained so that their expertise can be leveraged if need be.

Overall, even though GIS on an enterprise level was a first for the Ministry, the introduction into its operational environment not only happened seamlessly, it also leveraged existing assets and experience to the maximum. Compatibility and alignment with technology standards and Enterprise Architecture continue to ensure sustainable operation without any additional overhead. The impact of adding a major new application into the Ministry's portfolio was minimal.

Keeping up with Google's fast paced development of their Maps product has proven to be a challenge. Even without any business driven changes, the MLGIS application has to be periodically assessed and updated as part of the technology currency programme to avoid breaking the interface to GoogleMaps.

6 CONCLUSION

The experiences gained through the project lead to the following conclusions:

- Adding even a simple mapping element to visualise location information represents a significant leap in terms of boosting the users' understanding of this specific information. It enabled a much more natural and intuitive way of processing this data. It also allows for easier quality assurance as any anomalies immediately stand out. It is only logical to apply the benefits of traditional maps to the electronic delivery of information.

- A web mapping application that delivers substantial benefits can be assembled with reasonable effort and little bespoke development. It benefits from a high degree of re-use of standard building blocks which mitigates the implementation risk.
- Robust and mature Open Source building blocks are readily available for delivery of a significant degree of functionality thus allowing a cost effective implementation.
- Open Source GIS components are inherently designed for extensibility and interoperability so that they can be flexibly combined with other already existing or future elements of a GIS infrastructure thus avoiding vendor lock-in.
- Developers require no specialised expertise in GIS, cartography or surveying.
- Users do not require any formal cartography or surveying knowledge. The browser-based delivery with GoogleMaps or a GoogleMaps-like interface allows them to be immediately effective, without the need for any training.
- Many web mapping components are based on standard technologies such as JavaScript or Enterprise Java. Therefore they can be almost seamlessly integrated into an existing IT infrastructure. No specialised or dedicated servers, infrastructure components or deployment and maintenance models are required.

In summary, the new website was not only well received by Justice staff, it also achieved the goal of becoming an enabler that allows public users to become better and more easily informed. The development team and IT operations were impressed by robustness and maturity of the technologies used. As such, we are confident the system can serve as the basis of GIS activities at the Ministry of Justice for the foreseeable future.

REFERENCES

Beedle M., Bennekum A., Cockburn A., Cunningham W., Fowler M., Highsmith J., Hunt A., Jeffries R., Kern J., Marick B., Martin C., Schwaber K., Sutherland J., Thomas D. (2001). Agile Manifesto <http://agilemanifesto.org/>

Beehive website (2010). Māori Freehold Land Registration Project <http://www.beehive.govt.nz/release/project-draws-successful-close>

eSpatial website <http://www.e-spatial.co.nz/>

GeoServer <http://geoserver.org/>

Griggs Jill (2010), Māori Land Information System (MLGIS) project Business Case

Chief Judge Isaac W.W. (2011). Judges' Corner: Māori Land Today <http://www.justice.govt.nz/courts/maori-land-court/judges-corner>

Koordinates <http://koordinates.com/>

Law Commission (2001). Māori Custom and Values in New Zealand Law, Study Paper 9 http://www.lawcom.govt.nz/sites/default/files/publications/2001/03/Publication_112_288_SP9.pdf

Māori Land Court website <http://www.justice.govt.nz/courts/maori-land-court/about-us>

maorilandonline website & MLGIS system <http://www.maorilandonline.govt.nz/>

New Zealand Geospatial Office (2007), New Zealand Geospatial Strategy
<http://www.geospatial.govt.nz/sites/default/files/assets/Geospatial-Strategy/nz-geospatial-strategy-2007.pdf>

Open Geospatial Consortium website
<http://www.opengeospatial.org/>

OpenLayers
<http://openlayers.org/>

Scrum Alliance. SCRUM software development methodology
<http://www.scrumalliance.org/>

Te Puni Kōkiri (2009). Ngā Rawa ā Iwi: Tribal Assets
<http://www.tpk.govt.nz/en/in-print/our-publications/fact-sheets/tribal-assets/download/tpk-tribalassets-2009-en.pdf>

Terralink. Core Record System dataset
<http://www.terralink.co.nz/gis-data/core-record-system-crs/>

Te Ture Whenua Māori Act 1993
<http://www.legislation.govt.nz/act/public/1993/0004/latest/whole.html>

The Department of Internal Affairs (2009). Government Technology Services. New Zealand Government web standards 2.0
<http://webstandards.govt.nz/new-zealand-government-web-standards-2/>

Starting to Talk about Place

Stephan Winter, Allison Kealy, Matt Duckham, Abbas Rajabifard, Kai-Florian Richter

Department of Infrastructure Engineering
The University of Melbourne, VIC 3010, Australia
{winter, a.kealy, mduckham, abbas.r, krichter}@unimelb.edu.au

Tim Baldwin

Department of Computer Science and Software Engineering
The University of Melbourne, VIC 3010, Australia
tbaldwin@unimelb.edu.au

Lesley Stirling

School of Languages and Linguistics
The University of Melbourne, VIC 3010, Australia
lesleyfs@unimelb.edu.au

Lawrence Cavedon

Computer Science and Information Technology
RMIT University, Melbourne, VIC 3001, Australia
lawrence.cavedon@rmit.edu.au

Daniela Richter

Institute of Photogrammetry and Remote Sensing
KIT Karlsruhe, Germany
daniela.richter@kit.edu

ABSTRACT

Place descriptions are a common way for people to describe a location, but no current software tools are smart enough to understand them. This way, emergency call centres are risking lives, postal services are wasting billions of dollars per year through problems in addressing, and users of navigation or web services are frustrated about restrictive interfaces or prolonged search. This paper reports about a new project and the first steps of interdisciplinary research to automatically interpret human place descriptions. It will develop methods to capture place names with their meaning—their true location—for smarter databases and automatic interpretation procedures. The acquired knowledge will be an important step forward for data custodians and for service users.

KEYWORDS: spatial information, spatial cognitive engineering, spatial data infrastructure, language technologies, gazetteers

1 INTRODUCTION

Place descriptions are a common way for people to describe a location, or answer a “where” question. However, no current tools are smart enough to understand place descriptions in common language. Think of “I am opposite to the library”, or “I left the car park 30 minutes ago hiking North”, or “let’s meet at the usual place” to illustrate the challenges in the automatic interpretation of these expressions. Place descriptions occur in many different contexts, including emergency calls, local

search, and route directions. Each of these examples is nowadays supported by automated services, yet the sheer ubiquity, variety and complexity of human place descriptions exceeds the capacities of available tools. Emergency call centres still rely for safety reasons on human operators to interpret the place descriptions of callers; web services providing local search are still ignorant to most spatial reasoning (Winter and Truelove, 2010), and navigation services still frustrate their users with interfaces that restrict their input to predefined formats (Winter and Wu, 2009).

This paper presents a novel, interdisciplinary approach to tackle this problem. It reports about preliminary work and first results of a project *Talking about Place* that aims to enhance the content and functionality of place name databases by tapping human knowledge (ARC LP100200199). The project will collect large corpora of geotagged human place descriptions, and interpret them to reveal the use of place names, the spatial extent of places, and the relations between places.

The preliminary work so far consists of the collection of about 2000 place descriptions by a mobile location-based game between August 2010 and March 2011. The collected user-generated content reveals already interesting properties of place descriptions, and research challenges.

Place name databases (gazetteers or address files) form an essential backbone of any national spatial data infrastructure. Results of this project will support the data custodians of such databases to enrich their databases (e.g., PSMA's *Geocoded National Address File* G-NAF, or the Committee for Geographical Names in Australasia's *National Gazetteer of Australia*). Both acquired knowledge and richer place name databases, are essential for the design of more intelligent services. Hence, results will also support critical users of place name databases (e.g., the operator of Victoria's emergency call centre, ESTA). All three agencies are partners in this project. In future, enriching our data and knowledge about place will form an essential key for facilitating free dialogs about locations, e.g., in emergency calls, in navigation, or in local web search.

2 PLACE NAMES AND PLACE DESCRIPTIONS

Place plays a central role in human geography, urban planning and related disciplines (e.g. Tuan, 1977; Vanclay *et al.*, 2008). Place is also the way people perceive, memorize, reason and communicate about space—they rarely use geometry (or metric) expressions, but refer to named and unnamed places and qualitative spatial relations between them (Freundschuh *et al.*, 1990; Landau and Jackendoff, 1993; Levinson, 2003; van der Zee and Slack, 2003). Human place descriptions are linguistic expressions and hence externalisations of what is in the minds of people.

Human concepts of places are hard to formalize, and accordingly place is often absent from gazetteers of geographic names, in spatial databases, or in geographic information systems. Gazetteers collect (typically authoritative) names of some categories of geographic features, but not of others, and they describe the spatial semantics of these names by a point (Hill, 2006). Linking place names in gazetteers with polygons from other data sources is hampered by the fact that multiple data sources exist of different custodianship, e.g., cadastre data or topographic data, and their geometries do not match. Also, current spatial databases are built from geometric data models—points, lines, and areas—rather than from the semantics of named places, and cannot cope with the indeterminate nature of many places' boundaries (Burrough and Frank, 1996).

Place descriptions are complex expressions referring to places either by names of places ("Southern Cross Station") or by the names of their category ("the train station"). Since they are complex they also link different references by spatial relationships, either explicitly ("the hotel opposite the train station") or implicitly ("Carlton, Victoria" implying Carlton *in* Victoria). Generally the spatial relationships resolve ambiguities (Wu and Winter, 2011). For example, Carlton is disambiguated from the other 24 occurrences of this place name worldwide (according to the Getty Thesaurus of Geographic Names) by the qualifier "in Victoria". However, qualitative spatial relationships come in different flavours—distance relationships ("near"), orientation relationships ("left of", "opposite"), and topologic relationships ("in", "next to")—and reasoning with all these relationships is quite challenging (Cohn and Renz, 2008). Their qualitative nature resists metric

interpretations, and yet, people rarely use metrics to quantify these relationships (Denis *et al.*, 1999).

As a result, there is a fundamental gap in communications about places between humans and computers. Conflicts in address validation, local search, navigation services, location-based services, emergency calls and emergency management, or by colloquial or vanity addressing cost the economy billions of dollars and can even cost lives. In one tragic case, in December 2006 David Iredale tried multiple times to satisfy emergency operators' demands for details about his location before he died in the Blue Mountains, and similarly place descriptions in urban environments can delay an emergency response if the 000 caller does not know the official address of an incident.

One response of the geographic information community to the lack of place in spatial databases is standardization. They may be national standards, as for example the Geocoded National Address File (G-NAV). Also standardized exchange formats for location information exist, e.g., OpenLS (OGC, 2008).

Standardization is necessary but not sufficient. A recent think tank with Australian stakeholders revealed gaps between state-of-the-art systems and the in-depth knowledge required to understand human place descriptions, or to enter into a dialog about place descriptions (Bennett and Winter, 2009); see also (Davies *et al.*, 2009). Significant obstacles to achieving automated understanding of place descriptions include vernacular place names (or the incompleteness of place databases); the lack of spatial semantics of place names (spatiotemporal extent and vagueness of extent); and lack of tools to represent and reason about spatial relations between places, such as containment or nearness.

The project *Talking about Place* will address these obstacles, generating and integrating new knowledge using the resources from a range of related disciplines, and translating this into concrete applications for automated understanding of place descriptions.

3 TALKING ABOUT PLACE

Talking about Place is a research project funded by the Australian Research Council through its Linkage program (LP100200199). It commenced at the end of May 2011. The partners in this project are the University of Melbourne, RMIT University, PSMA Australia Ltd., the Committee for Geographic Names in Australia represented through two of its state sponsors (Victoria and New South Wales), and ESTA, the emergency call-taking and dispatch service in Victoria. The academic expertise in this project spans from spatial information science to computational linguistics and linguistics.

Talking about Place aims to significantly improve current understanding of place, and techniques for dealing with place, to enrich place name databases, to improve human-computer interaction about place, and to support Australia's spatial information sector in providing smarter products and services. The project in particular searches for contributions into:

- methods for large scale automated collection and storage of place names together with their spatial semantics and related metadata, using purpose-built mobile location-based gaming (crowd-sourcing) and emergency communication transcripts;
- enriching place databases further by the inclusion of prominence, dependencies and other spatial relationships that can improve place name search functionality; and
- methods for automated interpretation of complex natural descriptions of place, including resolving ambiguity and underspecification.

The approach and contribution of the project are by and large aligned with current mainstream research on tapping human knowledge by crowd sourcing (Surowiecki, 2004), citizens as sensors (Goodchild, 2007), user-generated content (Haklay and Weber, 2008; Krumm *et al.*, 2008), or volunteered geographic information (Elwood, 2008). This emergent field is held back by some challenges, which are mostly in the area of spatial semantics. User-generated spatial data consists of a combination of sensor data (e.g., coordinates) and user-added semantics (e.g., place descriptions). While databases are good at collecting and interpreting sensor data, the management and use of the user-added semantics is still in its infancy (Richter and Winter, 2011). This is the area where the current project will focus on.

Up-to-date information about the project, further resources and links can be found online (www.telluswhere.net).

4 TELL US WHERE

In preliminary work for *Talking about Place* we have developed, implemented and applied a location-based mobile game for collecting a larger corpus of geotagged human place descriptions: *Tell us where* (short for “Tell us where you are”). A mobile location based game can form an efficient way to tap human knowledge. It potentially reaches a large number of people, records interactions, and can analyse and learn (Richter and Winter, 2011). *Tell us where* (Figure 1) was a web-based location-based mobile game collecting human place descriptions from the players of the game. The game idea was deliberately simple. Wherever they are, players can submit an entry to the game in the form of a textual description of their location. The place descriptions are stored at the web-server together with the position of the mobile device, which is automatically obtained by GPS and confirmed by the player on a map. As an incentive each submitted place description had the chance to win a gift voucher. Tell-us-where was optimized for the web-browsers of different current smartphone operating systems; shown in Figure 1 is the appearance in Apple iPhones’ Safari browser.

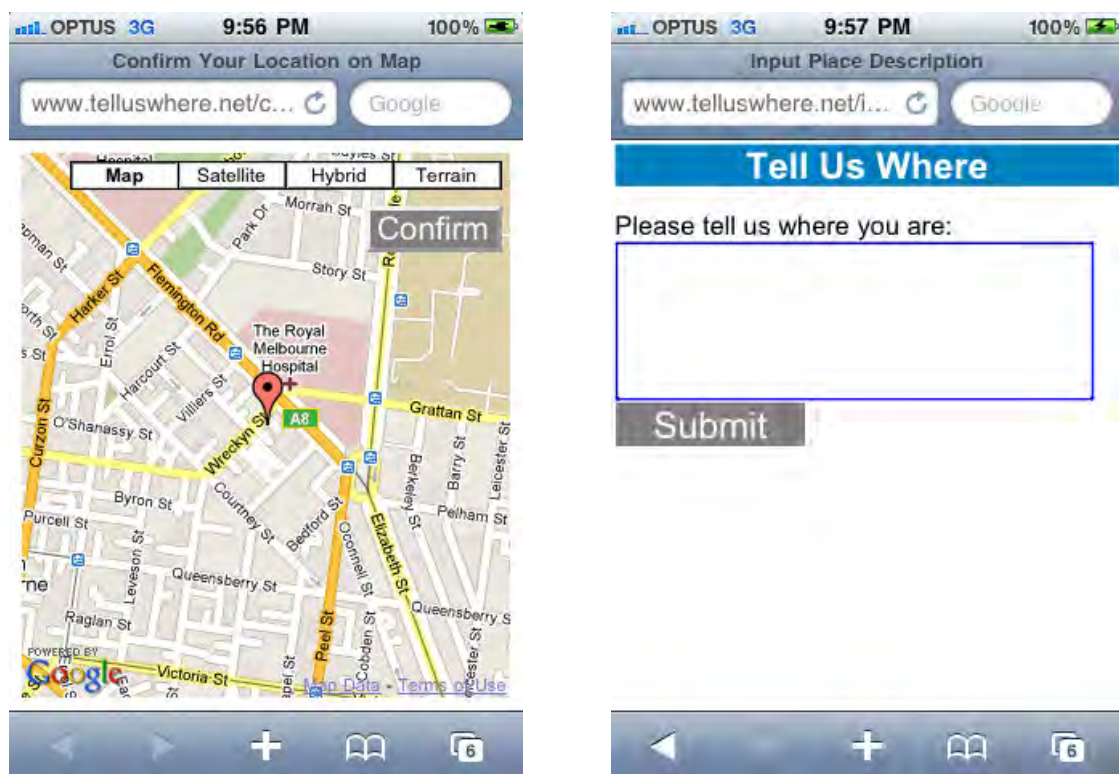


Figure 1: *Tell us where* started with a self-localization of the players (left), and then asked for a verbal description of where the players are (right).

The game was covered by the media and quickly attracted attention. We collected a large number of place descriptions over the six months of running the game. These place descriptions are distributed all over Victoria and beyond, but the majority is concentrated in Greater Melbourne (Figure 2).

The collected place descriptions are now being investigated (see next section for first results), but the game also allowed us to make a number of generic observations (Winter *et al.*, 2011).

The most important observation is a relative lack of context in our game. Place descriptions are context-dependent. For example, I would describe my location differently to a emergency call centre operator (“corner Barry St / Grattan St”), my colleague I intend to meet here (“Baretto cafe”), or my friend living overseas (“in Melbourne”). Hence to interpret the collected data, knowing (or

designing) the context is crucial. Generally, context is created by a role for the player in the game and by a purpose for specific actions. In *Tell us where* the context was set only by (a) a confirmation of the self-localization by GPS on a map, (b) the question: “Tell us where you are: [textbox]”, and (c) the purpose of pastime and a chance to win a voucher. This specification actually underdetermined the context. Players could enter place descriptions imagining any game role or purpose, and this contributed to the variety of the results. One additional and recorded factor to infer the context imagined by the player was the map zoom level the players had chosen for their self-localization.

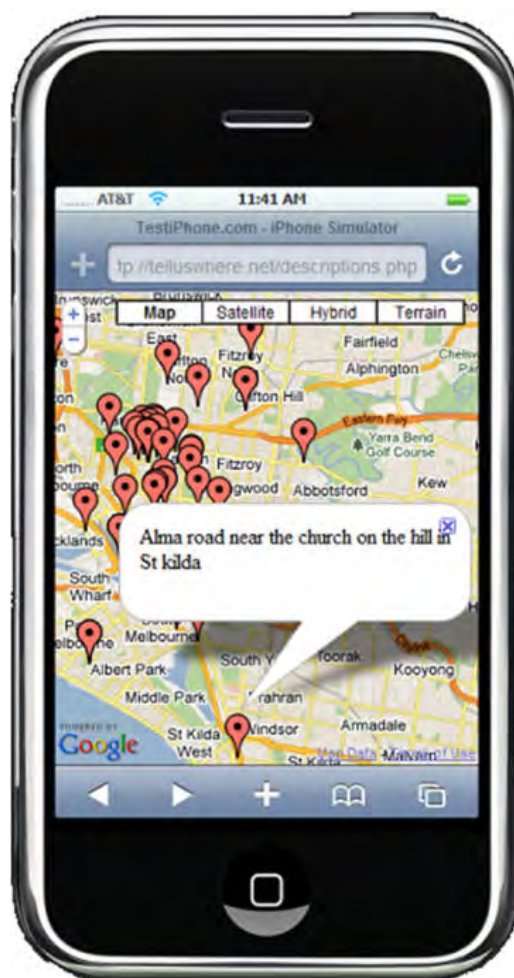


Figure 2: *Tell us where* showing some of the locations of place descriptions collected by the end of Week 1.

Tell us where was originally constrained to be played within the boundaries of Victoria. Within this area we observe an inhomogeneous sampling distribution, influenced by population distribution, mobile internet coverage, and the social networks through which the game was promoted. Depending on the type of knowledge the researchers are interested in an inhomogeneous distribution might be acceptable. If a more regular distribution is desirable other filter strategies can be implemented (Dykes *et al.*, 2008).

Not everybody plays location-based mobile games seriously. Location-based mobile games need other filter mechanisms to keep the collected data sets clean. Especially the opportunity to win vouchers made *Tell us where* vulnerable for fake participation, and we implemented several mechanisms to test in real-time whether the provided descriptions were actually valid place descriptions. Initially, one of these mechanisms was filtering out place descriptions that did not contain a place name found in a gazetteer of Victoria. These mechanisms constrained the context as well. They actually rejected some place descriptions that would have been acceptable by the rules; therefore, they were removed in a later stage of the game.

Tell us where's game aspect is minimal (restricted to the fortune element), and the immersion effect is accordingly limited. Nevertheless, the lessons learned are applicable to other location-based mobile games, despite the relatively small and local set of results.

5 PRELIMINARY RESULTS

A total of 2221 place descriptions have been collected within about seven months (15/8/2010 – 24/3/2011) of running *Tell us where*. To investigate granularity of place descriptions with respect to spatial and semantic context as well as to the chosen zoom level of the mobile game, we looked at different entities: the participant/player, the mobile game, as well as the described places and self-localization. Here we are reporting on preliminary findings, an in depth linguistic, grammatical analysis is in progress.

All records were directly attributed with a record number, latitude and longitude, the zoom level, the date, and an indication as to whether the submitted place description actually won a voucher. Filtering out datasets included the removal of datasets with empty strings or random characters (3.7% of total records) as well as double entries (same coordinates, date, and description / 3.7% of total records). After this filtering, we obtained 2057 distinct descriptions for 1915 unique places (coordinate pairs). In addition to the directly recorded attributes we derived text length, which is the number of characters in a given description and which serves as an indicator for the complexity or degree of specificity in our preliminary investigation.

Place descriptions were labelled with category information according to various criteria. Our preliminary analysis uses for this purpose a subset of 154 randomly selected records (maximal 10 per zoom level), in order to investigate different approaches, which will be explained in the following.

Looking at the player, initially we found three different methods as to how people dealt with the question “Tell us where you are?”: In 75% of cases in our subset they reported their positions using addresses, landmarks and spatial relations, for example “I am at the corner of Swanston and Bourke street”, or “near town hall” (all citations of collected place descriptions are shown here in their original spelling). 7% indicated that they were actually moving / getting around somewhere, e.g. “walking down Collins street”, “heading to emerald off road”, and 2% gave a rather complex route description, either how they have reached the respective location and/or how they will proceed to their destination, or, they tried to give detailed instructions allowing other people to find the place. All these methods incorporate official place names found in the Gazetteer of the State Government of Victoria, Google Maps or Microsoft Bing , including also vernacular place names and spelling errors or abbreviations that we manually checked, such as “st.kilda footy club” (St Kilda Football Club), “royal melb hospital” (Royal Melbourne Hospital), or “sa” (South Australia). Despite the fact that search engines may provide common spelling variations this task is challenging especially regarding vernacular places. The remaining 16% of records did not contain any specific geographical information, i.e., neither official nor vernacular place names. However, 13% involved semantic information normally related to functions of places like “the beach”, “at a little cafe of a side street”, to personal context (“at work”, “my house”), or business names and trademarks (“Anz Bank”, or “Apple store”). We filtered out the remaining 3% that contained no identifiable spatial information, this means records with information such as “bbq in the sun” or “i am lost”. The distribution of the different methods is outlined in Figure 3.

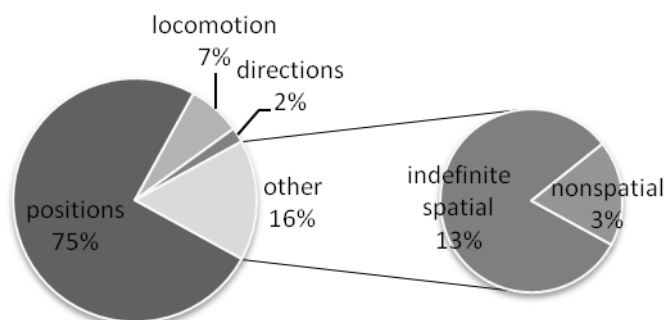


Figure 3 Classification of description types (subset of 154 records).

Analysing place descriptions of our subset, we found about 45% of reports that contain hierarchical structures referring to particular locations, for example, “I am on the banks of a pond in the botanical gardens”, “Esta support office 6/215 spring Street”, “131 Pelham st, Carlton, 3053, Victoria” or “building 174, Grattan St, Melbourne, Victoria, Australia”. This supports findings from literature in spatial cognition suggesting that people normally describe places in hierarchical ways (Shanon, 1979; Paraboni *et al.*, 2007). However, participants within Melbourne often did not explicitly tell they were in Melbourne, which is most likely related to the setup and promotion of the game. Only 11% of all records mention Melbourne in their descriptions, including all (implicit) cases such as “Melbourne Museum”, “Melbourne Town Hall”, or “Melbourne University”. Direct references to Melbourne mostly included specification of spatial relationships, e.g., “a suburb located in the east of Melbourne”, “1k south of Melbourne”. Often, categorical information is used to further specify locations “i am on the Docklands esplanade in the multi coloured 6 story building”. A general spatial distribution of all submitted place descriptions is shown in Figure 4. As already mentioned *Tell us where* was distributed from Melbourne, and thus we found 40% of descriptions located in the City of Melbourne and further 51% within a surrounding commuter belt of 100 km. The remaining 6% were spread throughout rural Victoria and a low 3% were located throughout the rest of Australia.



Figure 4: Spatial distribution of place descriptions (complete set of 2057 after filtering).

Investigating granularity and complexity of place description, we looked at two pieces of information: a granularity classification and text length. With respect to the finest granular place a person describes, we applied a hierarchical classification to group the descriptions into 6 different classes (in brackets are the observed percentages referring to our subset): country/state (7%), city (19%), street (15%), region (8%), building (33%) and room (2%). As we selected records from all zoom levels equally in this initial subset, the distribution will look different for the whole data set, since zoom level 15 in particular prevails in the data. However, so far we have been able to observe positive correlations between the classes and recorded zoom levels. For preliminary analysis using our subset we obtained a correlation coefficient $r=0.78$ ($R^2=0.9277$). This means that descriptions of coarse granularity such as “somewhere in westaustralia”, “going to perth along the coast”, or “in the simpson desert” also show low zoom levels, which is understandable considering the absence of landmarks in remote areas, and thus the necessity to choose a less detailed zoom level in order to get some location-related information.

Positive correlations were also noted between zoom level and text length (subset correlation coefficient $r=0.274$, $R^2=0.35$; total data: $r=0.13$, $R^2=0.59$). Figure 5 shows the correlation between the recorded zoom level of the mobile game and the text length of the whole set of given place description. We recorded in general shorter descriptions in lower zoom levels, whereas we found longer and thus presumably more detailed ones when people chose higher zoom levels e.g. to explain where to find a specific building or even where exactly in the building to find them, e.g., “In

my office on the 7th floor of the Babel building on the University of Melbourne Parkville campus. I am on the south side of the building...”

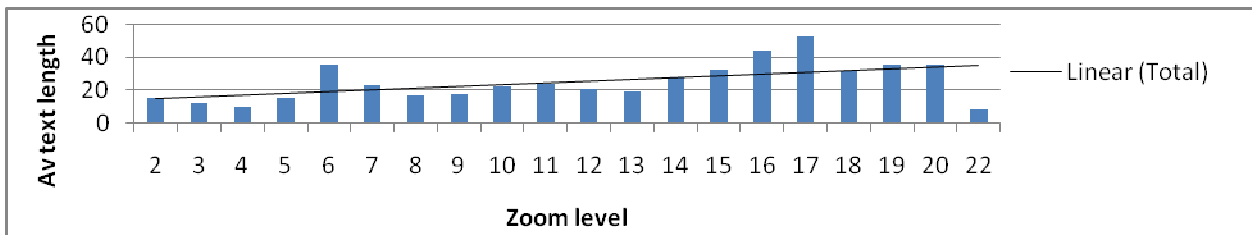


Figure 5: Average text length of description and zoom level for the complete dataset.

However, it is obvious that people were likely affected by the default zoom level 15 as 69% used this level. 6% chose level 16, 4 % level 13 and 3% each level 17, 14, 12 and 11. All other single zoom levels ranging from 2-22, above 17 and below 11 were chosen by 2 or less percent. This distribution can be explained in terms of the respective level of detail, since a zoom level of 5 will show the whole of Victoria, a level of 2 the continent of Australia whereas a high zoom level of, for example, 19 would only cover an area of approximately 60m x 80m on a smartphone.

In a last step we looked at context information related to the user’s activity such as “i am jogging at hillside park”, “fleetwood drive heading home”, or to surroundings/events such as in “burwood highway jam”, “at a loud street intersection”, or “...there is a busy road nearby”, context information in individual statements (“In a really nice restaurant on Victoria street”), as well as for indications whether the person was actual familiar with the place. People in unfamiliar situations often use semantic names such as “in the park”, “i’m sitting on a bench”, “near a building”, or “i am in north melbourne trying to find a car park”, instead of official place names or addresses. The link of place and the affordance of the environment for certain activities has to be further investigated (Scheider and Janowicz, 2010).



Figure 6: A player of *Tell us where* at leisure in a coffee house.

6 CONCLUSIONS AND FURTHER DIRECTIONS

This paper presented novel approaches to capture the names and the meanings of places. It suggested investigating corpora of human (common language) place descriptions, ideally geotagged place descriptions, to learn how people refer to places. One particular way to collect such a corpus is by mobile location-based games, and this paper has presented such a game, and analysed some of the preliminary observations made on the collected place descriptions.

In these preliminary observations it has become clear that games as a method to capture knowledge are shaping the context of people’s interactions. We made a number of corresponding observations in the *Tell us where* game. First, the given place descriptions can be categorized into different types (where, how to find, what activity), which reflects on the purpose of the place description, or the context of the player has imagined. Secondly, the given place descriptions can

be categorized by their finest level of granularity, which might be correlated with the zoom level in the self-localization. Granularity is also related to purpose or context imagined. Thirdly, the given place descriptions can be categorized by their length or complexity. The shorter place descriptions are either put at coarser levels of granularity, or they are more ambiguous.

Descriptions also often include further context information, e.g., respective activities or events. We observed that people are guided by their respective zoom level chosen for self-localization for choosing a context for their place description. Other factors were their activities and their familiarity with the place, as far as the latter could be inferred from their wording.

The project is now investigating the collected material and first experiences, but also addressing questions of making use of what has been learned. Our preliminary experiment, the mobile location-based game, showed some limitations that will shape our future research. The most significant one is the need to specify the communication context for the players more clearly, for example by assigning them a role and a communication purpose. In this respect, the default zoom level for self-localization already impacts on the kind of response we collected in the game. Similarly, if the purpose of the game is a collection of place names all over a specified area then the game must be playable everywhere in that area and promoted accordingly. Thirdly, cheating must be controlled and filtered out. For example, the discovery of a number of entries of identical coordinates suggests that they were given by the same persons just to raise their chances to win a voucher. Such opportunistic behaviour indicates that the players were not immersed, and hence their contributions were questionable (and not considered in the analysis).

Also the preliminary interpretations made on the place descriptions allow the formulation of future research questions. In particular we have started a more detailed linguistic and grammatical analysis of the place descriptions. This includes further semantic properties, such as the places' functions, but also the spatial relations used. Introducing and monitoring further parameters in future games to capture more about people's circumstances or their spatial context would be desirable.

ACKNOWLEDGEMENTS

This work was funded by the ARC under its Linkage Scheme (LP100200199), and by the Institute for a Broadband-Enabled Society (IBES). The game *Tell Us Where* was implemented and maintained by Hairuo Xie.

REFERENCES

- Bennett, R.; Winter, S. (2009): Future Research Directions on Computational Models of Geographic Places. Technical Report, Melbourne Engineering Research Institute (MERIT) - Spatial Information Science Research Group, The University of Melbourne, Australia.
- Burrough, P. A.; Frank, A. U. (Eds.) (1996): *Geographic Objects with Indeterminate Boundaries*. ESF-GISDATA, 2. Taylor & Francis.
- Cohn, A. G.; Renz, J. (2008): Qualitative Spatial Representation and Reasoning. In: van Hermelen, F.; Lifschitz, V.; Porter, B. (Eds.), *Handbook of Knowledge Representation*, Elsevier, pp. 551-596.
- Davies, C.; Holt, I.; Green, J.; Harding, J.; Diamond, L. (2009): User Needs and Implications for Modelling Vague Named Places. *Spatial Cognition and Computation*, 9 (3): 174-194.
- Denis, M.; Pazzaglia, F.; Cornoldi, C.; Bertolo, L. (1999): Spatial Discourse and Navigation: An Analysis of Route Directions in the City of Venice. *Applied Cognitive Psychology*, 13: 145-174.
- Dykes, J.; Purves, R.; Edwardes, A. J.; Wood, J. (2008): Exploring Volunteered Geographic Information to Describe Place: Visualization of the 'Geograph British Isles'. In: Lambrick, D.

(Ed.), 16th Annual Conference GIS Research UK: GISRUK 2008, Manchester, UK, pp. 256-267.

Elwood, S. (2008): Volunteered Geographic Information: Key Questions, Concepts and Methods to Guide Emerging Research and Practice. *GeoJournal*, 72 (3-4): 133-135.

Freundschuh, S. M.; Mark, D. M.; Gopal, S.; Gould, M. D.; Couclelis, H. (1990): Verbal Directions for Wayfinding: Implications for Navigation and Geographic Information and Analysis Systems. In: Brassel, K.; Kishimoto, H. (Eds.), 4th International Symposium on Spatial Data Handling. Department of Geography, University of Zurich, Zurich, pp. 478-487.

Goodchild, M. F. (2007): Citizens as Sensors: The World of Volunteered Geography. *GeoJournal*, 69: 211-221.

Haklay, M.; Weber, P. (2008): OpenStreetMap: User-Generated Street Maps. *Pervasive Computing*, 7 (4): 12-18.

Hill, L. L. (2006): *Georeferencing: The Geographic Associations of Information*. Digital Libraries and Electronic Publishing. The MIT Press, Cambridge, MA.

Krumm, J.; Davies, N.; Narayanaswami, C. (2008): User Generated Content. *Pervasive Computing*, 7 (4): 10-11.

Landau, B.; Jackendoff, R. (1993): "What" and "Where" in Spatial Language and Spatial Cognition. *Behavioral and Brain Sciences*, 16 (2): 217-238.

Levinson, S. C. (2003): *Space in Language and Cognition*. Cambridge University Press, Cambridge, 389 pp.

OGC (2008): OpenGIS Location Services (OpenLS): Part 6 -- Navigation Service. OGC Implementation Specification OGC 08-028, Open GIS Consortium Inc.

Paraboni, I.; Deemter, K. v.; Masthoff, J. (2007): Generating Referring Expressions: Making Referents Easy to Identify. *Computational Linguistics*, 33 (2): 229-254.

Richter, K.-F.; Winter, S. (2011): Citizens as Database: Conscious Ubiquity in Data Collection. In: Pfoser, D. et al. (Eds.), *Symposium on Spatial and Temporal Databases (SSTD)*. Lecture Notes in Computer Science, 6849. Springer, Berlin, pp. 445-448.

Scheider, S.; Janowicz, K. (2010): Places as Media of Containment, GIScience 2010 Extended Abstracts. University of Zurich, Zurich, Switzerland.

Shanon, B. (1979): Where Questions, 17th Annual Meeting of the Association for Computational Linguistics. ACL, University of California at San Diego, La Jolla, CA.

Surowiecki, J. (2004): *The Wisdom of Crowds*. Doubleday, New York, 296 pp.

Tuan, Y.-F. (1977): *Space and Place: The Perspective of Experience*. University of Minnesota Press, Minneapolis, MN, 235 pp.

van der Zee, E.; Slack, J. (Eds.) (2003): *Representing Direction in Language and Space*. Oxford University Press, Oxford, UK.

Vanclay, F.; Higgins, M.; Blackshaw, A. (Eds.) (2008): *Making Sense of Place*. National Museum of Australia Press, Canberra, 331 pp.

- Winter, S.; Richter, K.-F.; Baldwin, T.; Cavedon, L.; Stirling, L.; Duckham, M.; Kealy, A.; Rajabifard, A. (2011): Location-Based Mobile Games for Spatial Knowledge Acquisition. In: Janowicz, K. *et al.* (Eds.), *Cognitive Engineering for Mobile GIS*, Belfast, Maine, USA.
- Winter, S.; Truelove, M. (2010): Talking about Place Where It Matters. In: Frank, A. U.; Mark, D. M. (Eds.), *Las Navas 20th Anniversary Meeting on Cognitive and Linguistic Aspects of Geographic Space*, Las Navas, Spain.
- Winter, S.; Wu, Y. (2009): Intelligent Spatial Communication. In: Navratil, G. (Ed.), *Research Trends in Geographic Information Science*, Springer, Berlin, pp. 235-250.
- Wu, Y.; Winter, S. (2011): Interpreting Destination Descriptions in a Cognitive Way. In: Hois, J. *et al.* (Eds.), *Workshop on Computational Models for Spatial Language Interpretation and Generation (CoSLI-2)*. CEUR-WS.org, Boston, USA, pp. 8-15.

Harvesting User-Generated Content for Semantic Spatial Information: The Case of Landmarks in OpenStreetMap

Kai-Florian Richter, Stephan Winter
Department of Infrastructure Engineering
The University of Melbourne
Victoria 3010
Australia
kfr@kfrichter.org; winter@unimelb.edu.au

ABSTRACT

User-generated content currently revolutionizes data collection and maintenance. The advent of powerful mobile devices, Web 2.0 technology, and ubiquitous network connectivity enables and encourages users of a service to contribute their data to improve the utility of that service. This paper reports preliminary work on integrating landmarks in OpenStreetMap, illustrating the power of user-generated content for gaining semantic information. Landmarks are crucial in human wayfinding, but desperately lacking in navigation systems. Today's approaches to landmark integration require large amounts of geometric, attribute, and semantic data, rendering it impossible to include them in commercial systems due to the immense data collection costs. Distributing this task to users relieves these issues, but requires intelligent methods to avoid cumbersome interfaces that hinder collection of data or its future use.

KEYWORDS: user-generated content, spatial data infrastructures, navigation, location based services, OpenStreetMap

1 INTRODUCTION

A recent phenomenon points to dramatic changes in our information age. This phenomenon goes by several names, the most prominent ones being *crowd sourcing* (Surowiecki 2004), *citizens as sensors* (Goodchild 2007), *user-generated content* (Krumm et al. 2008; Haklay & Weber 2008), or *volunteered geographic information* (Elwood 2009). At the heart of this phenomenon is the idea of users participating in the collection and management of spatial data—the user as producer of information (Budhathoki et al. 2008). Another key characteristic of such an approach to Spatial Data Infrastructures (SDI) is that the contributing users are usually not trained in the matter at hand, but contribute data that they collected without a central authority managing or supervising the data collection process.

Individual approaches vary in their spectrum from conscious user actions ('volunteered') to passive modes ('citizens as sensors').

- *Volunteered* user-generated content is often used to replace existing commercial or authoritative datasets. A well known example is Wikipedia as an open encyclopaedia. OpenStreetMap as an open topographic dataset of the world is another example in the area of spatial infrastructures. User-generated content usually has rapid update cycles. This is exploited by some (commercial) providers to deliver improved services. For example, fixmystreet.com reports damages related to streets. Google, TomTom and other dataset providers encourage their users to report updates of their spatial data. In some cases, the database itself is the service. For example, Flickr allows users to upload and share photos.
- At the *passive* end of the spectrum, data mining methods can be used to further elicit hidden information out of the data. Researchers identified, for example, landmarks

defining a town from Flickr photo collections (Crandall et al. 2009; Schlieder & Matyas 2009), and commercial services track anonymized mobile phone locations to estimate traffic flow and enable real-time route planning.

In this paper, we explore how approaches to user-generated content can be exploited to collect data that is otherwise costly and hard to capture—namely semantic information. Semantic information attaches meaning to geographic features that is hard to infer from sensor readings. Often this information is personal, i.e., it refers to how (individual) humans conceptualize the environment they live in. Specifically, we address the challenge of integrating landmark information in navigation services. As will be discussed below, on the one hand landmarks are crucial in human navigation, on the other hand ‘landmarkness’, the aspects that let features stand out from other features nearby, is extremely difficult to assess automatically on a large scale. Thus, there are huge benefits in applying approaches of user-generated content in this domain.

The next section discusses the benefits and challenges of user-generated content. Section 3 then introduces the concept of landmarks, highlights their importance in human navigation, and describes the shortcomings of traditional approaches to integrating landmarks in navigation services. Sections 4 and 5 discuss the collection of semantic information through crowd-sourcing and exemplify by with collecting landmarks in OpenStreetMap. Section 6 concludes the paper.

2 BENEFITS AND CHALLENGES OF USER-GENERATED CONTENT

When applying user-generated content for managing SDIs, users expect free access, since the data is generated and shared by these users—they consider it to be ‘their’ data. Reproduction and distribution costs for this data are close to zero (Shapiro & Varian 1998). The traditional economy of spatial information is fundamentally challenged. In short, as has been introduced above, the major benefits of user-generated content are drastically reduced costs of data collection and time to next update.

In studying the motivation of people to contribute, Benkler demonstrates “that the diverse and complex patterns of behavior observed [...] are perfectly consistent with much of our contemporary understanding of human economic behavior” (Benkler 2006, p. 91). The phenomenon of user-generated content is not going away soon.

Furthermore, it has been proven that user-generated data can match the quality of authoritative datasets (Haklay 2010; Zielstra & Zipf 2010). However, as there is a lack of control in the data acquisition process, approaches to user-generated content are frequently criticized for heterogeneous quality, lack of validation, and potential impact on privacy, among other issues (Flanagin & Metzger 2008; Keßler et al. 2009). Currently, these concerns impact on the trust and acceptance of user-generated content in SDI, despite its economic and scaling advantages.

User-generated spatial data consists of a combination of sensor data (e.g. coordinates) and user-added semantics (e.g. place descriptions). Sensor data is easy to collect automatically (e.g. by GPS sensors). Sensor data is also easy to process automatically, for example, to create geometric representations of geographic features from a user’s GPS trajectories. Management, processing, and use of user-added semantics is still in its infancy, however.

Traditionally, semantics are inferred by data mining methods on sensor data, and the same has been done for user-added semantics (Purves et al. 2010). We believe that ultimately this traditional approach will fall short in capturing semantics in the required breadth and depth, since, fundamentally, adding semantics is a human intelligence task. Human Intelligence Tasks (HIT) seem to be simple tasks that appear to be ideally suited to automation (e.g. identifying objects on a photo), but often are in fact easier and cheaper to solve through crowd-sourcing. The term HIT was introduced by Amazon’s Mechanical Turk (<https://www.mturk.com>).

Collecting landmark information, the kind of semantic information discussed in this paper, goes well beyond the simple tasks of Mechanical Turk, so this information will need to be contributed by humans. The challenge is collecting these user-added semantics and developing novel information inference processes for semantically rich data.

3 LANDMARK INFORMATION

Landmarks are crucial in learning environments (Siegel & White 1975) and in forming mental representations of an environment (Hirtle & Jonides 1985; Couclelis et al. 1987). In communicating about an environment, for example, when giving route directions, people use landmarks to anchor actions in space or to provide confirmation that the right track is still being followed (Denis 1997; Lovelace et al. 1999; Michon & Denis 2001). Landmarks are the foundation for our 'image of the city' (Lynch 1960).

Not surprisingly, landmarks are a desirable addition to automatic navigation services, such as car navigation systems. They are a top feature request by users (May et al. 2003). Using prototypical research systems, landmarks have been found to improve users' performance and satisfaction with such systems (Burnett 2000; Ross et al. 2004). However, most commercial systems and services ignore them—with a notable exception of the Australian routing service whereis (Duckham et al. 2010).

3.1 What is a Landmark?

Lynch (1960) defined a landmark to be a readily identifiable object which serves as external reference point. This definition is frequently picked up in the literature. In fact, anything that sticks out from the background may serve as a landmark (Presson & Montello 1988). This includes individual buildings (e.g. a church), signage along a street, or linear features, such as rivers (Richter 2007), and even salient street intersections, such as roundabouts (Klippel et al. 2005). The Urban Knowledge Data Structure (Hansen et al. 2006; Klippel et al. 2009) provides an elaborate specification of which types of features may serve as landmarks in navigation instructions.

The "sticking out from the background" as a landmark is defined through a feature's salience, i.e., its suitability as a landmark (Raubal & Winter 2002; Elias 2003). Sorrows and Hirtle (1999) identified three key characteristics of landmarks that influence this salience: 1) singularity, i.e., contrast with its surroundings; 2) prominence of spatial location; 3) content, i.e., meaning or cultural significance. Several approaches to dealing with landmarks aim at covering these characteristics in calculating salience values for landmark candidates.

The inclusion of references to landmarks in navigation services requires two steps: 1) the identification of features that may serve as landmarks in principle (in the following, these features will be referred to as *landmark candidates*); 2) the selection of some of these candidates for inclusion in the navigation instructions. Often, these two steps are seen as independent (Elias et al. 2005). Landmark identification is the focus of this paper; therefore, the following discussion concentrates on this issue.

3.2 Landmark Identification

The first approach to the automatic identification of landmarks was presented by Raubal and Winter (2002). It inspired several extensions and further approaches. Their approach reflects the three landmark characteristics of Sorrows and Hirtle (1999) by taking into account different parameters of building facades (e.g. area, color, signs, visibility) in a weighted sum for calculating individual buildings' salience. These facades serve as point-like landmarks along a route; the required data is supposed to be stored in a spatial database. Employing a user survey, the weights for the individual parameters were set for specific situations, specifically accommodating for differences between day and night (Winter et al. 2005). New situations require an adaptation of these weights, which requires new user studies. Further, the location of a landmark has an influence on the ease of conceptualizing turning actions and, thus, determines the ease of understanding route instructions (Richter & Klippel 2007).

A similar approach to Raubal and Winter (2002) is taken by Elias (2003). She uses machine learning techniques to identify the most salient objects in a spatial data set. These objects are considered to be point-like entities. Winter et al. (2008) combine these two approaches of Raubal and Winter, and of Elias to construct a hierarchy of landmarks based on each individual candidate's salience. This hierarchy is used in the generation of destination descriptions (Tomko & Winter 2006), which presents an integration of landmarks into wayfinding instructions.

Others have explored data mining approaches to identify landmark candidates. Tomko (2004) used requests to Internet search engines looking for street names and subsequent filtering mechanisms to identify potential salient features (buildings) along a previously calculated route. In Tezuka and Tanaka (2005), text mining methods are used to mine WWW documents in order to identify prominent point-like features; prominence is based on how authors of these documents refer to the features. Others have mined Flickr photo collections to identify those landmarks that define a town, i.e., that are represent a typical view of that town (Crandall et al. 2009; Schlieder & Matyas 2009).

3.3 Landmark Integration

Caduff and Timpf (2005) presented an algorithm that calculates a route through a network based on the presence of point-like landmarks at decision points (nodes). It tries to navigate a wayfinder along a route that has a landmark at every decision point. The approach by Richter (Richter & Klippel 2007; Richter 2007) integrates landmarks into an abstract specification of route directions. His approach selects those landmarks from a set of landmark candidates that are best suited to describe actions to be performed (Klippel & Winter 2005).

The CORAL system (Dale et al. 2005) produces natural language instructions for route following, mimicking human principles of direction giving, which includes references to landmarks. Recently, Duckham et al. (Duckham et al. 2010) explored the use of categories of features instead of their individual properties to determine suitability as a landmark. They combined a category's general suitability, its uniqueness in an area and a feature's location along a route to select those features best suited to describe following the given route. This approach is implemented in the whereis route service (<http://www.whereis.com>), using yellow pages entries as landmark candidates.

However, several challenges prevent the integration of landmark references into commercial systems. In these systems, calculation and generation of route directions is based on simple, efficient algorithms. Metric distances and references to street names are easily calculable from a geo-referenced network representation of the street layout. Landmarks need to be embedded into this existing network structure in a seamless way, which requires annotating the network with additional features such that they are easily integratable into the directions.

Some systems combine metric distances with references to traffic lights to provide additional context¹. In some systems, points of interest (POIs), such as hotels or gas stations, are accessible, but they are hardly ever used for describing the route to take, but rather as selectable destinations or for advertising purposes. The seamless integration of landmarks requires a suitable data structure; the Urban Knowledge Data Structure (Hansen et al. 2006; Klippel et al. 2009) that is based on OGC's OpenLS specification (<http://www.opengeospatial.org/standards/ols>) might be such an approach. This structure can deal with different types of landmarks and offers mechanisms for structuring route information.

Given such a data structure, there is still the need to identify landmark candidates and then to integrate suitable candidates into route directions. However, as discussed previously, the identification of landmarks and the selection for route directions lacks integration. Elaborate approaches to landmark integration ignore the problem of identifying landmark candidates (Caduff & Timpf 2005; Richter 2007).

3.4 Shortcomings of Automatic Landmark Identification Mechanisms

Landmark identification mechanisms are highly data intensive. They identify individual features that may serve as a landmark. To gain useful results, these individuals need to be described in great detail and in multiple dimensions, as discussed above. Also, since 'landmarkness' is a relative characteristic—standing out from other features nearby—the other features nearby need to be part of the identification process. (Nothegger et al. 2004) have demonstrated and discussed the effort of identifying landmarks especially for building facades. The

¹ Often, however, this is done without taking into account the presence of other traffic lights. It is not uncommon to get instructions, such as 'in 500m, at the traffic lights, turn left,' with no consideration of other traffic lights before the one referred to.

required information is hard to collect automatically and, therefore, labor-intensive, needs to be specifically collected for each town and will result in large amounts of data. This makes it unlikely that it will appear in commercial databases due to the attached immense collection efforts and costs.

Using categories rather than individuals may be the more promising approach, as can be seen with the implementation of Duckham et al.'s (2010) approach in the whereis web service. Using categories, properties of individual features do not need to be known since they are inferred by some heuristics from a general assessment of a specific category's suitability as landmark. This approach requires much less data. The relevant information comprises location, geometry, and type of feature. Landmark candidates are, however, not evenly distributed across the environment. Duckham et al. state that in the current whereis implementation for all of Australia only 170 000 features from 66 categories were used as a landmark, resulting in a sparse distribution of landmark candidates throughout the country. And landmark density will be significantly higher in inner-city areas compared to suburbs or rural areas.

In summary, while generating landmark references based on category information rather than properties of individuals seems a promising option for automatic landmark identification, the challenge remains to pool sufficient information about enough features from an adequate number of categories so that landmark candidates cover all parts of an environment. This is where user-generated content shows much potential.

4 USER-GENERATED SEMANTIC CONTENT

User-generated content is ideally suited to collecting semantic information, rather than just data (Richter & Winter 2011). In the spatial domain, people are local experts of their neighbourhoods. They have detailed, up-to-date knowledge of its spatial layout and social structure. Using methods for user-generated content, this knowledge can be gathered. Locals may not only provide their tracks or locations, but also their local expertise for free. For example, users map their neighbourhoods in OpenStreetMap (see below), or annotate places in CityFlocks (Bilandzic et al. 2008), which can then be accessed by others for local decision making. Since contributors are in the environment when they trace geographic features' geometries and describe their semantics, they are ideal candidates to provide their insights and experience for others to use. Physical presence increases their credibility when reporting on the experience of spatial and social structures of environments. That is, compared to traditional methods of spatial data collection, user-generated content has another advantage: gathering information from users taps into their individual knowledge and experiences, thus, hugely broadening the kind of information that can be collected.

This collected information is partly available already. But it is fragmented, distributed, and often only implicitly inferable. It is scattered on various web sources, such as restaurant reviews, Flickr image collections, local news, or social networks. In order to be of real use for next-generation location-based services, dedicated approaches to collecting this information are needed. In addition, holistic, comprehensive ways of storing and managing this data are required. This information needs to be represented in a common data structure.

Such a common data structure would open up opportunities for location based services and human-machine communication that are not available today. Such future services would access the collective wisdom of multiple local experts at any location and, in this manner, replace the need to ask the 'local expert.' For example, such services would be able to deliver topographic data of the neighbourhood together with place annotations as in CityFlocks (Bilandzic et al. 2008), to make accessible gazetteers of official and vernacular place names, or to provide access to local and global landmarks for navigation. More generally, these services would communicate like humans (Winter & Wu 2009), which would be vastly beneficial since humans are not necessarily experts when it comes to communicating with machines.

In the following, we report on the first steps towards methods for collecting semantic information about landmarks for urban environments. This approach is implemented in the context of OpenStreetMap.

5 USER-CONTRIBUTED LANDMARKS: THE CASE OF OPENSTREETMAP

We implemented a prototype editor that allows for annotating OpenStreetMap data with landmark information. This section provides details of the chosen approach and its realization in a Python application (Ghasemi 2011; Ghasemi et al. 2011). First, however, the OpenStreetMap project itself is introduced.

5.1 OpenStreetMap

OpenStreetMap (<http://www.openstreetmap.org/>) is a project that allows people to contribute geographic data of their neighborhood to a growing collection of topographic data. OpenStreetMap provides user-contributed data sets and maps visualizing this data (Haklay & Weber 2008). The project was initiated in 2004 at University College London (UCL) by Steve Coast. Since then, it has grown rapidly in the number of users and the geographic data it provides.

OpenStreetMap's geographic data is mainly collected by volunteers using GPS, free satellite images and free data sources. Moreover, some organizations such as local and national governments, have donated geographic data to OpenStreetMap that was used to complete data in some areas of the world. Each contributor needs to have a username and password to be able to edit data. This ensures that the source of data can be traced, for example, in case of a copyright dispute. Users are able to add any type of information associated with the geographic data. Any contributor can edit the existing geographic data and its attached information if these are wrong, out of date or not accurate enough. Geographic data contains a variety of features including roads, buildings, points of interest, rivers, and administrative boundaries.

5.2 Characterizing Landmarks in OpenStreetMap

In the chosen approach, characterization of landmarks follows the classification introduced by Sorrows and Hirtle (1999). Each landmark object has a name and a type associated with it. A combination of these two attributes will be used to generate a description of the landmark object that is unambiguous in its surroundings. The attribute 'name' is general, ranging from proper place names, such as 'Flinders Street Station' to descriptive names, such as 'the large, old, yellow train station'.

Furthermore, according to Sorrows and Hirtle, the 'landmarkness' of an object is represented using three different attribute classes: visual characteristics, semantic characteristics, and structural characteristics. Visual characteristics include properties, such as shape and size, or texture and visibility. Semantic landmarks are characterized through their type, and cultural and historical significance. Finally, structural significance emerges from an object's centrality and location in an environment (e.g. Flinders Street Station in Melbourne's CBD). Figure 1 summarizes these landmark properties.

Landmark properties are stored as 'tags' in OpenStreetMap. A 'tag' is a key or value pair that associates a specific value to a given key. The OpenStreetMap data structure allows contributors to freely add new tags to any given feature in the database, which makes implementing landmarks in OpenStreetMap easy to achieve. While it would be possible to allow for dedicated specific values for each of the parameters, in the prototype each identifying factor is realized as a boolean value that indicates whether that factor defines or does not define the specified landmark.

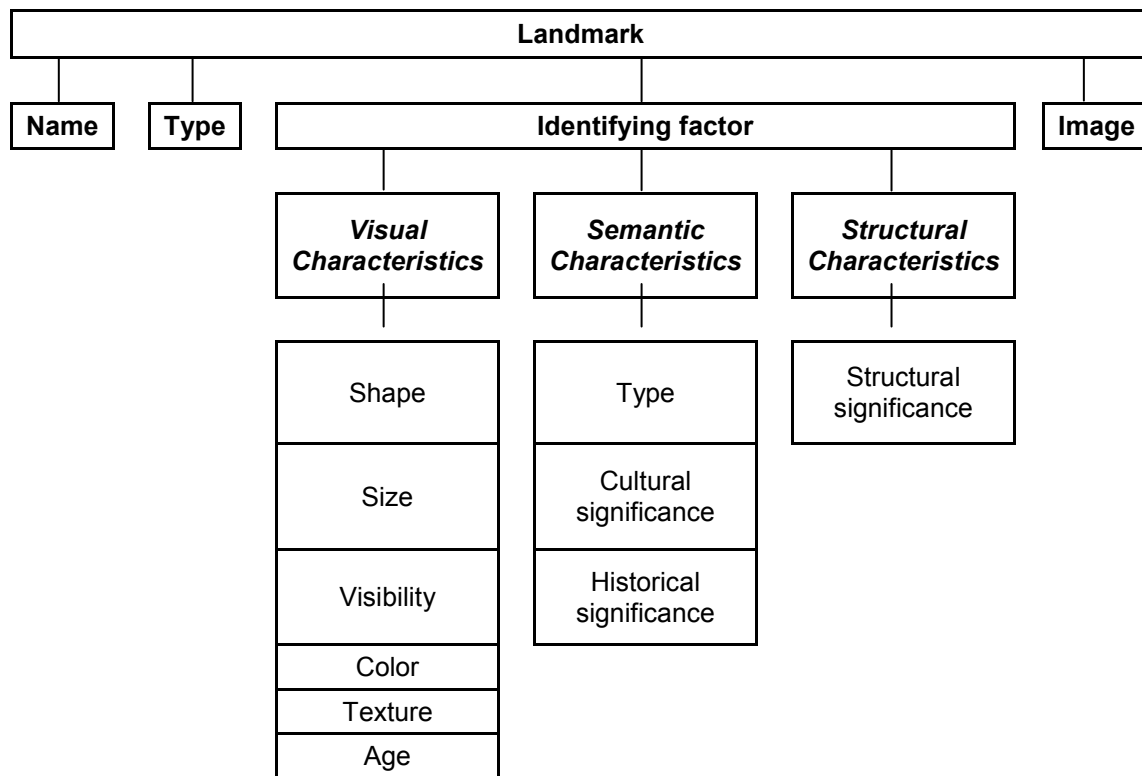


Figure 1: Properties defining the 'landmarkness' of a geographic feature.

5.3 The Editor: Marking OpenStreetMap Objects as Landmarks

Figure 2 illustrates the interface of the prototype editor developed for adding landmarks to OpenStreetMap. The editor is implemented in Python, using the pyGTK and GTK+ libraries for the graphical user interface, the tl.geodrawing library for displaying geographic data, and the PythonOSM API for connecting to OpenStreetMap.

Based on a user's current position, the application downloads OpenStreetMap data of the surrounding environment and renders it in the main part of the interface. 'Select Landmark' and 'Add Landmark' buttons both allow adding landmark information to OpenStreetMap—explained in more detail below. The 'Exit' button (unsurprisingly) closes the application.

There are two ways of adding landmark information to OpenStreetMap, both relating to two different scenarios: Firstly, users may want to identify an existing feature as a landmark. Secondly, a feature that users have identified in their environment as a landmark may not yet be present in OpenStreetMap. 'Select Landmark' caters for the former, 'Add Landmark' for the latter scenario.

In the first scenario, when identifying an existing feature, that feature is selected by clicking its defining node, i.e., the node of the feature that tags are assigned to. This will result in a new window popping up that asks the user to specify which of the identifying factors apply for the chosen feature (see Figure 3). Pressing 'Save' will update the feature's description in the OpenStreetMap database, i.e., tags that correspond to the chosen factors are added to the feature description.

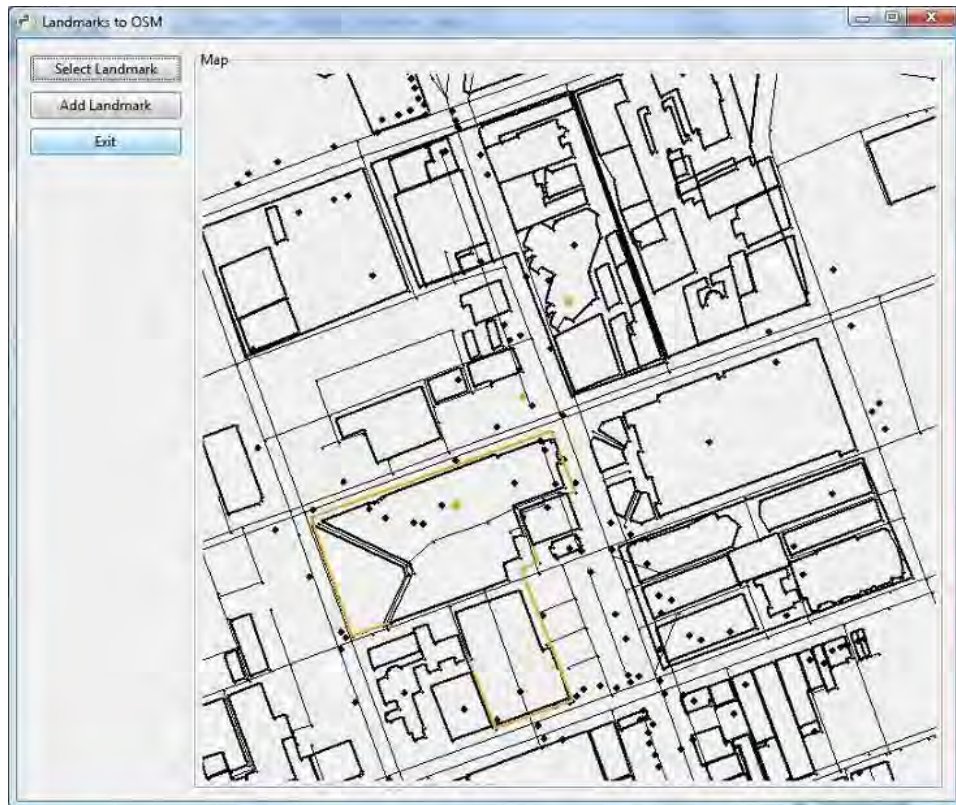


Figure 2: The application's main window.

Adding new landmark features works in a similar way. However, as a first step, users have to pinpoint the location of the new feature on the map, i.e., mark the position of the feature to be added by clicking on the map. This will define the geographic location of the new feature. Currently, only point-like features can be added this way. Users are also required to define the type of the new feature by selecting a value from a predefined list of possible types.

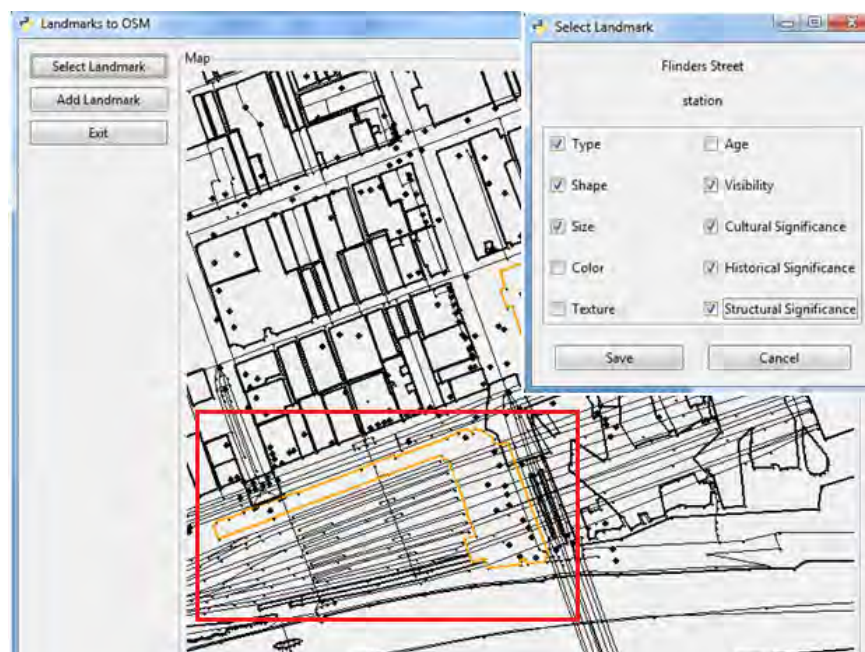


Figure 3: Selecting a feature (here, Flinders Street Station in Melbourne's CBD) in order to mark it as a landmark. The small in-lay window allows for specifying the identifying factors.

5.4 Discussion

Several issues remain for the presented crowd-sourcing approach to collecting landmarks: These are on different levels: 1) the editor itself; 2) the OpenStreetMap data; 3) the general approach.

Regarding the editor, the presented implementation is clearly only a prototype illustrating the general principles. In order to become of widespread use, it needs to be ported to run on current smart phones and other mobile devices utilising GPS positioning. Such a step would likely also require adaptations to the user interface in order to cope with touchscreen interactions and small screens. In addition, while boolean values for specifying the identifying factors are easy to implement and also allow for some reasoning as of why a given feature functions as a landmark, such values are poor when it comes to (automatically) generating descriptions of landmarks, for example, in verbal navigation instructions. Thus, some way of allowing users to enter more descriptive values for the factors is a desirable future development. Finally, a valuable additional feature would be to enable users to upload a photograph (or even several) of the identified landmark feature. Human navigators predominantly identify landmarks through visual perception; being able to display a photo representation of a landmark will increase their usability. Based on current smart phone technology, this feature is fairly easy to add.

As explained previously, the OpenStreetMap data specification is ideally suited for implementing functionality that adds new attributes to existing data. Since any user can add any tags as they see the need for them, the data structure is flexible with respect to how data is specified—all that is required is that the underlying XML structure is syntactically well-formed. This flexibility is also a drawback, however. Since there is no central authority validating the semantics of the underlying data, the OpenStreetMap approach relies on users rectifying errors of other users. One example which illustrates this is the way polygonal features, which are typically chosen as landmarks, are defined. The OpenStreetMap guidelines suggest modeling feature geometry by a sequence of points with each point marking a corner of the feature. Then, in order to define the features' attributes, an 'artificial' point is to be introduced (at the center of gravity), which gets all the tags attached. This avoids the need to attach all tags to every point of the geometry—the editor relies on such a point being present for adding landmark semantics. While this is a clean way of modeling features, in practice there are several features that violate this guideline. These features wait for other users to correct the modeling, a process that seems to be considered less rewarding than that of adding new data for previously unmapped areas, and hence is not necessarily resolved quickly.

Finally, the general approach of utilising users of a service to add landmarks to a topographic database is worth a closer look. Several commercial providers are currently moving in this direction (e.g. TomTom, Google). As discussed above, a major advantage of such an approach is that local experts add their knowledge to the service, which often is information that otherwise would be hard to acquire. Hence, a challenge of such an approach is finding experts that are willing to contribute. This will be easier in densely populated areas, simply because there are much more potential volunteers. Overall, contributing to a service must be seen as rewarding by the users to ensure continued contribution (Budhathoki et al. 2010). Part of this experience requires intelligent, easily usable interfaces, which at the same time are not intrusive in other activities and keep users in control of the data they contribute—'conscious ubiquity' in user interaction (Richter & Winter 2011).

6 SUMMARY

This paper discussed the power of user-generated content as one method for collecting and managing geographic data in Spatial Data Infrastructures. In particular, it described the possibilities such an approach opens up in terms of accessing semantic information of geographic phenomena, instead of just geometry or sensor data. Fully exploring the potential of exploiting human intelligence in collection and maintenance will lead to greatly improved location-based services. The paper presented first steps in this direction by discussing a prototype editor for adding landmarks to the OpenStreetMap topographic database.

ACKNOWLEDGMENTS

The OpenStreetMap interface was developed and implemented by Mahsa Ghasemi as part of her master research project. Her work was also supported by a MERIT Summer Research Experience grant.

REFERENCES

- Benkler, Y.: *The Wealth of Networks*. Yale University Press, New Haven (2006)
- Bilandzic, M., Foth, M., De Luca, A.: CityFlocks: designing social navigation for urban mobile information systems. In: *DIS '08: Proceedings of the 7th ACM Conference on Designing Interactive Systems*, New York, NY, USA 2008, pp. 174-183. ACM
- Budhathoki, N., Bruce, B., Nedovic-Budic, Z.: Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal* **72**(3), 149-160 (2008)
- Budhathoki, N.R., Nedovic-Budic, Z., Bruce, B.: An Interdisciplinary Frame for Understanding Volunteered Geographic Information. *Geomatica* **64**(1), 11-26 (2010)
- Burnett, G.: 'Turn Right at the Traffic Lights': The Requirement for Landmarks in Vehicle Navigation Systems. *Journal of Navigation* **53**(3), 499-510 (2000)
- Caduff, D., Timpf, S.: The Landmark Spider: Representing Landmark Knowledge for Wayfinding Tasks. In: Barkowsky, T., Freksa, C., Hegarty, M., Lowe, R. (eds.) *Reasoning with mental and external diagrams: computational modeling and spatial assistance - Papers from the 2005 AAAI Spring Symposium*, Menlo Park, CA 2005, pp. 30-35
- Couclelis, H., Golledge, R.G., Gale, N., Tobler, W.: Exploring the Anchor-Point Hypothesis of Spatial Cognition. *Journal of Environmental Psychology* **7**, 99-122 (1987)
- Crandall, D.J., Backstrom, L., Huttenlocher, D., Kleinberg, J.: Mapping the world's photos. In: *Proceedings of the 18th International Conference on World Wide Web*, Madrid, Spain 2009, pp. 761-770. ACM, 1526812
- Dale, R., Geldof, S., Prost, J.-P.: Using Natural Language Generation in Automatic Route Description. *Journal of Research and Practice in Information Technology* **37**(1), 89-105 (2005)
- Denis, M.: The Description of Routes: A Cognitive Approach to the Production of Spatial Discourse. *Cahiers Psychologie Cognitive* **16**(4), 409-458 (1997)
- Duckham, M., Winter, S., Robinson, M.: Including landmarks in routing instructions. *Journal of Location Based Services* **4**(1), 28 - 52 (2010)
- Elias, B.: Extracting Landmarks with Data Mining Methods. In: Kuhn, W., Worboys, M., Timpf, S. (eds.) *Spatial Information Theory*, Berlin 2003. *Lecture Notes in Computer Science*, pp. 375-389. Springer
- Elias, B., Paelke, V., Kuhnt, S.: Concepts for the Cartographic Visualization of Landmarks. In: Gartner, G. (ed.) *Location Based Services & Telecartography 2005*. *Geowissenschaftliche Mitteilungen*, pp. 1149-1155. TU Vienna
- Elwood, S.: Geographic Information Science: New Geovisualization Technologies --- Emerging Questions and Linkages with GIScience Research. *Progress in Human Geography* **33**(2), 256-263 (2009)
- Flanagan, A., Metzger, M.: The credibility of volunteered geographic information. *GeoJournal* **72**(3), 137-148 (2008)
- Ghasemi, M.: *Developing a user generated method to add landmarks to OpenStreetMap*. Master Thesis, University of Melbourne (2011)
- Ghasemi, M., Richter, K.-F., Winter, S.: Landmarks in OSM. Paper presented at the State of the Map, Denver, Colorado,
- Goodchild, M.: Citizens as sensors: the world of volunteered geography. *GeoJournal* **69**(4), 211-221 (2007)
- Haklay, M.: How Good is Volunteered Geographic Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design* **37**(4), 682-703 (2010)
- Haklay, M., Weber, P.: OpenStreetMap: User-Generated Street Maps. *Pervasive Computing* **7**(4), 12-18 (2008)

- Hansen, S., Richter, K.-F., Klippel, A.: Landmarks in OpenLS - A Data Structure for Cognitive Ergonomic Route Directions. In: Raubal, M., Miller, H., Frank, A.U., Goodchild, M.F. (eds.) *Geographic Information Science*, Berlin 2006. *Lecture Notes in Computer Science*, pp. 128-144. Springer
- Hirtle, S.C., Jonides, J.: Evidence of Hierarchies in Cognitive Maps. *Memory & Cognition* **13**(3), 208-217 (1985)
- Keßler, C., Janowicz, K., Bishr, M.: An agenda for the next generation gazetteer: geographic information contribution and retrieval. Paper presented at the GIS '09: Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, New York, NY, USA,
- Klippel, A., Hansen, S., Richter, K.-F., Winter, S.: Urban Granularities - A Data Structure for Cognitively Ergonomic Route Directions. *GeoInformatica* **13**(2), 223-247 (2009)
- Klippel, A., Richter, K.-F., Hansen, S.: Structural Saliency as a Landmark. In: *Workshop Mobile Maps 2005*, Salzburg, Austria 2005
- Klippel, A., Winter, S.: Structural Saliency of Landmarks for Route Directions. In: Cohn, A.G., Mark, D. (eds.) *Spatial Information Theory*, Berlin 2005. *Lecture Notes in Computer Science*, pp. 347-362. Springer
- Krumm, J., Davies, N., Narayanaswami, C.: User-Generated Content. *Pervasive Computing* **7**(4), 10-11 (2008)
- Lovelace, K.L., Hegarty, M., Montello, D.R.: Elements of Good Route Directions in Familiar and Unfamiliar Environments. In: Freksa, C., Mark, D.M. (eds.) *Spatial Information Theory*, Berlin 1999. *Lecture Notes in Computer Science*, pp. 65-82. Springer
- Lynch, K.: *The Image of the City*. The MIT Press, Cambridge (1960)
- May, A.J., Ross, T., Bayer, S.H., Tarkiainen, M.J.: Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing* **7**(6), 331-338 (2003)
- Michon, P.-E., Denis, M.: When and Why are Visual Landmarks Used in Giving Directions? In: Montello, D.R. (ed.) *Spatial Information Theory*, Berlin 2001. *Lecture Notes in Computer Science*, pp. 400-414. Springer
- Nothegger, C., Winter, S., Raubal, M.: Computation of the Saliency of Features. *Spatial Cognition and Computation* **4**(2), 113-136 (2004)
- Presson, C.C., Montello, D.R.: Points of Reference in Spatial Cognition: Stalking the Elusive Landmark. *British Journal of Developmental Psychology* **6**, 378-381 (1988)
- Purves, R., Jones, C., Clough, P.: *6th Workshop on Geographic Information Retrieval*. Zurich, Switzerland (2010)
- Raubal, M., Winter, S.: Enriching Wayfinding Instructions with Local Landmarks. In: Egenhofer, M., Mark, D. (eds.) *Geographic Information Science*, Berlin 2002. *Lecture Notes in Computer Science*, pp. 243-259. Springer
- Richter, K.-F.: A Uniform Handling of Different Landmark Types in Route Directions. In: Winter, S., Duckham, M., Kulik, L., Kuipers, B. (eds.) *Spatial Information Theory*, Berlin 2007. *Lecture Notes in Computer Science*, pp. 373-389. Springer
- Richter, K.-F., Klippel, A.: Before or After: Prepositions in Spatially Constrained Systems. In: Barkowsky, T., Knauff, M., Ligozat, G., Montello, D.R. (eds.) *Spatial Cognition V*, Berlin 2007. *Lecture Notes in Artificial Intelligence*, pp. 453-469. Springer
- Richter, K.-F., Winter, S.: Citizens as Database: Conscious Ubiquity in Data Collection. In: Pfoser, D., Tao, Y., Mouratidis, K., Nascimento, M.A., Mokbel, M., Shekhar, S., Huang, Y. (eds.) *Advances in Spatial and Temporal Databases*, Berlin 2011. *Lecture Notes in Computer Science*. Springer
- Ross, T., May, A., Thompson, S.: The Use of Landmarks in Pedestrian Navigation Instructions and the Effects of Context. In: Brewster, S., Dunlop, M. (eds.) *Mobile Computer Interaction- MobileHCI 2004*, Berlin 2004. *Lecture Notes in Computer Science*, pp. 300-304. Springer
- Schlieder, C., Matyas, C.: Photographing a City: An Analysis of Place Concepts Based on Spatial Choices. *Spatial Cognition & Computation: An Interdisciplinary Journal* **9**(3), 212-228 (2009)
- Shapiro, C., Varian, H.R.: *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press, (1998)

- Siegel, A.W., White, S.H.: The Development of Spatial Representations of Large-Scale Environments. In: Reese, H.W. (ed.) *Advances in Child Development and Behaviour*. pp. 9-55. Academic Press, New York (1975)
- Sorrows, M.E., Hirtle, S.C.: The Nature of Landmarks for Real and Electronic Spaces. In: Freksa, C., Mark, D.M. (eds.) *Spatial Information Theory*, Berlin, August 1999. *Lecture Notes in Computer Science*, pp. 37-50. Springer
- Surowiecki, J.: *The Wisdom of Crowds*. Doubleday, New York (2004)
- Tezuka, T., Tanaka, K.: Landmark Extraction: A Web Mining Approach. In: Cohn, A.G., Mark, D.M. (eds.) *Spatial Information Theory*, Berlin 2005. *Lecture Notes in Computer Science*, pp. 379-396. Springer
- Tomko, M.: Case Study -- Assessing Spatial Distribution of Web Resources for Navigation Services. In: 4th International Workshop on Web and Wireless Geographical Information Systems W2GIS04, Koyang, Korea 2004
- Tomko, M., Winter, S.: Recursive Construction of Granular Route Directions. *Journal of Spatial Science* **51**(1), 101-115 (2006)
- Winter, S., Raubal, M., Nothegger, C.: Focalizing Measures of Salience for Wayfinding. In: Meng, L., Zipf, A., Reichenbacher, T. (eds.) *Map-based Mobile Services - Theories, Methods and Implementations*. pp. 127-142. Springer, Berlin (2005)
- Winter, S., Tomko, M., Elias, B., Sester, M.: Landmark hierarchies in context. *Environment and Planning B: Planning and Design* **35**(3), 381-398 (2008)
- Winter, S., Wu, Y.: Intelligent spatial communication. In: Navratil, G. (ed.) *Research Trends in Geographic Information Science*. pp. 235-250. Springer, Berlin (2009)
- Zielstra, D., Zipf, A.: A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany. In: Painho, M., Santos, M.Y., Pundt, H. (eds.) *Geospatial Thinking: Proceedings of 13th AGILE International Conference on Geographic Information Science* 2010

Spatially Enabling Risk for Management of Land and Property

Katie Potts, Abbas Rajabifard, Ian Williamson

The University of Melbourne
Parkville, Melbourne, Australia

kpotts@pgrad.unimelb.edu.au, abbas.r@unimelb.edu.au, ianpw@unimelb.edu.au

Rohan Bennett

University of Twente
Enschede, Netherlands
bennett@itc.nl

ABSTRACT

Recent natural disasters such as earthquakes, flood, and fire, and other issues such as asbestos, pests, and fraud all highlight the spatial nature of risk and the relationship between place and risk. These examples show that location is a major factor in determining what or where is at risk and how to manage that risk. In this paper the term risk refers to hazards and threats that exist that have the potential to damage or destroy land and property. Using this definition some examples of risks include bushfire, earthquake, sea level rise, flood, fraud, pests, asbestos, drought, tsunami, and cyclone. These risks affect different scales: some hit locally, some hit entire states, and others cross state borders to affect people on a national scale. In order to reduce the uncertainty that risks bring they need to be managed. In order to manage these risks however, accurate and timely land, property, or spatial information is required. Land administration systems currently hold this valuable land and property information but are not arranged in a way to allow for coordination and sharing across jurisdictions. This paper is based on ongoing research and uses empirical research and a case study approach to discuss the spatial nature of risk, the relationship between place and risk, and proposes a spatially enabled approach to managing risks for citizens, government, and wider society.

KEYWORDS: risk, land administration, risk management, spatial enablement

1 INTRODUCTION

As the most recent natural disasters such as fire, flood, and earthquakes demonstrate – risk is all around us, and individuals, governments, and organizations alike can be affected. What is further demonstrated by these recent events is that many of these risks are spatial and could be better managed with more detailed spatial information. Currently however, not enough use is being made of spatial technologies. Moreover, many risks relate to property, so is the valuable information held within the land administration systems being utilised?

In order to manage these risks accurate and timely land, property, or spatial information is required. This is difficult however, because many systems for managing risk have emerged in an ad hoc fashion. Moreover, approaches to spatially describing risk are often undeveloped and simplistic. The information is often fragmented, based on point datasets and is limited in its utility. In order to effectively manage these risks this information needs to be combined with and make use of land, property and spatial information. That is, the management of risk needs to be spatially enabled.

This paper intends to explore the idea of risks as spatial objects and endeavours to develop a framework to explain this way of thinking. The framework will explore the three main obstacles

that need to be addressed: the coordination of this information, the aggregation of this information, and the dissemination of this information in a way that encourages consistency and efficiency. The framework will be directed towards the three main stakeholders identified which are governments – at all levels, citizens and business.

The first section discusses the risk theory and how the theory can be applied to land and property. Risk management is discussed and the stakeholders in the risk management process are identified. The background of land administration systems in the context of risk management is then discussed. The method is then outlined, followed by the results of the research and discussion.

2 METHOD AND APPROACH

Preliminary studies into the arrangements of current information infrastructures, particularly land administration systems, and risk management stakeholders have been undertaken. The research began initially with a literature review of risks and risk management, and land administration systems. A case study of the Australian context was then carried out. The case study was qualitative in nature. From the information gathered in the case study a framework was developed incorporating the findings from the literature and the results of the case study of the Australian context. This lead to the analysis of current problems, and the development of a new framework aimed at meeting the needs of stakeholders more effectively.

3 OVERVIEW OF CURRENT THEORY

Land administration exists fundamentally as a process to manage land. Management of land requires that effective risk management takes place in order to reduce effects should an event occur. In order to carry out effective risk management however, the information contained within land administration systems is required by government, business and citizens. This relationship between risk management and land administration highlights the importance of both, and the importance they represent for government, business and citizens.

3.1 Risk theory and Risk Management

A general definition of risk provided by Vaughan and Vaughan (1996) is a condition in which there is the possibility of an adverse deviation from a desired or expected outcome. More simply put this means the chance that a loss will occur. Critical to the concept of risk is the understanding of the two factors, perils and hazards, which when combined result in the chance of a loss. Peril in this instance is described as the immediate cause of a loss (Trieschmann et al., 2005; McLeman and Smit, 2006) and can be classified as either a human peril (theft, vandalism, fraud, war, terrorism) or a natural peril (earthquake, flood, storm, tsunami, bushfire). Hazard can be defined as a condition that increases the probability of a loss, the severity, or both (Athearn et al., 1989; Teale, 2008). Within the scope of this research the definition of risk is limited to pure risk – instances where there is no possibility of a gain such as fire, theft or flood where the result is either a loss is suffered, or at least no loss is suffered.

Risk management is defined by Standards Australia & Standards New Zealand (2009) as the policies and processes of communicating, establishing the context, identifying, analysing, evaluating, treating, monitoring, and reviewing risk. Within this framework identified in the standards, the communicating and establishing of the context sections are concerned with the necessary background information required to get an understanding of what are the potential risks. The location of land and property can be considered an important aspect in this section of the framework. The next part of the framework is centred on identifying, analysing and evaluating the risks. In terms of land and property, this section involves recognising and identifying the risks that threaten land and property, and analysing and evaluating the degree of risk each threat presents. Identification of possible hazards and perils would also occur in this section of the framework.

Once this section is completed the critical phase of treatment is next. Within this section the options of how to deal with each risk is presented. Available as choices are four options: avoidance of the risk, reduction of the risk, transference of the risk, and retention of the risk (Figure 1). To avoid the risk is the refusal to accept a risk exposure, and in terms of land and property an

example would be to refuse to live in an area that is known to be at risk – avoiding the coast to eliminate the risk of sea level rise. Reduction of a risk refers to mitigation where actions can be taken to reduce the cause of the risk or reduce elements that can contribute to the risk. An example of this would be to clear trees and dry plant matter from the immediate vicinity of a property if the risk of bushfire is known. Transference of a risk refers to the exchange of a risk from one party to another party such as taking out insurance. The final option is a result of no other option being available or no other options chosen as appropriate. When this occurs the risk is kept and no changes are made.



Figure 1: The risk treatment cycle

Within the risk management process three main stakeholders can be identified. They are citizens, business, and government (at all levels). Citizens are the people of a particular social, political, or national community and have a relationship with risk based on location. An understanding of the risks associated with a particular location is required in order for appropriate action to be taken to manage the risks and reduce the possibility of loss. Information about the spatial nature of the risks is required so that effective treatment of the risks can take place. Without an understanding of where each risk exists spatially, management and treatment is difficult.

Among all of the businesses that partake in risk management, insurance by far is the largest. Insurance companies have a direct relationship to risk management. They have a role to provide citizens with a means to protect the value of their property and assets. The traditional way for this to occur is for insurance companies to accept an agreed upon amount as a premium in exchange for protection from a financial cost should a risk event eventuate. In order for this exchange to happen however, insurance companies need extensive land and property information. This information needs to detail the location of the properties they insure so that types of risk that are present at a location can be understood before premiums are agreed upon.

As providing insurance is largely a maths game, the concept of pooling underpins the entire insurance sector. In order to effectively carry out pooling, where risks of similar nature are grouped together to pay for the losses that some might incur, adequate information about land and properties and the potential risks that affect these locations is absolutely critical. The ability to be able to use the laws of large numbers to determine the chance and likelihood of a risk occurring, the average cost of damages, and then spreading this cost over a large number of people allows for society today to operate smoothly and function effectively. However, without understanding the nature of each risk, offering insurance of any type would be too great a risk for an insurance company. Only with adequate information are insurance companies able to carry out calculations which determine the likelihood of a risk occurring, and the cost required in premiums to cover the cost of the risk event occurring. The better the information available, the easier it is to offer protection in the case of loss (Palmer, 1998). Not having access to this critical land and property information causes considerable problems for insurance companies. It means that they either issue

insurance without knowing the full extent or likelihood of a risk occurring, or don't offer insurance at all simply because the unknown present too much risk.

Without the existence of insurance the risk of owning any asset that has the potential to be destroyed, damaged or lost is too high and the benefits of owning it are far outweighed. If this critical spatial information, which supports decisions made by the insurance sector, is not available, which may be as a result of a variety of reasons such as political, institutional, technological, or legal, then society as a whole suffers.

Governments at all levels have a role in managing risk. As the overarching authority in society they have a role to protect citizens and maintain basic security and public order. As a part of this role, the government manages the use of land and property. Making decisions regarding what can be built and where it can be built is largely up to governments. As such they need to be well informed and able to take into account all relevant factors such as risks. The decisions made by government in terms of land and property are recorded in land administration systems, however access to this information for other parties is not always permitted. Allowing access to this critical information is necessary for the management of risks and governments should take action to allow dissemination of this important information to the public.

3.2 Land Administration Systems

Most risks have a relationship to land. The location and nature of the land determines what and where will be at risk, and people develop connections to land based on these factors. This is because land is not just the earth that people walk on; it is fundamentally the way people think about place. Human life depends on having land on which to live and work. It represents territory, opportunity, wealth and prosperity, and can be an essential element for survival. This fundamental and crucial function that land holds in society has lead to the development of rules dictating how land should be managed over time. As time progresses the complexity of the relationships between people and land evolves and a system is required to explain the workings of the arrangements and to provide an understanding of the relationships. The discipline of land administration is the key. Land administration is the growing and evolving discipline about land administration systems (LAS) and is the process of administering the complex rights, restrictions, responsibilities, and (increasingly) risks related to land and its use. A land administration system provides a country with the infrastructure to implement land-related policies and land management strategies (Williamson *et al.*, 2010).

A commonly used tool within the land administration discipline is the land management paradigm which defines the different elements, their relationship, and their role within a land administration system (Figure 2).

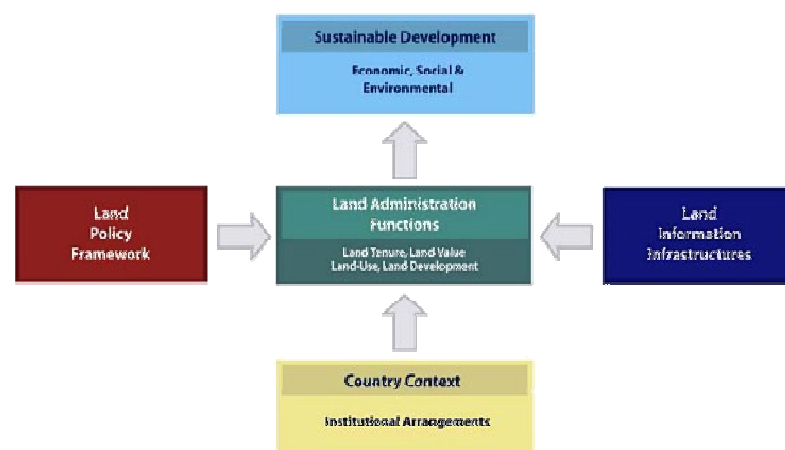


Figure 2: The land management paradigm (Enemark et al., 2005)

The land management paradigm incorporates the country context, land policy framework, land information infrastructures, and the land administration functions in support for sustainable development. At the centre of the paradigm, embedded within the land administration functions, are the components land tenure, land value, land use, and land development. These components,

when implemented effectively lead to efficient land markets and effective land use management which in turn contributes to economic, social, and environmental harmony.

As mentioned earlier, a primary focus of land administration involves administering the complex right, restrictions, responsibilities and risks related to land and land use. The first three elements (rights, restrictions, responsibilities) are considered traditional to property and are linked to the land. These elements are able to be identified due to their nature which requires that they are regulated by law. The fourth element risk is a new consideration for land administration. Unlike the traditional rights, restrictions, and responsibilities which are well defined, risks remain largely unknown and thus can exist with a relationship to land without this relationship ever being revealed. Only recently has the importance for understanding and identifying risks to property received attention. Land administration systems that are based on 19th century models are in charge of managing land and property, consideration for 21st century issues such as social, economic, and environmental risks have not been taken into account.

4 RESULTS – A NEW APPROACH

Analysis of the Australian context revealed the stakeholders involved and the infrastructures available for managing risks to land and property. The variety of risks that affect land and property has been shown. The arrangements and interactions between these components in the Australian context are discussed here.

4.1 Case Study of the Australian Context

Increasingly in Australia there are risks that are of national significance. Examples of these risks are flooding, bushfire, sea level rise, drought, asbestos, and insulation. Events that are a result of these risks are becoming more frequent. We are seeing more frequent storm events which results in flood damage to many properties, the worst bushfire events in the history of the country causing severe damage to property and loss of life, and asbestos concerns as a result of bad management in the past.

In order to manage these risks, spatial information and spatial technologies as well as land administration systems need to be taken advantage of. The accurate and timely land and property information that is kept within the land administration systems underpin risk management decisions made by governments, business, and citizens, however is not accessible by these stakeholders because of a variety of institutional, technical, political and economic reasons. The limited sharing and aggregation of this fundamental land and property information between jurisdictions and organisations is impacting on the ability of stakeholders to make effective risk management decisions and threatens financial stability.

An example that demonstrates this is the current lack of comprehensive insurance available in Australia. Most notably, flood insurance against riverine flooding is not generally available despite an estimated 160 000 Australian homes being at risk of a 1 in 100 year flood. Insurance companies as a major stakeholder in the risk management process have a role to provide individuals with a means to protect the value of their property and assets, however, current insurance arrangements mean that consumers who are aware of the flood risk and want to manage the risk are unlikely to be able to obtain insurance at all, or at an affordable premium. The Insurance Council of Australia wish to resolve the problem and achieve greater market availability of flood cover for the majority of households, but in order to do so they require access to fundamental land and property information such as floodplain information (Insurance Council of Australia, 2006). This information for a variety of institutional, technical and economic reasons is not available in Australia for use by insurance companies, and prevents citizens and business from being able to manage their risk.

The reality of this problem has been highlighted with the recent Queensland flood events. Hundreds of properties were inundated by the water that overflowed from the swollen Brisbane River which resulted in houses being severely damaged. Insurance for these victims is limited for this type of event, and reconstruction of the damaged properties and the replacement of damaged contents in many cases will be left as the responsibility of each individual owner.

Another example that demonstrates the limited sharing and aggregation of land and property information between jurisdictions, and the effect that this has on effective risk management decisions is climate change. In Australia recent predictions have indicated sea level rise will be a result of climate change. These predictions have labelled some coastal areas vulnerable to inundation and consequently, as areas that are unfit for development. Decisions that are made regarding development are the responsibility of the local council. Of late however, some of these decisions made by the local councils have been overturned by a higher planning authority who claim councils are ignoring the risks of predicted sea level rise to developments in the area. As a result councils are confused and argue that the level of information and guidance by governments for councils isn't complete and doesn't indicate how sea level rise and climate change ought to be dealt with by councils. This information is critical for effective risk management decisions to be made and must be made available so that the risks associated with climate change can be understood by decision makers to ensure the safety of Australians living in these vulnerable areas.

4.2 Development of a Risk Framework

The issues raised in the case study and in the previous study of current literature all conclude with the same message: that accurate and timely spatial information and land and property risk information is fundamental for effective risk management. However, currently in Australia there is no infrastructure that facilitates the coordination, sharing, aggregation and dissemination of consistent information on risk. As a consequence the ability of government, business and citizens to manage risk to land and property is limited. Figure 3 below demonstrates the current arrangement of land and property information, risk information and stakeholders in Australia.

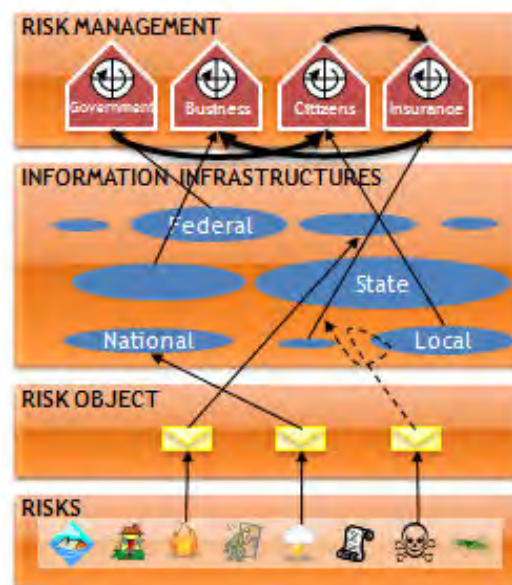


Figure 3: The current situation – Australian context

The framework above (Figure 3) was developed to give an overarching view of the current arrangements of all stakeholders, information infrastructures and risks in Australia currently. Within the Australian context the problems associated with the coordination, sharing, aggregation and dissemination of information is exacerbated by the out of date administrative arrangements that exist. Australia as a federated country is made up of six states and two territories – all with their own land administration systems. On top of these arrangements exists the three levels of government which includes federal government, state government – one for each of the eight jurisdictions, and local governments, which sit within each separately defined area in a state or territory.

Within the figure sits four separate layers, each representing a different element in the complex problem. The first layer from the bottom is the risk layer which represents the risks that

are present in society today that affect property. Examples of these risks are flooding, earthquakes, bushfire, fraud, storms, rights restrictions and responsibilities, asbestos, and pests. The second layer of the diagram represents the risk object. The risk object represents a relationship between a person and a risk. A risk object is created when a person that is affected by a risk (has a relationship with the risk) recognises that relationship. The risk object reflects this recognition because without an understanding of the risk or acknowledgement of the existence of the risk the risk is unable to be managed. Managing a risk is an active choice so the creation of the object is necessary before the stage of risk management can take place. The third layer of the diagram represents the information infrastructures that exist. These infrastructures are present in all levels of government – federal, state and local, as well as in businesses – national, state and local, and range from authoritative land administration systems, to other infrastructures which hold important land and property information. The final layer of the diagram is the risk management layer. This layer represents all the stakeholders that are involved in the risk management process. These stakeholders are able to carry out risk management by making choices about how to deal with risk that is relevant to them. As a part of this process the stakeholders are able to transfer the risk – one treatment option, to each other. For example the citizen might transfer the risk to an insurance company, which the insurance company may then transfer to another business to reinsure the risk.

5 DISCUSSION

Results from research in the areas of risk management, risk information and land administration system all argue that good information is required for the risk management process to be effective. The necessity of making this information available is also strongly put forth. This message is relevant for all major stakeholders, whether they are government, business or citizens.

The information that is required when making risk management decisions is ‘where’ is the risk. If where the risk exists can be understood, then it can be more effectively managed. Moreover, understanding ‘where’ the risk exists allows for an understanding of ‘who’ is affected by the risk to be discovered. With this information smart risk management choices can be made by stakeholders.

Spatially enabling the information about risk to property is a possible solution. Spatial enablement of this type of information would allow access to the information in a simpler and easier way based on a specific location. The capability to retrieve information about potential risks to land or property based on the location of that land or property enhances a stakeholder’s ability to make smarter, safer choices about risk and to more effectively manage risks. A model showing how spatially enabled information would operate is depicted in Figure 4.

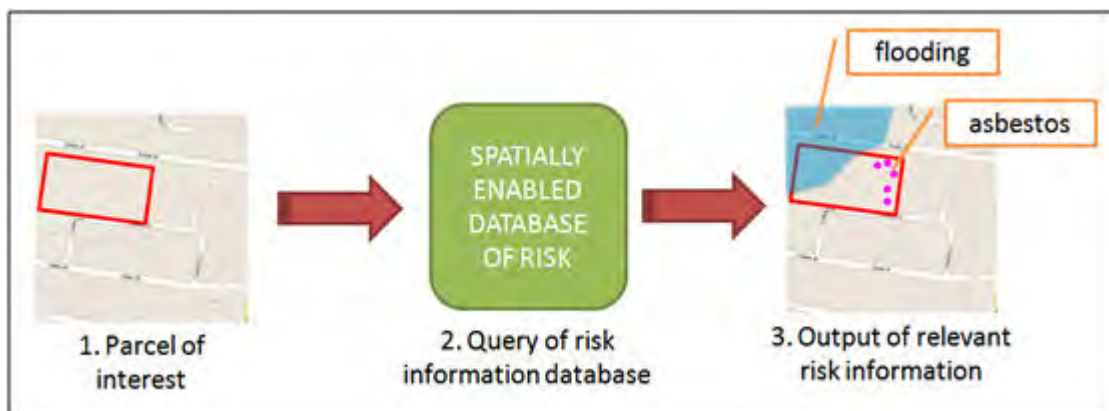


Figure 4: Spatial enablement of risk information

Spatially enabling the information makes it easily accessible and interpretable for stakeholders. They can visualize how different disasters or risks affect them specifically based on location.

From the preliminary research carried out which aimed to develop a framework to explore the idea of risks as spatial objects Figure 3 was developed. Within this framework the risk object concept was expanded on and the relationship between the risks, the information infrastructures

and the stakeholders was explored. The problems associated with the current model are discussed further below.

- **Risks**

The risks layer represents the risks and the information available describing the nature of each risk. This information is needed by government, business, citizens and insurance companies for effective risk management to be carried out. In order for this risk information to make its way to the risk management level, it needs to pass through an information infrastructure situated in the middle layer. The problem is that in Australia many different information infrastructures currently exist, so when the risk information is passed into the information infrastructure it may end up in one of many infrastructures, or it may end up in no infrastructure at all. Because there is currently no aggregated system each information infrastructure is operating in an ad-hoc and uncoordinated fashion which contributes to existing problems associated with obtaining important information about risks related to land and property.

- **Risk Object**

The risk object represents the recognition of a particular risk by the person who has a connection or relationship to that particular risk. Problems arise when a risk object is created for a certain risk; however the information about this risk does not belong anywhere or is not recorded in a way that the stakeholders can access this information. An example of this might be asbestos – where location information about the risk might not be known and therefore may not be recorded. Challenges such as this need to be addressed to ensure information about all relevant risks can be made available for stakeholder use.

- **Information Infrastructures**

The information infrastructures layer represents all of the locations where information about land and property is stored. This includes information about rights, restrictions, responsibilities and risks. This information is critical to all stakeholders, especially the citizens who in most cases seek out this type of information to help them make decisions about property they intend to purchase and/or live in. For a citizen this type of investment is a big deal and could result in financial ruin if smart decisions are not made. As a result of the potential consequences citizens require as much information about the land and property as possible. The ability to be able to access this information quickly and easily would be of great benefit to citizens and also to other stakeholders however, information of this type is not always centrally located or easily accessible due to the current arrangements of the information infrastructures. When information describing risks is passed into the information infrastructures situations arise where the risk information may end up in one of the many infrastructures, or it may end up in no infrastructure at all. Attempts by stakeholders to find this information then become futile, because where to look is not known, nor whether it even exists.

- **Stakeholders**

The stakeholder layer represents the parties involved that have an interest in risk management. Within this layer the stakeholders gather information required for them to make risk management decisions and to carry out their chosen risk management treatments. The information they require about risks affecting their land and property is contained within the information infrastructures. A problem that arises from this scenario is that the relationship between the stakeholders and the individual information infrastructures determines what information is available and to which stakeholders the information is available to. Not all information is available to all stakeholders, however all available risk information is needed for effective risk management. Access to this information is critical because the risk management cycle relies on timely and accurate information about the nature of property risks and who they apply to. The current information infrastructure arrangements prevent this from happening.

5.1 Proposed System

Responding to the problems raised regarding the current arrangements, a framework addressing these problems has been developed. The framework (Figure 5) takes into account the current situation as shown in Figure 3 and transforms the information infrastructures layer of the diagram into a one connected infrastructure that uses a land administration system as the platform.

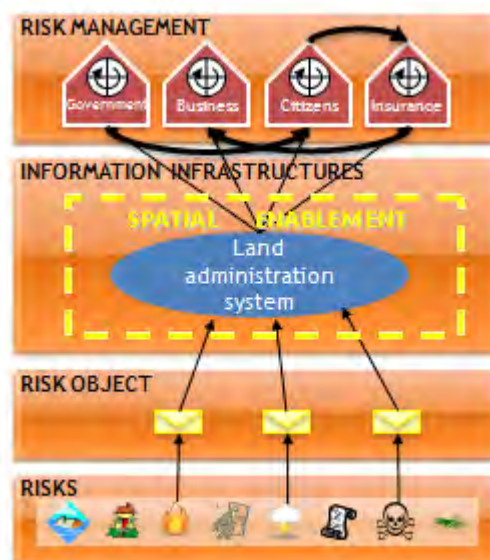


Figure 5: The proposed framework for aggregating and disseminating consistent information about risk to all stakeholders.

For this proposed model a land administration system was selected as the platform because it is shown to be an effective infrastructure for managing and integrating land information (Williamson and Wallace, 2007). Already land administration systems have the capability to manage rights, restrictions and responsibilities relating to land, so extending the role of these systems towards managing risks is a logical progression. The proposed framework demonstrates an infrastructure that is able to coordinate all of the current information infrastructures to allow for timely, accurate, and aggregated information about risks that relate to land and property to be facilitated.

To address the problems highlighted, further spatial enablement of the information within the land administration system has been included in the proposed framework also. The aggregation of the information infrastructures would allow for the information to be searched across one platform, and spatial enablement of the information would greatly enhance the effectiveness of the system for the stakeholders attempting to gather information. A search queried into the platform would return all of the relevant information based on the spatial located searched. Further research into the specific architecture of this framework is required.

6 CONCLUSION AND FUTURE DIRECTION

This research has given an overview of the current understanding of the way that risks to land and property and the information that describes these risks relate to stakeholders and information infrastructures. The framework developed provides an understanding of the wider context and uses Australia as a case example; however the framework requires further empirical testing. How the situation can be improved needs to be explored further.

Two case studies are planned with the intention of gathering an understanding of the problem in more detail. The first case study will be aimed at investigating and understanding each different layer of the diagram (Figure 3) to gain an understanding of how each layer operates and how the different layers interact with each other. The second case study will look specifically at land administration systems with the aim to understand how land and property information is

managed in each different land administration system that already exists and to investigate the capability for these systems to expand their operations to be able to manage risk information. The information from both these case studies will be brought together and will provide further insight into the issues, as described in this paper, and will highlight possible solutions to resolve issues surrounding this problem.

ACKNOWLEDGEMENT

We would like to acknowledge the assistance of our colleagues in the Centre for SDIs and Land Administration and the Department of Infrastructure Engineering at the University of Melbourne in the production of this paper. We would also like to thank our research partners for their support and provision of qualitative data. These partners are: Land and Property Management Authority (NSW Government), Land Victoria (Victorian Government), Landgate (Western Australian Government), PSMA Australia Limited, and the Australian Research Council. It should be noted that the views expressed in this paper are those of the authors and may not represent those of the research partner organizations.

REFERENCES

- Athearn, J. L., Pritchett, S. T., & Schmit, J. T. (1989). *Risk and Insurance* (6th Edition ed.). St Paul, Minnesota: West Publishing Company.
- Enemark, S., Williamson, I. P., & Wallace, J. (2005). Building Modern Land Administration Systems in Developed Economies. *Spatial Science*, 50(2), 51-68.
- Insurance Council of Australia (2006). *ICA Response to the Report to the Council of Australian Governments on Natural Disasters in Australia*. Insurance Council of Australia.
- McLeman, R., & Smit, B. (2006). Vulnerability to climate change hazards and risks: crop and flood insurance. *The Canadian Geographer*, 50(2), 217(210).
- Palmer, D. (1998). Security, risk and registration. *Land Use Policy*, 15(1), 83-94.
- Standards Australia, & Standards New Zealand (2009). *AS/NZS ISO 31000:2009 Risk management - Principles and guidelines*.
- Teale, J. (2008). *Insurance and Risk Management*. Milton, QLD: John Wiley & Sons Australia Ltd.
- Trieschmann, J. S., Hoyt, R. E., & Sommer, D. W. (2005). *Risk Management and Insurance* (12 ed.). Mason, Ohio: South-Western, Thomson.
- Vaughan, E. J., & Vaughan, T. M. (1996). *Fundamentals of risk and insurance* (7th Edition ed.). New York, USA: John Wiley & Sons, Inc.
- Williamson, I. P., & Wallace, J. (2007). *New Roles of Land Administration Systems*. Paper presented at the International Workshop on Good Land Administration - Its Role in Economic Development, Ulaanbaatar, Mongolia, July 2007
- Williamson, I. P., Enemark, S., Wallace, J., & Rajabifard, A. (2010). *Land Administration for Sustainable Development* (1ed.). Redlands, California: ESRI Press.

Analysing the Effect of Drought on Net Primary Productivity of Tropical Rainforests in Queensland using MODIS Satellite Imagery

Sipho S.T. Shiba

Ministry of Agriculture
Department of Land Use Planning
P. O. Box 5893, Manzini, M200 Swaziland &
Faculty of Engineering and Surveying
University of Southern Queensland
Toowoomba 4350 QLD Australia
Email: sthokozane@yahoo.co.uk

Armando Apan

Faculty of Engineering and Surveying &
Australian Centre for Sustainable Catchments
University of Southern Queensland
Toowoomba 4350 QLD Australia
Email: apana@usq.edu.au

ABSTRACT

The changing climate is characterised by increased frequency and severity of extreme weather events that include cyclones, drought, flooding, heatwaves and snowing. Such climatic patterns are predicted to have negative effects on tropical forests. Given the role played by these forests in hydrological cycle, carbon cycling and biodiversity, it is crucial to quantify these effects to ensure adequate monitoring and management of these forest ecosystems. The aim of this study was to analyse the effect of drought on the Net Primary Productivity (NPP) of tropical rainforests. The specific objectives were to a) determine existing trends of NPP over the study period, b) assess the relationship between NPP and climatic - biophysical factors, c) analyse the effect of drought on NPP, and d) evaluate inter-annual variation of NPP.

Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery, thematic data sets and climatic data sets were acquired and analysed using GIS techniques, image processing tools and statistical techniques. Correlation and regression analyses were used to establish relationship between NPP and climatic-biophysical factors. NPP had a positive trend over the study period even though there was a significant reduction of NPP in 2003 and 2006 as a result of drought events in 2002 and 2005, respectively. A significantly high r^2 (0.8) was achieved when comparing rainfall of current year and NPP of subsequent year. In this study, it was concluded that rainfall and drought had a lag effect on NPP. The interannual variation of NPP in this region was significantly ($P < 0.001$) influenced by the rainfall amount ($r = 0.8$), maximum temperature ($r = 0.8$), solar radiation ($r = 0.84$) and potential evaporation ($r = -0.61$). The influence of temperature was evident within distinct periods (cycles). The index representing soil water holding capacity in the B horizon was found to have a positive relationship ($r = 0.82$) with NPP. The other biophysical factors (slope, DEM, aspect and terrain wetness index) did not exhibit strong correlations with NPP, indicating a weak influence. The influence exerted by the interaction of biophysical and climatic factors on NPP requires more research.

KEYWORDS: Net Primary Productivity, tropical rainforests, drought, MODIS

1 INTRODUCTION

Tropical forests play a significant role in the carbon budget. They store significant amount of carbon in their woody biomass. The rate at which energy is converted into plant biomass is called Primary Productivity (Roy & Saugier, 2001). Net primary productivity (NPP) represents the net amount of carbon added to plant biomass per unit of space and time. Terrestrial NPP is the initial step in the carbon cycle. Studies have indicated that terrestrial ecosystems are becoming a major sink in the global carbon cycle (Houghton, 2007). A major contribution in sequestering carbon comes from tropical forests and they have become part of the mitigation strategy proposed by Intergovernmental Panel for Climate Change to address increasing carbon emissions (IPCC, 2007).

The role played by tropical forests seems to be threatened by changing climatic patterns. Changes in climate include predictions of increased frequency of extreme weather events such as cyclones, droughts, floods, heat waves and snowing (Whetton, 2003). Some studies have reported that extreme droughts, usually associated with El Nino Southern Oscillation (ENSO), causes higher rates of tree mortality and increase forest flammability (Nepstad et al., 2004). However, contrasting observations have been made where excessive greening has been associated with drought in the Amazon tropical forests (Saleska et al., 2007). It is evident that the sensitivity of tropical forests to drought is poorly understood.

Climate change has been associated with increased frequency and severity of drought episodes (Salinger et al., 2000). Drought has been defined as a period of abnormally dry weather, sufficiently prolonged because of a lack of precipitation, causing serious hydrological imbalance and has connotations of a moisture deficiency with respect to water use requirements (McMahon and Arenas 1982). The deficiencies have impacts on both surface and groundwater resources and leading to reductions in water supply and quality, reduced ecosystem productivity, diminished hydro-electric power generation, disturbed riparian and wetland habitats, and reduced opportunities for some recreation activities (Riebsame et al., 1991).

In Australia, drought is defined as a condition when the rainfall over a three-month period is in the lowest decile (i.e. lowest 10%) of what has been recorded for that region in the past (BoM, 2011a). The worst drought in the past ten years occurred in 2002 with the Australian average rainfall of 472mm. It was the fourth lowest on record since 1902 where it was 317mm (BoM, 2011a). Drought has been identified as a precursor of forest fires, thus generating more carbon emissions. This is downplaying forests' role in the mitigation strategy to reduce carbon concentration. The prediction of increased frequency of extreme weather events such as drought will continue altering species abundance, composition and dominance, and reducing favourable habitats to the point of extinction in some cases (Thomas et al., 2004).

It has been argued that measuring and quantifying Net Primary Productivity (NPP) is one of several methods to monitor response of tropical rainforests to the recurring extreme weather events (Cramer et al., 1999). The variation of NPP driven by global change has been one of the most important aspects in climate-vegetation studies. Understanding the suite of factors influencing NPP is one of the main interests for scholars in this field. This will allow development of models that will predict variation of NPP with changing climate.

The launch of MODIS in 1999 has provided an array of data outputs that give useful information in the monitoring of certain phenomenon in our planet, including monitoring biophysical aspects of vegetation (Zhao et al., 2004). The operational MODIS Production Efficiency Model (MOD17) on the Terra satellite is used to generate a composite of 8-day near-real-time vegetation primary production (Heinsch et al., 2003; Zhao et al., 2005). MOD17 is based on the radiation use efficiency logic which state that the NPP of well-watered and fertilized annual crop plants is linearly related to the absorbed photosynthetically active radiation (APAR) (Monteith, 1972). The study by Zhao et al. showed that tower-based and MODIS estimates of annual GPP compare favourably for most biomes, although MODIS GPP overestimates tower-based calculations by 20%–30%.

There are several factors that influence net primary productivity. Studies have shown that forests growing on soils with relatively high fertility and moisture content will have high productivity (Field et al., 1995; Delucia et al., 2007). These factors have an indirect impact on the process of photosynthesis. Different vegetations react differently to changing seasons and rainfall patterns (Huete et al., 2006). Some plants have better coping strategies during drought periods, thus

minimising the negative impact on net primary productivity. It has been argued that the effects of climatic factors on NPP are sometimes delayed in certain biomes due to interaction of climatic and physical factors. The current study is exploring the lag-effect concept in tropical forests ecosystems.

The response of tropical rainforests to drought has been a subject of debate in the past few years. Some have shown that tropical forests green up during drought implying an increase of net primary productivity during the drought episodes (Saleska et al., 2007). However, recent studies (Samanta et al., 2010) have refuted these results indicating that only 11- 12 % of the forests greened up during the 2005 drought in the Amazon; 28-29% showed browning or no change; and the rest of the forest could not be categorised due to poor quality of the data used.

A study by Nemani et al. (2003) which explored variation of NPP between 1982 and 1999 showed a positive trend in NPP in spite of the occurrence of drought episodes within the study period. The increase was attributed to global warming and increasing concentration of CO₂. Recently, a study by Zhao et al. (2010) reported a decline in NPP between 2000 and 2009 as a result of drought episodes. However, their finding has caused alarm in some quarters (Sherwood and Craig, 2011). The ongoing debate indicates that the sensitivity of tropical rainforests to drought episodes is still poorly understood. The debate has motivated the current study.

The aim of the study was to analyse the effect of drought on the Net Primary Productivity (NPP) of tropical rainforests. The specific objectives were to a) determine existing trends of NPP over the study period, b) assess the relationship between NPP and climatic - biophysical factors, c) analyse the effect of drought on NPP, and d) evaluate inter-annual variation of NPP. MODIS satellite imagery (2000-2006), thematic data sets, climatic data sets and other relevant data sets were acquired and analysed using GIS techniques, image processing tools and statistical techniques.

2 DATA AND METHODS

The study was confined to the tropical rainforests found in Queensland, Australia, within a biome region known as the Wet Tropics. The Wet Tropics Bioregion covers approximately two million hectares. Two main sites were identified, namely Daintree National Park and Wooroonooran National Park which are relatively close to Cairns, one of the major cities in Australia (Figure 1). Both national parks are listed under the World Heritage Area. The rainforests in these areas are closed forests characterised by dense foliage in the upper layers (> 70% foliage cover), with vines, epiphytes and mosses form a conspicuous and important element of the structure (Department of the Environment and Water Resources, 2007). The main reason for selecting protected areas as a study area was to minimise the influence of potential disturbance (e.g. forest clearing) during the study period (2000-2006).

The mean annual rainfall in Cairns is about 2,200 mm. However, rainfall is highly variable in the wet tropics ranging from 1,200 mm to 6,000 mm. The area has a distinct wet and dry season with most of the rainfall received during the summer months. The mean minimum temperature is about 18°C while mean maximum temperature is approximately 31°C. Daintree National Park is about 100 km from Cairns. It has been described as one of the oldest forests in the world. It has about 57,000 ha of tropical rainforest. On the other hand, Wooroonooran National Park is located about 25km from Cairns. The two highest mountains in Australia are located within the Wooroonooran National Park boundary. It has about 115 000 ha of tropical forest.

The Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery (1km spatial resolution) was downloaded from the NASA MODIS Land Science team website for the period 2000 to 2006. In this study, the MOD 17A3, version 4.8 GPP/NPP product was used and it is designed for the study of the net amount of carbon fixed in vegetation by the photosynthesis process. Each pixel of the MODIS NPP product is expressed as grams of carbon per m² per year (gC/m²/yr). In the production of MOD 17A3, corrections for gases and clouds are implemented. The MODIS GPP product showed seasonal variation that was generally consistent with the in situ observations (Turner et al., 2006). Validation at "stage 3" has been achieved for the global and net primary production products. The global annual estimates of GPP and NPP are within 10.4% and 9.0% of average published results, respectively (<http://landval.gsfc.nasa.gov>).

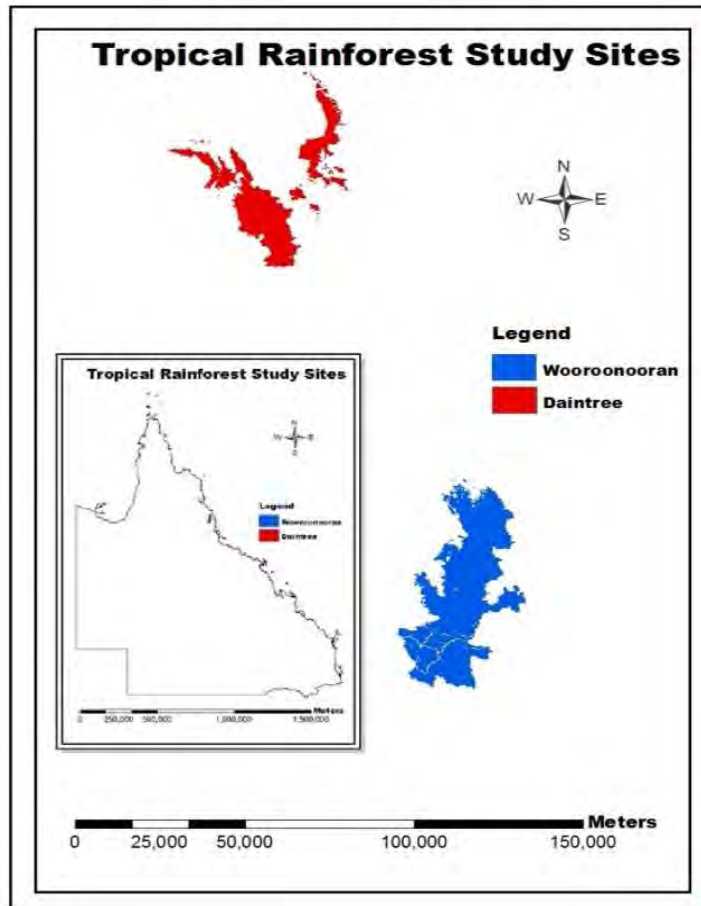


Figure 1: Study sites for tropical rainforests in Queensland

Thematic map data sets were acquired from the following data custodians or agencies:

- Queensland Department of Environment and Resource Management (regional ecosystems map, national park and reserves boundary, land use, DEM, locality, geology, land use, land cover, foliage projective cover, etc.);
- Australia's Bureau of Meteorology (mean annual rainfall, mean maximum annual temperature); and
- CSIRO (available soil water holding capacity, incoming solar radiation, and potential evaporation).

The pre-processing step included mosaicing within the ERDAS IMAGINE software, combining the four MODIS image tiles into a big digital image covering Queensland. The Queensland boundary map layer was used to clip the mosaic image to eliminate areas outside Queensland. The image was further clipped by the national park boundary map to focus on the wet tropics bioregion. All other GIS datasets were geo-referenced and clipped to a common map projection (Geocentric Datum of Australia 1994) using ArcGIS 10.

The "sample" tool within ArcGIS was used to extract pixels values from these raster maps. The generated tables were exported to Excel for statistical analysis. Regression and correlations analysis were used as statistical tools to investigate relationships between NPP and climatic factors. The t-test was used to compare NPP of drought years and NPP of non-drought years. There has been a suggestion by Steele et al. (2005) that effects of drought may be delayed. This was incorporated in the analysis by comparing NPP of a drought year with NPP of the subsequent year.

3 RESULTS AND DISCUSSION

3.1 NPP Trend

The results showed a positive trend in NPP from 2000 to 2006 even though there was a drop in 2003 and 2006. This trend was illustrated by the fitted line on the NPP graph (Figure 2). The obtained results are similar to a conclusion reached by Nemani et al. (2003). Their study investigated the trend of NPP between 1982 and 1999, and concluded that recent climatic changes have enhanced plant growth in northern mid-latitudes and high latitudes. The enhanced plant growth resulted in an increase of 6% NPP globally. The largest increase occurred in tropical ecosystems. Several factors were attributed to the positive trend, including the increase in carbon dioxide concentration which has been described as having fertilization effect (Norby et al., 2005).

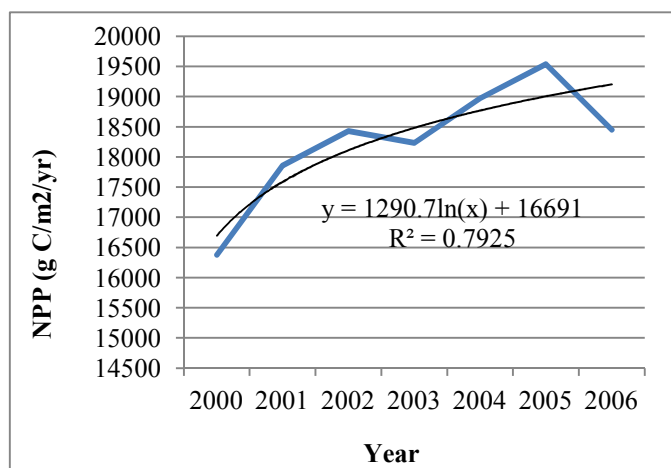


Figure 2: NPP trend between 2000 and 2006

However, recent studies have reported a gradual change of the global trend observed by Nemani et al. (2003). The change was attributed to increased frequency of drought events (Zhao & Running, 2010). In their study, Zhao and Running observed a reduction in the global NPP of 0.55 pentagrams of carbon over the period 2000-2009. The drought episodes that induced the reduction occurred in 2002 and 2005. The positive trend observed in the current study may be due to regional differences in the intensity and frequency of drought events. The drought year in the wet tropics had an annual rainfall of 1591 mm, yet in other regions of the world the drought threshold was around 500 mm annually. Therefore, the difference in the severity of the drought may be the explanation for the different trend. Given the prediction of increased occurrence of drought episodes in Australia in the next 50 years (CSIRO, 2001; IPCC, 2001), a steady decline in NPP may be expected

3.2 NPP and Rainfall

NPP ranged from 16,373 gC/m²/yr in 2000 to 18,448 gC/m²/yr in 2006 (Figure 3). The lowest NPP in the year 2000 surprisingly corresponded to the wettest rainfall of 4307 mm in the same year. According to the reports from Bureau of Meteorology, the year 2000 was Australia's second wettest year on record (<http://www.bom.gov.au/info/leaflets/nino-nina.pdf>). The rainfall decreased, reaching the lowest rainfall in 2002 of 1,591 mm. Unexpectedly, the NPP in the same year (2002) increased, reaching 18,426 gC/m²/yr.

El Nino conditions within the study period occurred in 2002-2003 and 2004-2005. These conditions resulted to a severe drought in 2002 and a mild drought in 2005. The expectation was to have a reduced NPP in 2002. The rainfall increased in 2003 while NPP reduced in the same year. A similar trend was observed from 2004 to 2005 where rainfall reduced by 1,076 mm while NPP increased by 566 gC/m²/yr in 2005. Again, the NPP reduced the subsequent year (2006) even

though rainfall increased. The drop in rainfall seems to have lag effect on NPP where the negative effects of reduced rainfall amount are realized in the subsequent year.

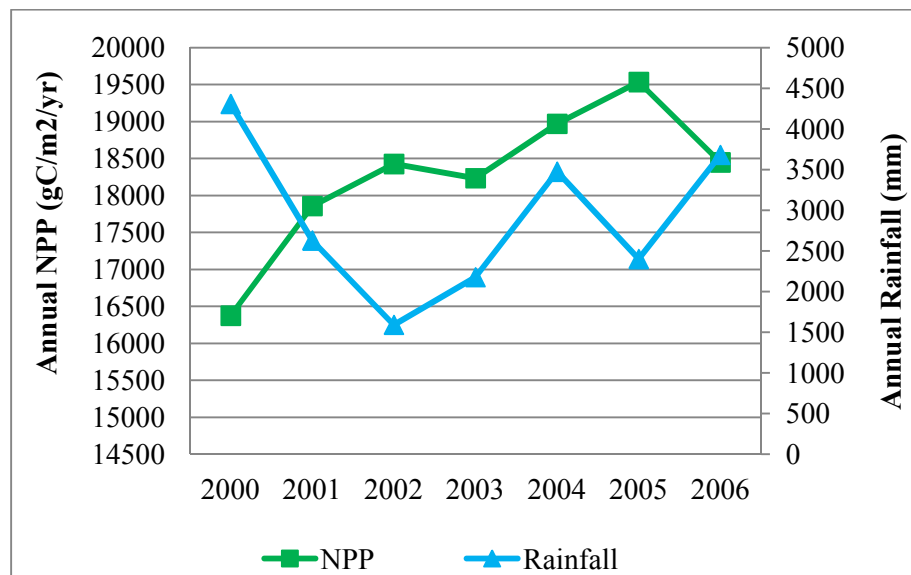
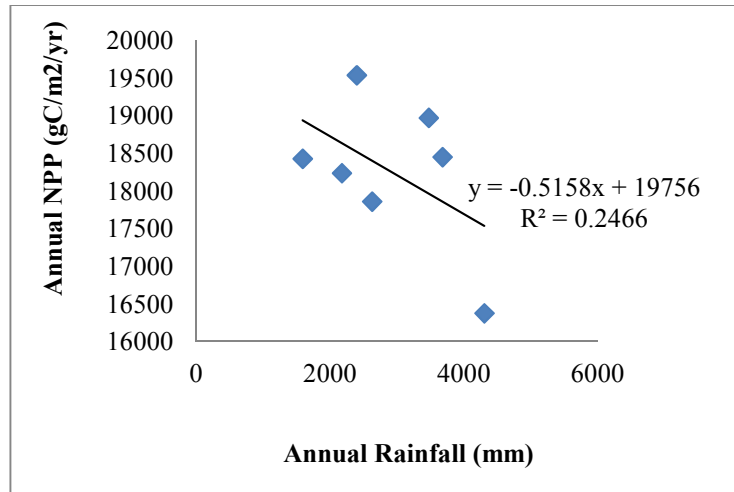


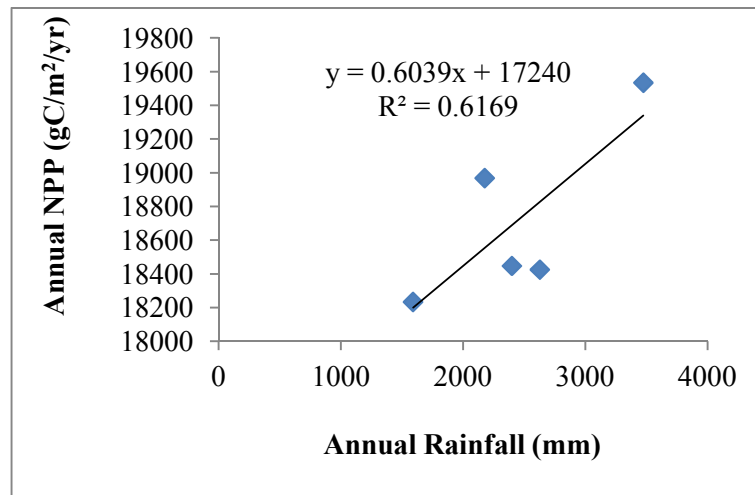
Figure 3: Variation of NPP and rainfall between year 2000 and 2006

3.3 Lag Effect of Rainfall

The lag effect was explored using regression analysis. The first step was comparing means of NPP and rainfall in the same year (i.e. without lag effect). The second step was comparing means of rainfall in the current year with means of NPP in the subsequent year (i.e. lag effect). Figure 4 (a) and (b) show results of the two conditions. The results showed that without lag effect, the relationship was poor with only 25% explained by the fitted regression equation. However, with the lag effect, the relationship improved with 62% explained by the equation. Individual years were compared on pixel basis to assess the lag effect, and it was discovered that the 2001 rainfall was significantly ($p < 0.001$) correlated to the 2002 NPP (Figure 5). The rainfall of 2002 (a drought year) was significantly correlated with 2003 NPP (Figure 6). There could be many factors related to this observation including the complexity of forest and canopy structure, reducing light penetration to the floor and thus reducing water loss through evaporation.



a) Comparison of annual NPP and annual rainfall without lag effect ($r = -0.49$)



b) Comparison of annual NPP and annual rainfall with lag effect ($r = 0.79$)

Figure 4: Comparing annual NPP and annual rainfall

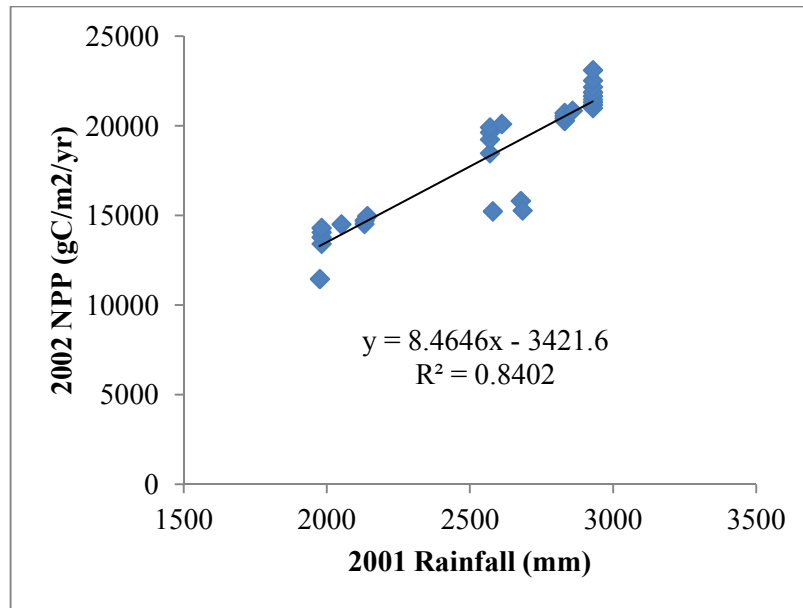


Figure 5: Comparing 2001 rainfall and 2002 NPP

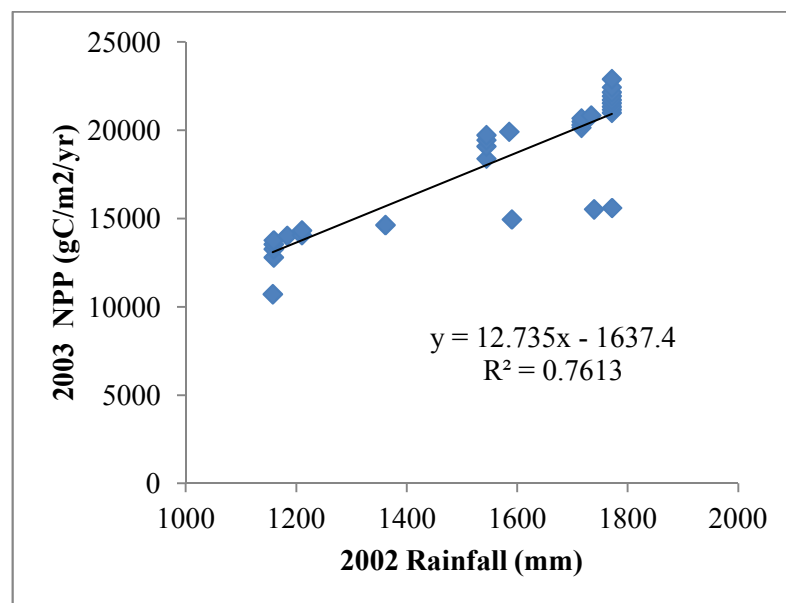


Figure 6: Comparing 2002 rainfall and 2003 NPP

3.4 The Lag Effect of Drought

The study has showed that effects of drought do not normally occur in the same year in which the episode occurs. The years 2002 and 2005 were drought years within the study period. For the rainforests in the study area, the NPP increased during 2002 which was one of the worst drought episodes in many parts of Australia. The observation was similar to that made by Saleska et al. (2007) who studied the effect of drought in Amazon forests in 2005.

The annual NPP in the year 2000 (wet year) was 18,426 gC/m²/yr while in 2002 (drought year) it was 20,479 gC/m²/yr. There was a significant ($P < 0.001$) increase of NPP during the drought year by 2,053 gC/m²/yr when compared to the wet year (Figure 7). The increase in NPP in 2002 was against expectation, as the rainfall significantly reduced due to El Nino conditions that brought severe drought over most parts of Australia.

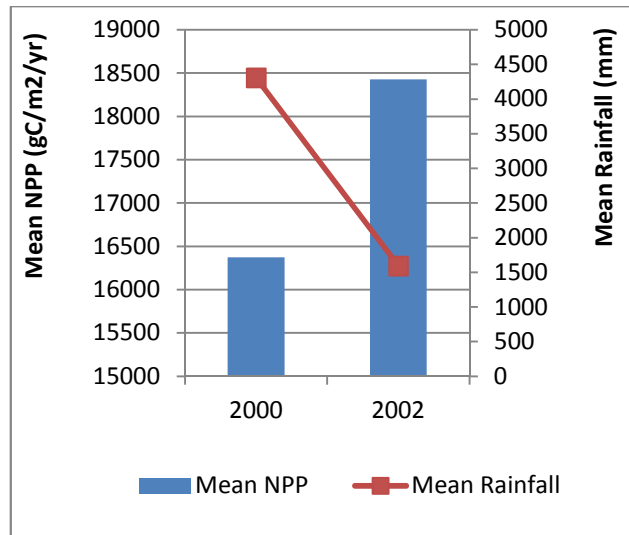


Figure 7: Comparing NPP of drought year and non-drought year

The debate surrounding the behaviour of tropical rainforest during drought conditions has resulted in proposals of assessing the effects in the long term rather than in the same year in which drought occurs. Figure 8 shows a comparison of NPP and rainfall between the drought year and the subsequent year. The NPP during the drought increased to 18,426 gC/ m²/yr when rainfall reduced to 1,591 mm. In the subsequent year (2003), NPP significantly dropped ($P < 0.001$) by 192 g C/ m²/yr compared to 2002. This was in spite of the significant increase ($P < 0.001$) in rainfall in 2003 of about 586 mm. This drop in NPP was associated with drought conditions of the previous year, indicating delayed effects of drought on the tropical rainforest.

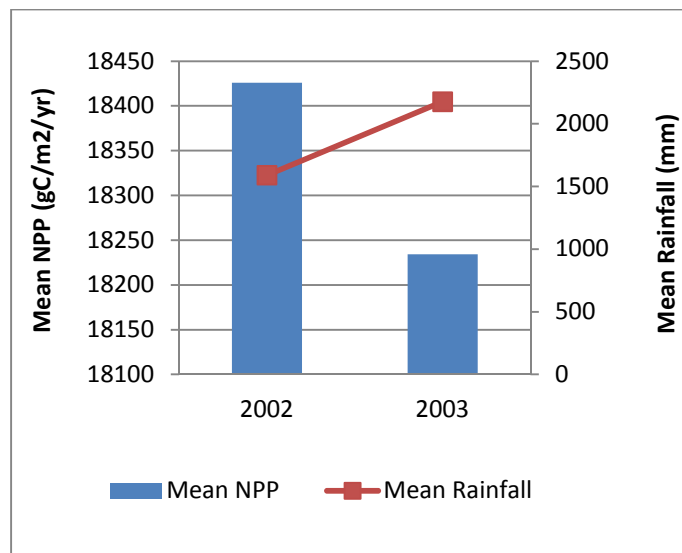


Figure 8: Comparing drought year and subsequent year

A trend similar to 2002 and 2003 was observed between 2005 and 2006. The NPP in 2005 increased even though rainfall significantly dropped by 1,076 mm from previous year. However, in the subsequent year (2006), NPP reduced significantly by 1,087 gC/m²/yr ($P < 0.001$) in spite of the significant increase in rainfall by 1,281 mm (Figure 9). The impact of reduced rainfall in 2005 was realised in the subsequent year.

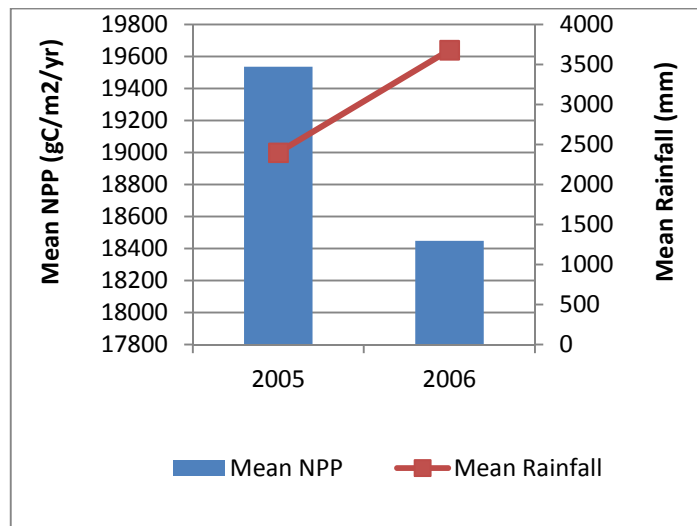


Figure 9: Comparing NPP between 2005 and 2006

It has been shown that in short time scales, climatic factors affect the physiological processes that control plant photosynthesis and growth (Christopher et al., 1995). However, in some vegetation types the effects of climatic factors on NPP are not instantaneous, exerting delayed effects and serial correlation ageing (Steele et al., 2005; Peng et al., 2008). Drought is found to significantly affect ecosystem carbon exchange processes (Krishnan et al., 2006). Analysis of measured atmospheric carbon dioxide and satellite-derived measurements of temperature and the vegetation index suggests that nutrient effects on the carbon cycle can delay ecosystem response to changing climate by as much as two years (Braswell et al., 1997). The current study shows delayed effects of drought events. Therefore long term studies are essential to establish the effect of drought and other climatic and physical factors on NPP.

3.5 NPP and Other Climatic Factors

The overall correlation coefficients of all the climatic factors with NPP are shown on Figure 10. The interannual variation of NPP in this region was significantly ($P < 0.001$) influenced by rainfall amount ($r = 0.8$), maximum temperature ($r = 0.8$), incoming solar radiation ($r = 0.84$) and potential evaporation ($r = -0.61$). Similar results were obtained by Peng et al. (2008). The increasing temperature appears to reduce the cloud cover, subsequently increasing the photosynthesis activity in some regions within the Wet tropics.

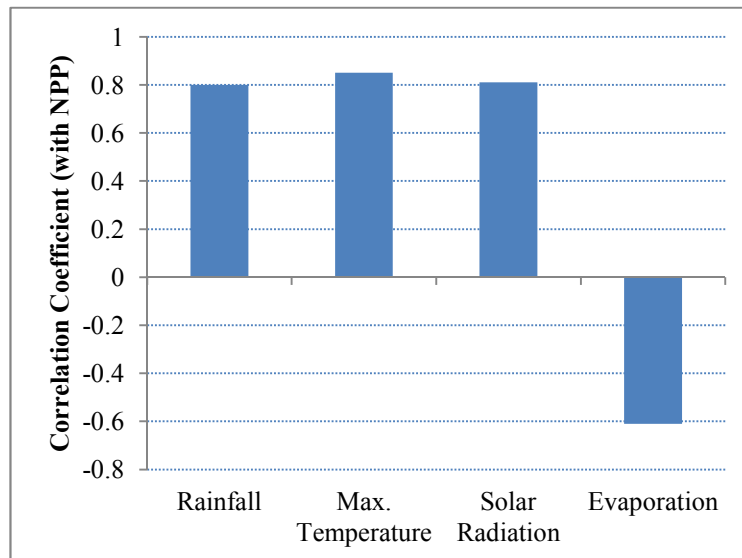


Figure 10: Correlation coefficient of climatic factors with NPP

3.6 NPP and Biophysical Factors

The overall correlation coefficients of all the biophysical factors are shown on Figure 11. The soil water 2 ($r=0.82$), which is an index for water holding capacity of the sub soil (B horizon), was found to have a significant influence on NPP. The tropical forests are made of tall trees with deep roots that extract moisture from deeper layers of the soil. Therefore, indices that indicate water availability in deep layers are likely to give best prediction of NPP (Eamus, 2003). Churkina and Running (2008) indicated that while soil moisture availability may be the most influential variable for setting the upper limit on NPP, multiple environmental factors may influence NPP in nonlinear and perhaps discontinuous ways. Therefore, areas observed to have high soil water holding capacity are likely to avail moisture for longer periods and thus enhance productivity.

Aspect ($r=0.48$), slope ($r=-0.14$), DEM ($r=0.26$) and terrain wetness index ($r=0.37$) were found to have weak to moderate correlation with NPP (Figure 11). Their influence on NPP was significant ($p<0.05$) in spite of the weak correlation. This indicated that their influence may be linked to other factors, and therefore their study requires assessing interaction of a suite of factors. A study by Chen et al. (2007) showed that topography has an effect on NPP, with greater effects observed near 1,350 m above the sea level. Their study further showed that an elevation increase of 100 m above this level reduces the average annual NPP by about 25 g C/m². Also, they discovered that there was a 5% change in NPP as a result of terrain aspect for forests located below 1,900 m as a result of its influence on incident solar radiation. However, in the current study the influence of biophysical factors did not reflect similar results. Their influence could be overshadowed by other factors that affect NPP.

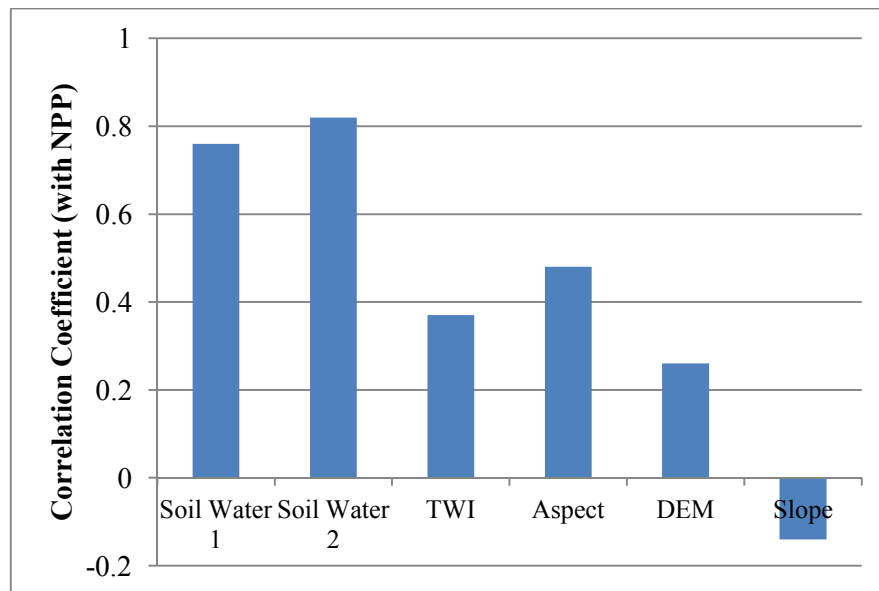


Figure 11: Overall correlation coefficient of Biophysical factors with NPP

4 CONCLUSION

The study has shown that significant drop in rainfall, including a drought event, has a lag effect on tropical rainforests. The effect of a drought episode in a particular year is delayed by one year as shown in this study. This was not tested in other vegetation types. Droughts episodes that occur over a period of more than one year consecutively may depict a different scenario in terms of the lag effect. Therefore studying just the year of drought event (i.e. 2002) may have led to a misleading conclusion that tropical forests are resilient to drought episodes.

The NPP in this region depicted a positive trend, a direct opposite to the recent global scenario where NPP is reported to decline due to frequent drought events. It is an indication that the severity of drought vary from region to region. The observed trend is likely to change given the prediction that frequency of severe drought will increase in the next 50 years in Australia. These predictions indicate future threats to the functions of forest ecosystems given the general negative impact of drought events on NPP.

The interannual variation of NPP in this region seems to be influenced by rainfall amount, maximum temperature, solar radiation and potential evaporation. This was depicted by strong correlation between these factors and NPP. There could be other environmental factors interacting with these factors in the processes driving NPP in the region. However due to time and resource limitation it was not possible to explore directly the influence of nutrients, soil pH, cloud cover, soil type, and plant species.

The index representing the soil water holding capacity in the sub soil was found to have a better positive relationship with NPP than the index representing the top soil. It provided an indication that tropical rainforests benefit more from water available in deeper layers as they generally have deep roots. Other physical factors which include slope, DEM, aspect and terrain wetness index did not exhibit strong correlations with NPP, indicating a weak influence.

Research assessing the effect of drought on tropical forests over longer time periods is essential in the future given the improved monitoring systems and availability of data sets. The challenge with the research establishing effects of past events dating backward to the 1900 is data availability. This future studies will enhance understanding the behaviour of these forest ecosystems in light of repeated episodes of different extreme weather events including drought.

REFERENCES

- Bowman D.J.M.S. (2000). Tropical rainforest progress. *Physical Geography*, 24, 103 – 109.
- Braswell B.H., Schimel D.S., Linder E. & Moore B. (1997). The response of global terrestrial ecosystems to interannual temperature variability. *Science*, 278, 870-872.
- BoM (2011a). Living with drought. Bureau of Meteorology (BoM), Commonwealth of Australian. Viewed 21 May 2011, <<http://www.bom.gov.au/climate/drought/livedrought.shtml>>.
- BoM (2011b). Climate variability and El Nino, Bureau of Meteorology (BoM), Commonwealth of Australian. Viewed 21 May, 2011, <<http://www.bom.gov.au/climate/enso/enlist/index.shtml>>.
- Christopher B.F., James T.R. & Carolyn M.M. (1995). Global net primary production: combining ecology and remote sensing. *Remote Sensing of Environment*, 51, 74–88.
- Cramer W., Kichlighter D.W. & Bondeau A. (1999). Comparing global models of terrestrial Net Primary Productivity (NPP): overview and key results. *Global Change Biology*, 5, 1–15.
- CSIRO. (2001). Climate Change Projections for Australia. CSIRO Atmospheric Research, Aspendale, Victoria.
- DAFF. (2003). Australian forest profile: rainforest, Department of Agriculture, Fisheries and Forestry (DAFF), Bureau of Rural Science, Canberra, Australia. Viewed 23 March 2011, <<http://adl.brs.gov.au/forestsaustralia/facts/water.html>>.
- Department of the Environment and Water Resources (2007). Australia's Native Vegetation: A summary of Australia's Major Vegetation Groups, 2007. Australian Government, Canberra, ACT.
- Falge E., Baldocchi D., Olson R., Anthoni P., et al. (2001). Gap filling strategies for long term energy flux data sets. *Agricultural and Forest Meteorology*, 107, 7, 71-77.
- Fang J.Y., Chen A.P., Peng C.H., Zhao S.Q. & Ci L.J. (2001). Changes in forest biomass carbon storage in China between 1949 and 1998. *Science*, 292, 2320-2322.
- Hashimoto H., Melton F., Ichii K., Milesi C., et al. (2009). Evaluating the impacts of climate and elevated carbon dioxide on tropical rainforests of the western Amazon basin using ecosystem models and satellite data. *Global Change Biology*, 16, 255-271.
- Heinsch F.A., Zhao M., Running S.W., Kimball J.S., et al. (2006). Evaluation of remote sensing based terrestrial productivity from MODIS using tower eddy flux network observations. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1908-1925.
- Houghton R.A. (2007). Balancing the global carbon budget. *Annual Review of Earth and Planetary Sciences*, 35, 313–47.
- Huete A.R., Didan K., Shimabukuro Y.E., Ratana P., et al. (2006). Amazon rainforests green-up with sunlight in dry season. *Geophysical Research Letters*, 33, L06405.
- IPCC (2001). Climate change 2001: the scientific basis. Report of the Intergovernmental Panel on Climate Change, J.T. Houghton, Y. Ding, D.J., et al. (eds). Cambridge University Press, United Kingdom.
- IPCC (2007). Climate change 2007: the physical science basis. Report of the Intergovernmental Panel on Climate Change, S. Solomon, D. Qin, M. Manning, et al. (eds), Cambridge University Press, United.

- Krishnan P., Black T.A., Grant N.J., Barr A.G., et al. (2006). Agricultural and forest meteorology impact of changing soil moisture distribution on net ecosystem productivity of a boreal aspen forest during and following drought. *Agricultural and Forest Meteorology*, 139, 208-223.
- Malhi Y. & Wright J. (2004). Spatial patterns and recent trends in the climate of tropical rainforest regions. *Philosophical Transactions of the Royal Society B*, 359, 311–29.
- McMahon T.A. & Arenas A.D. (1982). *Methods of computation of low stream flow*, UNESCO, Paris.
- Nemani R.R., Keeling C.D., Hashimoto H., Jolly W.M., et al. (2003). Climate-driven increases in global terrestrial Net Primary Production from 1982 to 1999. *Science*, 300, 1560-1563.
- Nepstad D., Lefebvre P., Dasilva U.L., et al. (2004). Amazon drought and its implications for forest flammability and tree growth: a basin wide analysis. *Global Change Biology*, 10, 704-717.
- Norby R.J., DeLucia E.H., Gielen B., et al. (2005). Forest response to elevated CO² is conserved across a broad range of productivity. *Proceedings of the National Academy of Sciences*, 102, 18052–18056.
- Peng D.L., Huang J.F., Cai C.X., et al. (2008). Assessing the response of seasonal variation of net primary productivity to climate using remote sensing data and geographic information system techniques in Xinjiang. *Journal of Integrative Plant Biology*, 50, 1580–1588.
- Saleska S.R., Didan K., Huete A.R. & DaRocha H.R. (2007). Amazon forests green up during 2005 drought. *Science*, 318, 612.
- Samanta A., Ganguly S., Hashimoto H., et al. (2010). Amazon forests did not green-up during the 2005 drought. *Geophysical Research Letters*, 37, L05401.
- Sherwood K. & Craig I. (2011). CO₂ Science. Centre for study of carbon dioxide and global change. Viewed 21 May 2011, <http://www.co2science.org/articles/V13/N36 /EDIT.php>
- Steele B.M., Reddy S.K. & Nemani R.R. (2005). A regression strategy for analysing environmental data generated by spatio-temporal processes. *Ecological Modelling*, 181, 93–108.
- Thomas C.D., Cameron A., Green R.E., et al. (2004). Extinction risk from climate change. *Nature*, 427, 145–148.
- Turner, D. P., Ritts, W.D., Zhao, M., Kurc, S.A., Dunn, A.L., Wofsy, S.C., E.E., Small, and Running, S. W. (2006). Assessing Interannual Variation in MODIS-Based Estimates of Gross Primary Production, *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1899-1907.
- Whetton P. (2003). Projected future climate change for Australia. *Proceedings of the Climate Impacts on Australia's Natural Resources*, Brisbane, Australia, November 2003, 12-14.
- Zhao M. & Running S. (2010). Drought-Induced reduction in global terrestrial net primary production from 2000 through 2009. *Science*, 329, 940-943.

Implementing the Phoenix Fire Spread Model for Operational Use

Gillian Paterson

Land and Fire, Department of Sustainability and Environment, Victoria
L4, 8 Nicholson St, East Melbourne, Victoria 3002
gillian.paterson@dse.vic.gov.au

Derek Chong

University of Melbourne, School of Forest and Ecosystem Science
Burnley Campus, 500 Yarra Boulevard, Richmond, Victoria 3121
derekmoc@unimelb.edu.au

ABSTRACT

In a major fire event accurate and timely information is essential for a range of time critical processes such as community fire warnings, resource decisions and suppression allocation. In the 2009/10 fire season the Department of Sustainability and Environment (DSE) in Victoria trialed using the Phoenix fire characterization model to automatically generate fire spread predictions. In the 2010/11 fire season the trial was extended and additional functionality added for users. PHOENIX models complex fire behaviour that occurs in the Australian landscape using weather, topography, fuel loads, fire history, roading and suppression. PHOENIX has been integrated into DSE FireWeb and Firemap web based applications and produces a 6 hour fire spread prediction for reported ignitions. These predictions support the existing manual process currently employed by DSE for this purpose

This paper will discuss some of the challenges and limitations of the process used to generate the spatial data required by the underlying PHOENIX models, the FireWeb and Firemap user interfaces as well as the steps taken to improve the implementation after the 2009/10 trial.

KEYWORDS: Fire Predictions, Fuel / Landuse Datasets, Data Limitations, Web based services

1 INTRODUCTION

Timely and accurate information is critical in the decision making process during a major bushfire event. There are a large number of information sources and processes that are called upon during an event in emergency bushfire management. In response to the recent catastrophic Victorian fires the Victorian Department of Sustainability and Environment (DSE) Land and Fire bushfire management team has been working towards streamlining decision support systems that utilize these data sources and processes. One of these developments has been the evaluation and subsequent operational implementation of the PHOENIX fire characterization model developed by the University Of Melbourne to support the existing manual fire spread prediction process. Named Phoenix Firemap, it produces a 6 hour fire spread prediction as tabular and map outputs for every fire that is reported to the FireWeb information portal. These outputs are then made available to the approved FireWeb registered users.

PHOENIX models the progression of wildfire in the landscape in response to a range of inputs including fuel, weather and topography. Temporal and spatial accuracy of these inputs is a key factor in allowing the model to provide reliable and accurate representation of fire behaviour originating from a point ignition. Obtaining spatial datasets that could be used as these inputs was

in some instances relatively easy; however others, namely fuel types, proved to be more challenging as it had to be sourced from a range of different datasets to provide state wide coverage.

2 THE PHOENIX MODEL

The PHOENIX model is a “spatially and temporally explicit fire characterization model” that responds to the dynamic nature of the interaction between fire and its environment (Tolhurst, Shields and Chong 2008). It was developed by the University of Melbourne in conjunction with the Bushfire Cooperative Research Centre (CRC) as one component of a bushfire risk management model for southern Australia and continues to be maintained and improved by the Bushfire CRC. PHOENIX requires a number of inputs describing the weather and landscape, and using a grid the model characterizes the fire for each cell, evaluating the changes in conditions – fire, fuel, weather and topographic. The outputs can include the rate of spread, flame height, fireline intensity and ember density falling into the cell (Figure 1). The DSE implementation utilizes a raster output from Phoenix showing flame height and ember density and a vector format for hourly isochrones (fire progression perimeters) from the ignition point.

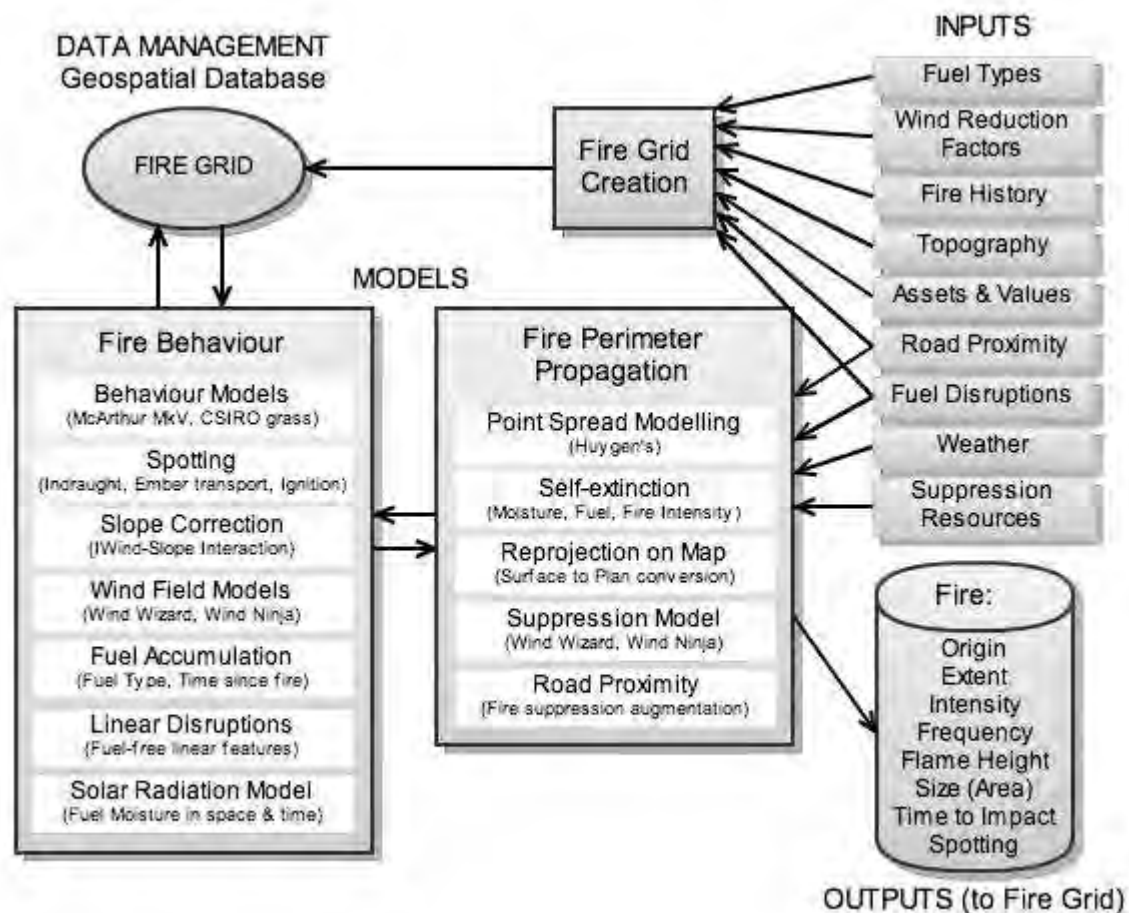


Figure 1. Schematic diagram of PHOENIX showing the inputs, outputs and data storages (Tolhurst, Shields and Chong 2008)

3 OPERATIONAL USE OF THE PHOENIX MODEL IN DSE

A desktop version of the PHOENIX model was trialled during the 2009 bushfires by Kevin Tolhurst in DSE's State Control Centre (SCC). The success and potential of the model prompted it to be developed into an automated process that generated a prediction for every new fire report entered into FireWeb, DSE's Land and Fire information portal. A key to this implementation is the use of gridded forecast weather data sourced from the Bureau of Meteorology; this negated the need for time consuming and error prone manual input of the weather forecast data. This state wide hourly forecast data is provided twice a day at approximately 2km resolution.

The automated process was a black box where all inputs were preconfigured from external sources and no interaction between end users and the process was accommodated. The predictions were of a mixed success, both over and under predicting fire spread. In analysing the results, the main reasons behind the unreliability of the outputs were:

- Incorrect ignition point coordinates, either entered incorrectly, or due to the rounding of using the MGRS coordinates used to report fires
- The reported coordinates placed the fire origin in a no-fuel area indicated by the data e.g. on a road or river and would self extinguish in the model
- Incorrect weather forecast data, e.g. time of change, wind speed or direction, can have dramatic affects on the resulting fire spread behaviour.

(Cummings, 2010)

A high degree of expertise is required to manually predict fires; in addition, the end users of the model outputs needed to understand the limitations of the models underlying PHOENIX. Due to the unreliability of the input data, the trial status of the implementation and the lack of a mechanism to quality control the predictions, access to the predicted outputs was limited to a few select roles, Fire Behaviour Analysts (FBAN), Incident Controller (IC) and the mapping team. The approval process was manually performed by FBANs and the ICs who were considered to have the required training to evaluate prediction accuracy.

An updated user interface to the "black box" system was designed and implemented for the 2010/11 season to address the lack of a systematic verification process for filtering out erroneous predictions and to provide a mechanism of altering some of the model inputs. This new user interface allows the FBANs to view all predictions created by the PHOENIX model, alter the starting time, ignition coordinates or the weather that has been used and re-run the prediction if necessary. Once an FBAN has approved the prediction as a reasonable approximation of fire spread, it is then shown in Firemap, for all to see (Figure 2). If by chance a prediction has been created and a FBAN is not able to review it within 30 minutes of it appearing in the user interface, it is given a status of unvalidated and shown in Firemap as a hashed version of a validated prediction. This still allows information to be disseminated in a timely fashion, but alerts users that greater care needs to be used if considering the information to be used for any decisions. FBANs are still able to review the unvalidated prediction and update it to an approved status if appropriate.

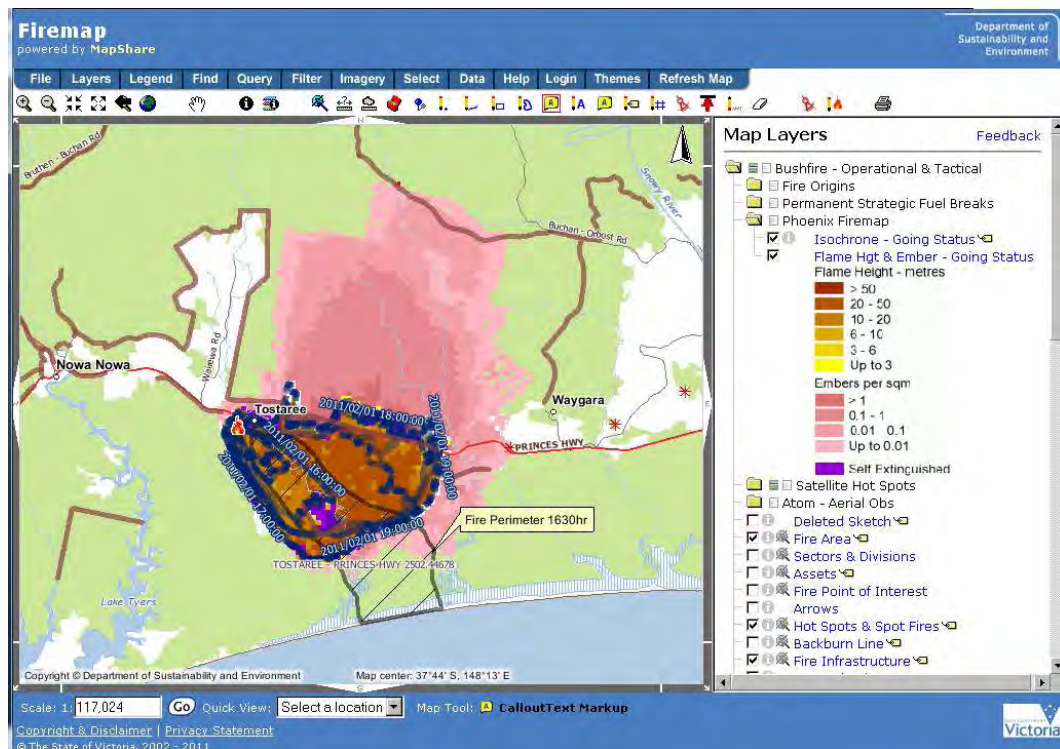


Figure 2. Fire spread prediction output shown in Firemap

The validation process (Figure 3) allows the other roles that do not have specific training in fire behaviour to have access to this information and have a fair degree of certainty that the prediction provides an approximation of the next 6 hours of fire behaviour and can therefore plan and act accordingly. In the case of February 2009 where there were more than 100 fires reported in one day, this would add another layer of information at the state level for resource allocation and hopefully forward planning that could limit devastation to a landscape and its inhabitants.

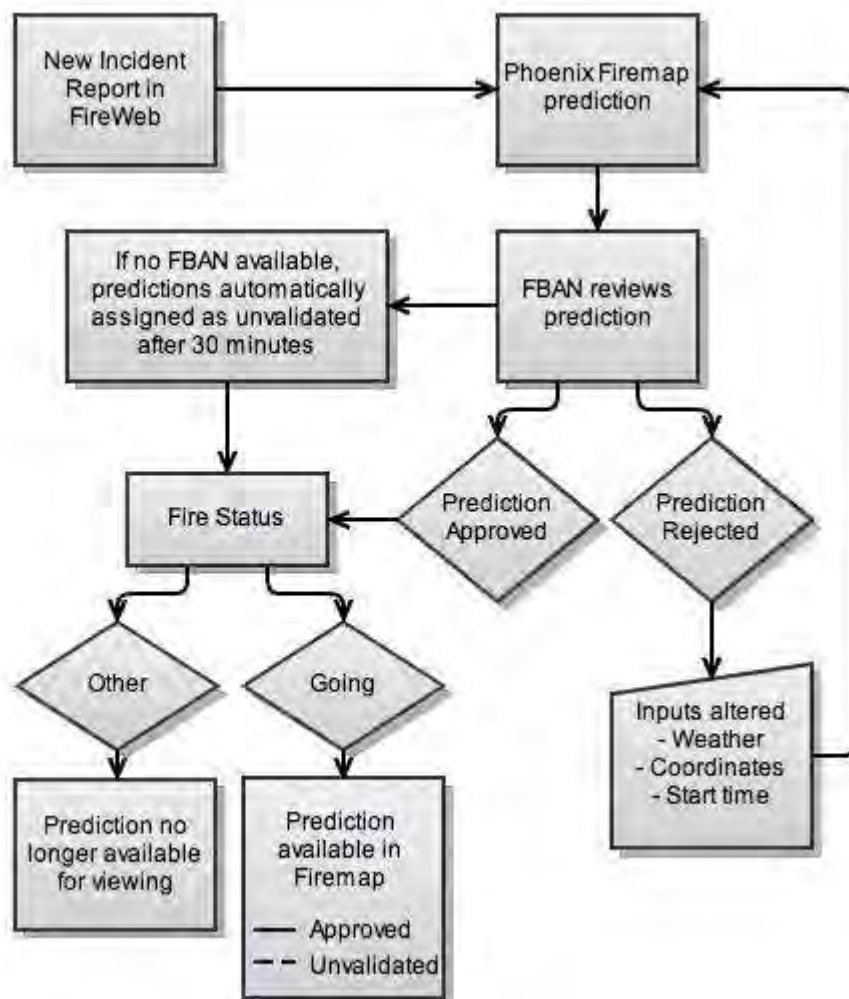


Figure 3. Phoenix Firemap Validation Process

4 INPUTS INTO PHOENIX FIREMAP

Phoenix Firemap relies on a number of different inputs in order to give a reliable fire spread prediction. In Figure 1 nine inputs are listed in the PHOENIX model. Phoenix Firemap only uses 7 of these; Fuel Types, Wind Reduction Factors, Fire History, Topography, Road Proximity, Fuel Disruptions and Weather. The other two, Assets & Values and Suppression Resources, are optional inputs and are currently only used in the desktop version used by the FBANs.

Apart from weather, the input layers are stored as ASCII grids that are updated at various intervals by python scripts. The update frequency is based on the datasets that are used to create the layers and how often they themselves are updated. The creation process for each of the input layers was scripted in order to be able to run updates at regular intervals and with an output that was guaranteed to be consistent. The datasets and information used in the different layers creation was decided upon after much discussion of the suitability of various datasets, and tried to match up the requirements of the model with the information that we had available. For each of the input layers the different contributing datasets are processed into one vector layer with any lookup tables and attribute calculations applied to create one field of the necessary information required for PHOENIX before being converted into an ESRI raster with a grid size of 30m, which is in turn converted to the ASCII grid.

The use of PHOENIX in a desktop version in 2009 meant that some of the decisions on which particular datasets to collate into input grids had already been done. However, some of it needed to be expanded and added to, particularly with the Fuel Type layer which is discussed below.

4.1 Fuel Type

The Fuel Type input layer was the most difficult input grid to create. Unfortunately there wasn't an existing Victoria wide dataset that described different fuel types at a scale that was useable with PHOENIX. The original Fuel Type input layer that was used in the desktop version of PHOENIX was the whole of Victoria assigned to one fuel type, grass, overlaid with a Vegetation dataset which only covered roughly 47% of Victoria. Due to the over simplification of information, a grass type fuel classification covering areas like towns as well as crops, plantations and other different types of vegetation, gross errors occurred in the system's outputs giving some fire spreads a much larger area than would actually be expected and highlighted the need for a more detailed layer. Therefore a number of different datasets were used to fill in various areas and fuel types, each with different accuracies, currency and suitability of use (Table 1). The different positional accuracies of the contributing datasets weren't as much of an issue as the process included a re-sampling into a 30m grid using a centre point methodology resulting in a reduction of the spatial precision of the final layer. The reason that the grid size of 30m was chosen was due to the contributing vegetation dataset being based on a 30m grid, and this was carried through to the rest of the input layers.

Table 1. Contributing datasets of the Fuel Type layer

Dataset	Currency / Update Freq.	Positional Accuracy	Assumptions
Roads	Continual	5m to 20m	
Railways	Bi-Annually	10m to 30m	
Plantations	Annual	Unknown	<ul style="list-style-type: none"> - May not be a complete dataset of all plantations - Polygon boundary only of plantation - Attributes indicate a plantation type of hardwood or softwood - Approximately 50% of features have a plantation type of "undefined" - Year of planting obtained by overlaying another dataset
Vegetation	As required (major dataset update coming late 2011)	10 to 100m	<ul style="list-style-type: none"> - Descriptions of the ecological vegetation community codes were used to assign Fuel Types
Land Use	1996 to 2005	25m to 100m	<ul style="list-style-type: none"> - Descriptions of land use were used to assign Fuel Types
Planning	Weekly	Less than 25m	<ul style="list-style-type: none"> - Used to fill in the gaps around the Melbourne urban area where the Land Use dataset didn't have coverage - Descriptions of the planning zone codes were used to assign Fuel Types
Base	-	-	Final fill for any gaps left from the above.

To assign fuel types to the various dataset and attributes, assumptions were made based on simple descriptions within the datasets and knowledge of different areas of Victoria. The resulting Fuel Type input layer seemed to give the PHOENIX model an accurate representation of fuel load to give reasonable fire spread predictions. A recent review of these fuel type assignments has indicated that a few changes to the land use to fuel type look up table will greatly improve fire spread predictions in certain areas.

Concerns were raised about the contributing land use dataset as this is the largest contributor, by area, to the Fuel Type input layer. The land use dataset was collected using a wide range of methods across a period of around 9 years. Although a good source of information, in some parts of Victoria it is a decade out of date and some land uses were not easily interpreted to

fuel types, hence the decision to include the land use dataset was a hesitant one. It was included based on the fact that it was the best that was available. It is hoped that for future developments of Phoenix Firemap that a better source of fuel classification and/or land use information can be found and that it is regularly updated.

The biggest issue in creating the Fuel Type input layer was the time and memory required to process the huge datasets from polygons to rasters and during the testing phase the process often didn't reach a conclusion. The vegetation dataset alone has more than 370,000 records and the land use dataset has more than 450,000. The original processing time was more than 8 hours and took up more than 80% of the memory capacity. As a result it is set to run overnight and is only updated monthly. Given that the majority of the contributing datasets are not regularly updated, it is perhaps surprising that the Fuel Type input layer is updated as frequently as it is. The main reason that it is updated monthly is due to the inclusion of the planning zones. The planning zone information is also being used to cover an important area, the urban fringe, where change is occurring frequently and it is necessary to keep the fuel information as up to date as possible. This Fuel Type input layer python script is being reviewed and the processing time and resources will hopefully be lessened. Information is also being reviewed to replace the plantations and vegetation datasets.

4.2 Fuel Disruptions

Fuel Disruptions is a buffered layer of roads, rail and hydrology that describes lineal disruptions to fuel loads. Assumptions have been made that the class of each these features is in direct relation to the width of the feature on the ground. Roads and Railways are fairly standard and due to the nature of their use, the classifications are a reliable indication of the on-ground disruption. However, hydrology was not so straight forward as the dataset that was readily available had one classification for a river from beginning to end. This doesn't take into account that at the source a river could be a small creek of less than a metre wide, and by the time it reached the end of its course it could be a major watercourse of several metres width. A dataset had been put together previously that did classify watercourses with a reasonable indicator of on-ground width, and this was the one that was used in creating the Fuel Disruptions layer. The only issue with this hydrology dataset is that it is not part of a database that is regularly maintained. As hydrology is not known to change dramatically at regular intervals this issue was noted but it was decided that the benefits of a better representation of on-ground width were greater than the detriments of it being a static dataset. This may cause issues in the future in areas where flooding has occurred and changed river courses or new water distribution channels dug or existing channels put underground. If the Phoenix Firemap system continues, a method of classifying rivers and applying it automatically annually or bi-annually will need to be researched.

4.3 Fire History

Fire History is information that DSE Land and Fire has been keeping and regularly maintaining for a number of years and our fire history dataset contains information from 1903 to the last fire season, 2011. It is regularly maintained with ad hoc requests for updates from people in the regional offices, and annually on the change over of fire seasons. However, planned burns and bushfire areas are still missing and for some that are included the positional accuracy is quite poor, giving a misrepresentation of what has been previously burnt.

To create the Fire History input layer for PHOENIX, the fire history information is filtered to only contain information from 1975 to the current year, and is then combined with current operational data for bushfires and planned burning, as well as logging history. The process to create this input layer is run nightly and optionally for ad hoc updates, to capture any new information that is drawn into the operational datasets for that day to ensure that any areas burnt act as fuel breaks for the next day's predictions.

4.4 Weather

Traditionally forecast weather data for fire spread prediction was provided by Bureau of Meteorology staff who interpret a range of weather model outputs and current observations to

provide a 'spot forecast' for the fire area at a 3 hour resolution. These spot forecasts have to be formally requested and were faxed to the requester. This manual forecasting system was seen as the major obstacle to an automated spread prediction system. This system suffers from several limitations when applied to events like the recent 2009 Victorian fires:

- Scalability, over 100 fires were reported on one day, all of which would have required a spot forecast
- Large, rapidly moving fires require updated spot fire forecasts as they move across the landscape
- Time required to manually enter the spot forecast data
- Easily made transcription errors, especially in a high pressure emergency context
- Temporal resolutions of 3 hours is too coarse

For these reasons the PHOENIX model was designed to accept gridded weather forecast data from the Bureau of Meteorology. This state wide hourly gridded forecast data is generated from a numerical weather model (ACCESS), is updated twice a day and is available via a subscriber ftp site. The gridded forecast data utilizes the netCDF file format, at approximately 2km resolution with a 7 day outlook.

The two major updates appear at approximately 6am and 6pm and are part of the GFE (Gridded Forecast Editor) approach to producing and updating forecast data. As part of this process forecast grids are manually edited and uploaded during the day to reflect any changes observed during the day.

This manual editing process is intended to negate the need to rerun the ACCESS model during the day when observed conditions significantly deviate from the forecast. This introduced a limitation to the use of the forecast grids as GFE edits were only forward looking which would result in forecast grids that started at the time of the last edit. PHOENIX requires weather data for two hours before the ignition time to capture temperature and RH lag conditions and for the calculation of fuel moisture so any fires starting within 2 hours of an edit would not have the required weather to process.

The other problem observed with the manually edited GFE grids was the spatial and temporal artefacts often introduced by the editing tools used as well as the often variable level of smoothing applied by the different operators. To get around this problem it was decided to only use the 6am update and ignore any subsequent edits, the downside of this approach is that any significant changes to the weather patterns during the day are missed. Discussions with the Bureau of Meteorology are ongoing to improve the temporal and spatial resolution of the forecasts as well as the frequency of model runs.

4.5 Wind, Topography and Roading

Wind Reduction Factors, Topography and Road Proximity are direct interpretations of datasets that are readily available, have reasonable accuracy and currency levels and are therefore the more reliable of the input datasets into Phoenix Firemap. Wind Reduction Factors has a wind model applied to a DEM, Topography is a straight export of a DEM and Road Proximity calculates for each cell the distance to the nearest road.

4.6 Improvements

In order to improve the input layers for PHOENIX two main datasets are required, fuel classification from a dedicated fuel dataset or land use/ vegetation dataset for the whole of Victoria and an accurate fire history layer. As a working dataset, the fire history is constantly improving and will continue to do so, removing gaps and inconsistencies as well as hopefully improving in spatial accuracy as the benefits of accurately capturing previously burnt areas is seen in the wider fire history user community. However, the fuel classification dataset will take collaboration with a number of other departments and a major project to ensure that it is captured at an appropriate scale with the required attributes and is a dataset that is maintained for the future.

5 USER INTERFACE

The verification process of the predictions is enabled through the user interface (Figure 4). Each fire report is shown in a table with all predictions produced listed along with their status. As the user interface can be seen by all FBANs in the state it is necessary to show whether there are predictions being processed (“Running” status) or any predictions “Pending” for approval so that duplicates of predictions being run are limited. The information stored in this table also becomes an archive for the decisions made for each prediction. If weather is changed, FBANs are required to add comments as to the nature of the change. A separate report is written that gives greater explanations if the prediction is of a major fire, however this user interface provides a quick way of capturing the decision making process for future reference if necessary. The documentation produced by the PHOENIX model as well as the spatial data is also stored and archived for future reference.

- TEST 20101215-1017 (15 Dec 2010 11:12)								3 Predictions 0 Running... 1 Pending 0 Approved 0 Rejected	
No.	Start Time	End Time	Status	Author	Status Time	Comments	Status Comments		
3	15 Dec 2010 11:12	15 Dec 2010 11:16	Pending	pr08	15 Dec 2010 11:16	try again again	try again again	+ P P P P P	
2	15 Dec 2010 10:45	15 Dec 2010 10:49	Failed	pr08	15 Dec 2010 10:49	try again	try again	+ P P P P P	
1	15 Dec 2010 10:18	15 Dec 2010 10:22	Approved Unvalidated	system	15 Dec 2010 10:55			+ P P P P P	

Opens Firemap in a new window and zooms to the prediction 


View the prediction details – ignition date, start time, weather as well as a picture of the prediction 

Figure 4. User Interface for Phoenix Firemap allowing the FBANs to interact with the PHOENIX model

6 CASE STUDIES

6.1 Roses Gap - 2009/10

Roses Gap was included as a case study in the Phoenix Fire Simulator – Implementation Review 2009/10 (Cummings, 2010) as it proved that given the correct weather information the automation of the PHOENIX model could provide an accurate fire spread prediction. The differences between the prediction fire boundary (Figure 5) and the final fire boundary (Figure 6) were due to changes in the actual weather on the ground, the wind changes came in earlier than expected meaning that the fire burnt more area to the north, towards the town of Dadswells Bridge.



Figure 5. Predicted fire spread for the Roses Gap fire 2009/10

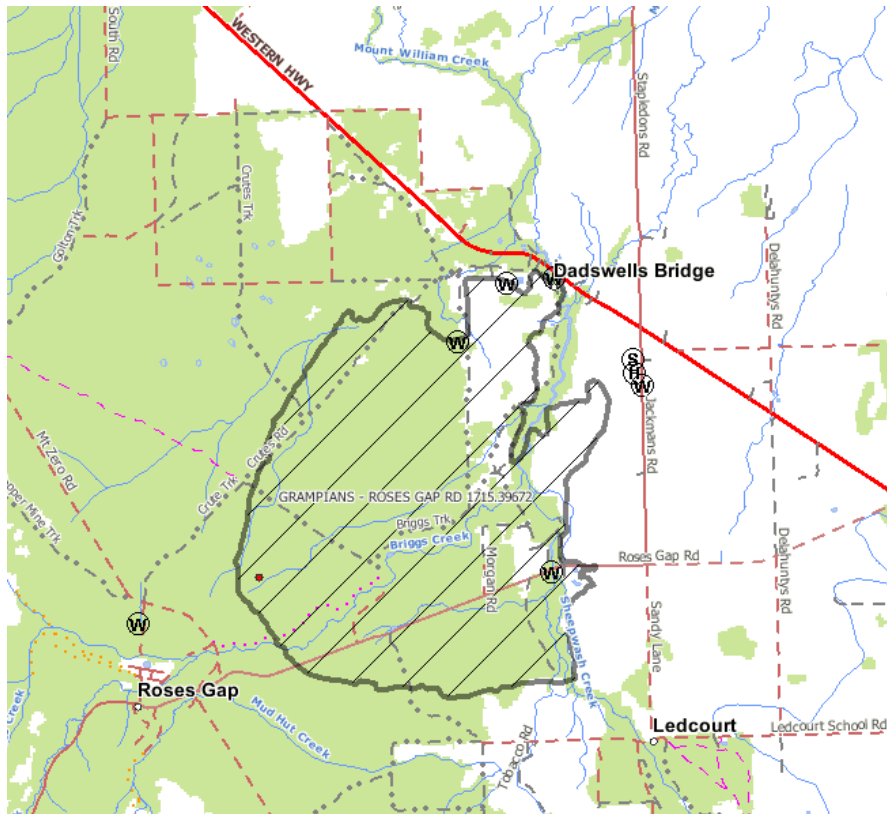


Figure 6. Final fire boundary for the Roses Gap fire 2009/10

The validation from this review of the output being a reliable source of information went towards the justification of creating the Phoenix Firemap user interface and verification process, where changes in the weather were able to be fed back into the PHOENIX model and predictions re-run with updated information.

6.2 Tostaree - 2010/11

The Tostaree fire was an example from the 2010/11 fire season, where the new user interface and verification system was used to approve a prediction from Phoenix Firemap. The first reported coordinate had it on the wrong side of the road. The resulting prediction indicated that the fire self extinguished. The coordinates were corrected, and the prediction process run for a second time. Again, weather played the major role in the differences between the predicted fire spread and actual on ground behaviour (Figure 7). By 16:00 the fire had progressed much further south than predicted and spot fires were a couple of kilometres east of where the impacted area was predicted to be.

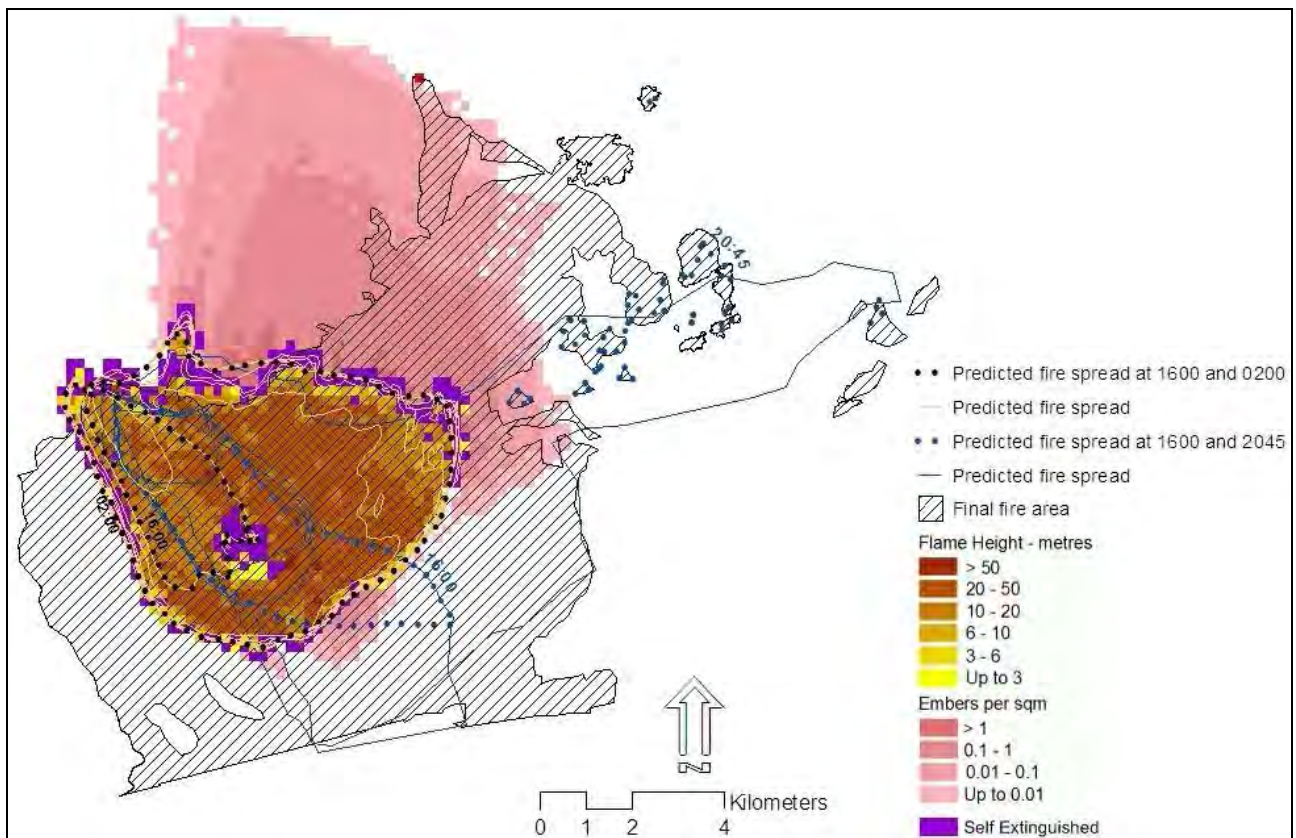


Figure 7. Predicted and actual fire spread for the Tostaree fire 2010/11

7 FUTURE DEVELOPMENT

There are a number of future developments that the Phoenix Firemap project team would like to implement to increase efficiency of the process, yet also take further advantage of the ability to process spatial information to provide products that can be used to inform bushfire management. This year's developments include:

- Ability to parallel process (run) many fires simultaneously
- Summary report for each prediction
- Upgrade to latest PHOENIX model version

Each prediction from Phoenix Firemap, which has been given a status of “approved”, can be relied on as a reasonable representation of the potential impact of a bushfire. Thus automated systems that provide town names for media and public alerts, lists of infrastructure that is likely to be impacted for government and utility companies could be written and a barrage of emails sent out as soon as a prediction has been verified. This would inform operational staff, emergency services and utility companies of potential impacts. It would involve some raster analysis and simple overlay processes, but the wealth of information that could be gained may become indispensable.

Currently the outputs of Phoenix Firemap include the information that has been used to create the prediction – start time, coordinates, weather and some fuel information for the grid cells that have been impacted. There is more information that could be captured from the prediction itself and a summary report has been proposed that will provide for the FBAN and other users a more in depth look at the predicted impact of the fire. Information would include:

- A map of the prediction
- Summary statistics for the whole prediction (eg area burnt, maximum spotting distance)

- Summary statistics for certain time intervals (eg X ha of area burnt at one hour, spot fires)
- Criticality scores of the cells impacted (affects on the environment, economy, infrastructure and people)
- Convective outputs
- Fuel types burnt by area
- Fire history
- Weather
- GFE (Gridded Forecast Editor) aerogram

A python script is being proposed to create this report that will again run some simple spatial analysis to provide further information on the potentially impacted area.

One of the aims of introducing PHOENIX was to obtain timely fire spread predictions. Currently predictions are being processed anywhere between 8 minutes and 30 minutes, longer if there is a queue of predictions waiting to be run. Although the 8 minute predictions are acceptable, the longer ones are not and add a significant amount of time to obtaining information on the potential impact, especially if the information needs to be reviewed and the prediction re-run. The short term goal is to decrease the processing time for creating a prediction by putting in more powerful processors and the long term is to increase capacity for processing predictions concurrently, whether this is “in the cloud” technology or our own farm of servers is yet to be determined.

8 OTHER USES IN THE DEPARTMENT

As well as an operational tool, the PHOENIX model is being used by other parts of Land and Fire to create statewide fuel hazard maps, daily forecast pictures of the potential fire spread and statewide impact scenarios for a two day period. Each application is relying on the same input layers and the development of these applications has contributed to the testing of these layers and the processes that have created them. These applications provide information for forward planning rather than the “need to know now” information required for operational use.

9 CONCLUSION

Introducing the PHOENIX model into the operational user interface of the Phoenix Firemap system has allowed greater access to verified information that can contribute to critical decisions during an emergency operation. However there is improvement to be made in the timeliness in producing the predictions and how the prediction queue is handled by the servers. Improvements can also be made to the inputs to Phoenix Firemap, in particular the Fuel Type layer, however this will not be achieved without great cost and time put into creating a new dedicated dataset. In the mean time, alterations to the processing of the input layers can be made to make them more efficient and able to cope with changes to contributing datasets.

It has been shown that Phoenix Firemap can produce predictions that are an accurate representation of fire spread and will be a valuable source of information for future bushfire management. Further development utilising the spatial information output from Phoenix Firemap to inform of potential impacts will only add to the timely dissemination of information to relevant parties.

REFERENCES

Cummings, F. (2010), Phoenix Fire Simulator – Implementation Review 2009/10, Department of Sustainability of Victoria, Office of Land and Fire.

Tolhurst, K., Shields, B. and Chong, D., (2008), 'Phoenix: development and application of a bushfire risk management tool', The Australian Journal of Emergency Management, Vol.23 No. 4, pp. 47-58

Modelling Habitat Networks Using the Concept of Graph Theory

Fatemeh Poodat, Colin Arrowsmith

School of Mathematical and Geospatial Sciences RMIT University
GPO Box 2476 Melbourne Victoria 3001 AUSTRALIA

Fatemeh.poodat@student.rmit.edu.au, colin.arrowsmith@rmit.edu.au

Elizabeth Farmer

School of Mathematical and Geospatial Sciences RMIT University
GPO Box 2476 Melbourne Victoria 3001 AUSTRALIA

elizabeth.farmer@rmit.edu.au

ABSTRACT

Urbanisation in many major cities results in changes to the ecological integrity of habitats within these areas. Through the process of urbanisation, habitats including native vegetation have been degraded, reduced in size, fragmented, isolated or even lost. This can disrupt the proper functioning of natural areas and inhibits wildlife interaction and gene flow between different habitats. In these areas, the majority of the remaining biological diversity is located in small fragments of native vegetation that have been set aside during development. The movement of species between these habitats in the urban environment is subject to their functional connectivity. Lack of suitable connectivity in networks would cause populations to be isolated by surrounding urban areas and threaten their long-term viability. It is important for planners to recognise what the networks of individual species are so that future development can take these into account. GIS modelling of habitat networks provides input data for connectivity analysis, which can be used for effective conservation planning, landscape restoration and ecological impact assessment studies.

Graph theory is an effective way of modelling habitats and the ecological interactions that occur between them. The objective of this study is to represent habitat networks for three threatened faunal species: *Varanus varius* (Lace Monitor), *Sminthopsis crassicaudata* (Fat-tailed Dunnart), and *Litoria raniformis* (Growling Grass Frog) based on the concept of graph theory. This paper presents the application of two algorithms for the development of a node layer in order to simulate habitat networks for each of the target species in Metropolitan Melbourne: the resilient centroid model, and the maximum distance to edge algorithm.

The results for the three species differ in terms of the number of nodes and their distribution pattern which were determined by the shape and orientation of the patches used to them. The more linear or perforated a polygon was the more centroids are derived. Deriving the centroids using the shrinking polygon algorithm detects the central cores of polygons. We suggest that the resilient centroid model is suitable for perforated compact patches rather than elongated patches. Further research will show how the links in the network will be represented to reveal the connectivity of the network.

KEYWORDS: Habitat networks, graph theory, resilient centroid model, Polygon shrinking algorithm, GIS.

1 INTRODUCTION

The process of urbanisation alters not only the landscape structure, but also its ecological functionality. Urbanisation refers to densification and outward spread of the built environment (Forman, 2008). As a result of this process, habitat patches in the landscape have been altered critically (Forman and Godron, 1986). Such modifications in habitat patches degrade ecological

values, shrink habitat size, and modify their shape (Alberti, 2004; Young and Jarvis, 2001). Fragmentation or complete eradication can also occur (Alberti, 2004; Young and Jarvis, 2001). This highlights the need for conservation planning for urban landscapes that incorporate habitat mapping.

One consequence of change in the structure of urban landscape is the disruption of ecological functioning. This process radically influences both natural habitat and wildlife communities through the conversion of the extensive areas into fragmented and isolated islands (Hough, 2004). For instance, movement of faunal species between neighbouring habitats can be severely affected (Wheater, 1999). Most species need to disperse in order to provide their daily or lifetime needs such as food, energy, refuge or breeding. They might even disperse to adapt themselves to new climate conditions. The dispersal of species in this modified landscape allows for the recolonisation of areas that have suffered from local declines in population or extinctions, and also reduces “the negative effects of inbreeding” (Forman et al., 2003). Small fauna populations with the ability to move between these discrete islands of habitat form a metapopulation (Merriam, 1990) that is more resilient to urbanisation.

In the urban environment, habitats must be functionally connected so that species can readily move between them. Functional connectivity refers to “the degree to which landscape facilitates or impedes movement among resource patches” (Taylor et al., 1993). The combination of the habitats and the availability of a functional connection between them is known as a habitat networks (Forman and Godron, 1986).

Functional connectivity is species-specific (Crooks and Sanjayan, 2006). Species may have different habitat requirements as well as different dispersal abilities. For example, the particular distance between two habitat patches which seem to be connected for a large mammal may not be connected for a small amphibian with less dispersal ability and more sensitivity to the type of landscape structure. Each species therefore has a unique habitat network which may be completely separate from the network of other species, or may be adapted to or used in conjunction with the networks of other species.

Lack of suitable connectivity in the networks is likely to cause extinction of populations in isolated habitats especially when they are under threat from the surrounding urban context. In this context, functional connectivity determines how the network facilitates the movement of species. Therefore urban planning authorities can benefit from a deeper understanding about habitat networks of species so that developments do not have negative impacts on species dispersal.

There is also the potential to determine which elements (eg. habitat patches or functional corridors) of the networks are more critical in the dispersal of the species. GIS modelling of habitat networks provides input data for connectivity analysis, which can be used for effective conservation planning, landscape restoration and ecological impact assessment studies.

In this paper, we aim to simulate the habitat networks of these threatened species using the concept of graph theory which is part of PhD research in progress at RMIT University, Australia.

2 APPROACHES TO MODELLING HABITAT NETWORKS

Two approaches applied for modelling species habitat networks are the gradient and binary modelling approaches. Binary modelling refers to classification of landscape elements into two broad clusters of habitat and non-habitat, whereas in the gradient approach each part in the landscape will have a value of probability of being habitat for a certain species.

One of the applications of gradient approach is least-cost modelling. This approach has been widely applied to modelling habitats and their functional links (Adriaensen et al., 2003; Drielsma et al., 2007). In least-cost modelling, a resistance (friction) layer is developed so that for each cell of raster data, landscape resistance to the dispersal of a species is determined. Considering habitats as the least resistant part of the landscape, the least-cost path between two habitats is derived. Cost refers to the accumulative amount of resistance value between two habitats and the least-cost path is the path with the lowest accumulative resistance value. The result of applying this approach will be a habitat network and the least-cost path for dispersal of species. The representation of the network in this approach is dependent on the arrangement of lowest resistant

cells, which sit beside each other forming a homogenous group of cells interpreted as habitat. The group can also form a string of cells which is assumed as a corridor in the network.

In the binary approach, a landscape element is considered either as habitat or non-habitat and there is no probability which shows the importance of habitat patch. One application of this approach is modelling based on graph theory concept. Using graph theory concept, habitat and their interactions can be represented as sets of nodes and links in GIS environment.

2.1 Graph Theory

Graph theory is a mathematical concept based on a finite set of nodes and links. This concept was initially introduced by Harary in 1969 (Harary, 1969). Graph theory has been applied in a variety of disciplines including ecology (eg. Keitt et al., 1997). Bunn et al. (2000) demonstrated the first application of graph theory in simulating connectivity in habitat networks which resulted in suitable scenarios for conservation biology. It has become commonly accepted that graph theory is an effective way of modelling habitats and ecological interactions among them (García-Feced et al., 2011; Urban and Keitt, 2001; Zetterberg et al., 2010). Graph based modelling is a rapid tool in conservation assessment (Urban and Keitt, 2001) and is not data demanding (Saura and Rubio, 2010).

Habitat networks have been simulated on the basis of graph theory concept in several studies. For instance Bun et al. (2000) modelled the habitat networks for two fauna species which share the same habitat but have different dispersal ability in the Coastal Plain of North Carolina. Saura et al. (2011) simulated the habitat networks for forest habitats over the European continent. The forests and the links between them were represented in two time spots in order to reveal how the connectivity in the forest network changed over time.

There are various ways to represent a network using graph theory. Zetterberg et al. (2010) grouped them in three types of patch-corridor, patch-link and nodes-links. In the patch-corridor type of modelling, habitat patches and corridors will be represented in their actual size, shape and location. Patch-link modelling also represents habitats in the same extent whereas the possibility of species dispersal is represented as links between the habitats. In the node-link representation, habitat patches are represented as nodes and functional connectivity between the nodes is represented as links. The choice of each modelling type for analysis is dependent on the objective of the study, the condition of the input data and the availability of ecological information (eg. dispersal information of the focal species) to support the model.

The objective of this study is to represent habitat networks for three threatened fauna species, *Varanus varius* (Lace Monitor), *Sminthopsis crassicaudata* (Fat-tailed Dunnart), and *Litoria raniformis* (Growling Grass Frog). The input data of this study are binary sets of habitat maps for these target species, and the aim is to apply the graph theory concept to model the species networks. In this model, nodes and links represent the habitat patches and the possibility of dispersal between them respectively. This paper presents the application of two algorithms for developing the node layer in order to simulate habitat networks for each of the target species in Metropolitan Melbourne.

3 METHODS

3.1 Study Area

Melbourne, the capital of the state of Victoria, is one of the major cities in Australia. Metropolitan Melbourne is located on Port Philip Bay on the southern coast of south-eastern Australia (Figure 1). Melbourne has a temperate climate with distinctly dry and warm summers (Bureau of Meteorology, 2010). More than 500 native vertebrate fauna and invertebrate species and more than 17,000 vascular native plants have been recorded within the study area since 1990 (Victorian environmental Assessment Council, 2010).



Figure 1: The location of the Metropolitan Melbourne study area in the State of Victoria, Australia

Since the British settlements in the 1880s, the Melbourne environment has been greatly modified. Native vegetation and habitats in Melbourne are fragmented, degraded, and lost in some parts. This has resulted in a large proportion of Victoria's threatened flora and fauna species occurring in Melbourne. About 300 plant species, vertebrate fauna and invertebrate fauna species are listed as threatened in Australia and Victoria (Department of Sustainability and Environment, 2007a, 2009). Three of the threatened species in Metropolitan Melbourne were selected for network modelling in this study.

3.2 Target Species

The criteria for this study were that the native threatened species chosen for habitat modelling have frequently occurred in the study area and have overlaps in their spatial distribution. The fauna species are considered as threatened species which are affected by loss of connectivity in their network. In fact, the long-term viability of these species is dependent on how connected their habitat networks are. Due to the special conservation statuses of these species, there is valuable information and expert assistance available regarding their dispersal and spatial ecology. They were selected from different types of fauna species with different ability and behaviour in dispersal in order to reveal different aspects of movement by fauna species and to do a comprehensive analysis. The target species are Lace Monitor (*Varanus varius*), Fat-tailed Dunnart (*Sminthopsis crassicaudata*), and Growling Grass Frog (*Litoria raniformis*).

The Lace monitor (*Varanus varius*) is a very large carnivorous, predatory lizard characterized by an elongated muscular body (King and Green, 1999). The Lace Monitor is the second largest terrestrial carnivore in south eastern Australia (Weavers, 1993). Monitors occur in well-timbered areas from dry woodlands to cool temperate southern forests (Wilson and Swan, 2003). The minimum viable size for the Monitor's habitat is considered to be 10 Hectares (Guarino, 2002) and minimum suitable width of habitat patch is set to 1000 metres (Nick Clemann, 2011, Pers. Comm, July). According to the advisory list of threatened vertebrate fauna of the Victorian government, the conservation status of the Lace Monitor is considered vulnerable, that is considered to be facing a high risk of extinction in the wild (Department of Sustainability and Environment, 2007a).

The Growling Grass Frog (*Litoria raniformis*) is a large threatened frog distributed in south eastern Australia. The Growling Grass Frog generally inhabits permanent and ephemeral wetlands and their immediate surrounds (Heard, 2010). According to Heard (2011, Pers. Comm, July) the minimum viable size for Growling grass frog habitat is between 1000 and 4000 square metres. Also, the minimum width of the habitat patch is 100 metres (Geoff Heard, 2011, Pers. Comm, July). The range and numbers of the frog have declined since the 1980s with many populations disappearing (Department of Sustainability and Environment, 2007b). *Litoria raniformis* is now listed as threatened according to the Victorian Flora and Fauna Guarantee Act 1988 (FFG Act) and is vulnerable to extinction according to the Environment Protection and Biodiversity Conservation Act 1999 (EPBC Act).

The Fat-tailed Dunnart (*Sminthopsis crassicaudata*) is a small mammal which inhabits only lowland areas in the western half of Victoria (Menkhorst, 1995). The Fat-tailed Dunnart predominantly lives in open grass land habitats (Morton, 1976) such as tussock and hummock grassland, giber plain, saltbush and bluebush plains and rough pasture (Menkhorst and Kenight, 2001). The threshold for the minimum size of habitat patches of Dunnart is considered 20 hectares and for the minimum width of habitat patches is set to 500 metres (Peter Menkhorst, 2011, Pers. Comm, July). *Sminthopsis crassicaudata* is under the advisory list of threatened vertebrate fauna in Victoria (Department of Sustainability and Environment, 2007a).

3.3 Input Data

In Metropolitan Melbourne, potential habitat maps for a number of rare or threatened fauna species were modelled. This potential habitat modelling was part of a study for conservation zoning in Metropolitan Melbourne (Gordon et al, 2009). The habitat maps were produced using an expert based process in which the potential habitats for each specie were derived from a main question, 'what land uses, vegetation type and/or wetland types are defined by the available spatial data, would never compromise habitat for this species?' (Gordon et al., 2009). The result of the process was a set of binary maps with two broad classes – habitat and non-habitat – with 20*20 metres pixel size. Deriving the potential maps with only a simple habitat requirement model and the above mentioned question, resulted in some cases where the potential map overlaps with a certain land use (eg. industrial) (figure 2). Despite this generalised view to the reality of presence of species, the habitat maps for the three target species was prepared from Gordon et al. (2009). Theses maps were found to be suitable for the objective of this study and application of graph theory because they are binary set of maps which are developed from the same process by the same expert. Therefore, there is a hope that less uncertainty was achieved by this comparative analysis.



Figure 2: Input habitat maps overlaps with urbanised lands such as roads, industrial and residential uses. Habitat patches of target species are shown in pink.

There are a number of modifications made to adapt the input data to the aim of this analysis. The modifications are: setting thresholds for the minimum suitable size and width of habitat patches, and assigning separate values for each habitat patch. A local expert in each of the target species was consulted in order to identify suitable size and width of habitats.

As the suitability of a habitat is a function of different landscape elements, the resulting habitat maps differ in terms of fractal dimensions (e.g. shape and size, and perforation of habitat patches). For example, the habitat patch of the Dunnart has a large and compact shape with lots of perforation which is less complicated in terms of fractal dimension, while the habitat patches of the lace monitor has an elongated linear shape which would be difficult to serve as long-term habitat for the species. A habitat map for the Growling Grass Frog was developed by buffering around the water bodies and wetlands which resulted in soft and smooth borders of long patches with a good width to support the species.

Patches in each habitat map had the same value. Applying the nearest neighbour algorithm, a patch refers to a group of raster cells connected in any 8 directions. By applying the Region

Group tool of Esri's *ArcGIS* (version 9.3.1), which is based on the nearest neighbour algorithm, separate values were assigned to each patch. The resulting patches are complex in their fractal dimensions and it is difficult to differentiate whether a certain patch serves ecologically as a habitat or as a functional corridor. Next, having determined the minimum viable size of a habitat patch for each target species, patches that were smaller than the threshold were filtered and eliminated for the derivation of centroid analysis. However, these small patches may serve as corridors or stepping-stones for the target species and therefore will be considered in future studies of developing the links between nodes. Because of complex fractal dimensions of habitat patches, especially for the case of the Lace Monitor, it was decided to consider the minimum suitable width after achieving the node layer for each species. The next part will explain how to derive the centroid of habitat patches by applying the shrinking polygon algorithm.

3.4 Centroids

In a graph, network habitat patches can either be represented as polygons or nodes. Due to the excessive border of habitat patches in the input data, depicting the habitats by nodes demonstrates the core of each patch and simplifies representation of the shapes. In fact, because of the complex fractal dimensions of patches, it is important that the nodes of the habitat patches be representative of dense areas inside each patch. This helps with the patches which vary in width and clarifies which parts of the patch are dense enough to support the nesting of species.

A variety of algorithms has been used to derive the centroid of a polygon in GIS. Centre of gravity, geographic centre, and centroid are the terms utilised for the centre of a polygon. Deakin and Bird (2002) name and define seven centroids: the moment centroid, area centroid, average centroid, the minimum bounding rectangle centroid, minimum distance centroid, negative buffer centroid, and the circle centroid (Deakin and Bird, 2002). Deakin and Bird (2002, page 161) suggest that moment centroid is "the best measure to the centre of complex polygons". Moment centroid or centre of gravity is the centre of mass for a polygon which is bounded by x- and y- axis. Farmer et al. (in-press) argue that these algorithms, when applied to highly perforated complex polygons, do not reveal the reality inside habitat patches. Their argument is that the centroid may fall outside the patch or near the edge of the polygon and not represent the polygon mass which is the most homogenous part of the patch (Farmer et al., in-press).

Farmer et al. (in-press) suggest that a polygon reduction algorithm is the most suitable method for deriving the centroid of a complex shape polygon such as a vegetation patch. This algorithm is based on topological thinning of a patch where the iterations may result in several small sub-patches for each patch. Rather than a single centroid, these reduction algorithms generate multiple centroids that are representative of the topological cores inside patches. This study aimed to apply the resilient centroid model (Farmer et al., in-press), based on the reduction polygon algorithm, to derive the centroid of habitat patches for three selected target species.

The term 'resilient centroid' refers to the centroid or centroids which survive after thinning iteration of the external and internal edge of a polygon. The resilient centroid model is a *Python* integrated scripting environment using Esri's *ArcGIS*. The resilient centroid model is applied to the raster data and works by shrinking a polygon from its external and internal edges. Shrinking iterations may result in smaller fragmented sub-patches. Sub-patches are subsequently eroded to achieve the pixel most resilient to the shrinkage. The resilient centroid is guaranteed to be located in the area of greatest polygon density or greatest homogeneity and has the maximum distance from the internal and external polygon edges (Farmer et al., in-press). Farmer et al. (in-press) proved that the resilient centroid occurs inside the patch and is not influenced by the complexity of polygon boundaries.

Two factors must be set in the resilient centroid algorithm: erosion increment and a polygon size filter threshold. Erosion increment refers to the number of cells eroded in each stage. The polygon size filter is applied after an iteration, to test the size of the resulting sub-patches so that only those larger than the threshold are considered for the subsequent process. In this study, the erosion increment is set to 1 so that only one pixel would be eroded in each stage, and no threshold is considered for the polygon size filter to get the most nodes out of the patches. We argue that because of the different sizes and fractal dimensions of habitat patches, setting a limit on the size of a polygon in the resilient centroid model will mean that the model is unable to give an

accurate picture of the largest homogenous cores inside the habitat patches. The reason is that the most resilient centroid in a set of linear habitat patches may not be as important as the most resilient centroid in a large and compact-shaped habitat. Therefore, it is preferable to determine all the resilient centroids from a single polygon and then, by considering their minimum distance to edge, choose the centroids which are representative of the larger cores (sub-polygons).

In order to assign the value of each polygon to the resulting centroids, the 'extract value to the point' tool of *ArcGIS* (version 9.3.1) was applied. Using the Maximum Distance to Edge algorithm (Farmer et al., in-press) and *Python* scripting, a Euclidean raster for each habitat map was prepared in which each pixel has a value of minimum distance to the edge of the specific patch. By assigning the Euclidean raster value to the resulting centroids, each centroid has an additional value which shows its minimum distance to the border of the polygon. Centroids for each habitat map were filtered based on equation (1) below so that centroids with value (minimum distance to edge) less than w were omitted from further analysis.

$$w = \frac{s}{2} \quad (1)$$

Where s refers to the suitable width for each target species, w refers to the width threshold we considered for each species and respective centroid layer. This equation is based on the assumption that centroids are at the same distance from both side edges. Through this process, every effort was made to differentiate between the corridors and habitat patches and depict the networks in an ecologically realistic way.

3 RESULTS AND DISCUSSION

Many centroids were derived for each habitat using the resilient centroid model. More centroids resulted from linear patch habitats rather than from compact shapes. This was specifically the case for the Growling Grass Frog and the Lace Monitor habitats where there were many linear patches. This was not the case for the Fat-tailed Dunnart where the habitats were compact. Also the more varied the direction and orientation of the linear patches, the more fragmented the sub-patches and hence more centroids. Figure 3 shows some examples of derived centroids for the Lace Monitor compared to the Growling Grass Frog habitats.

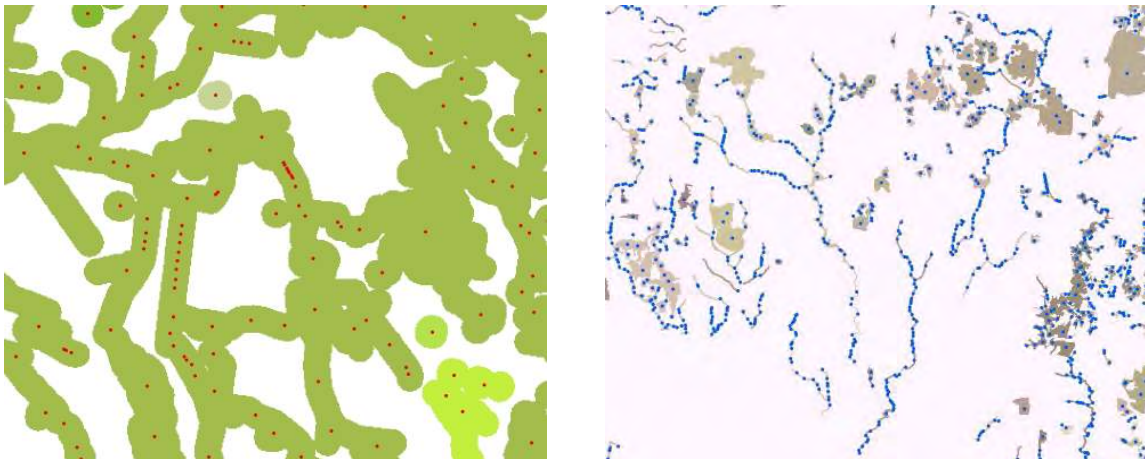


Figure 3: (a) Derived centroids for the Growling Grass Frog (b) Derived centroids for the Lace Monitor: Many centroids resulted for linear shape patches.

After filtering the centroids based on a suitable width using equation (1), the number of nodes dramatically declined for the Lace Monitor whereas almost no change occurred in the number of centroids for Growling Grass Frog habitats (Figure 4).

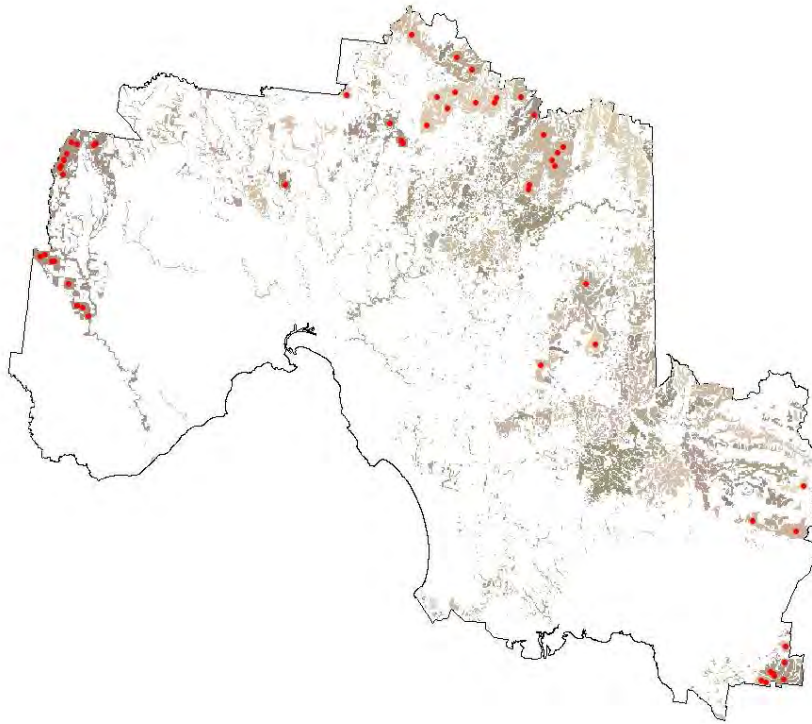


Figure 4 (a): The nodes layer for the Lace Monitor: The number of resultant centroids declined after applying the width threshold

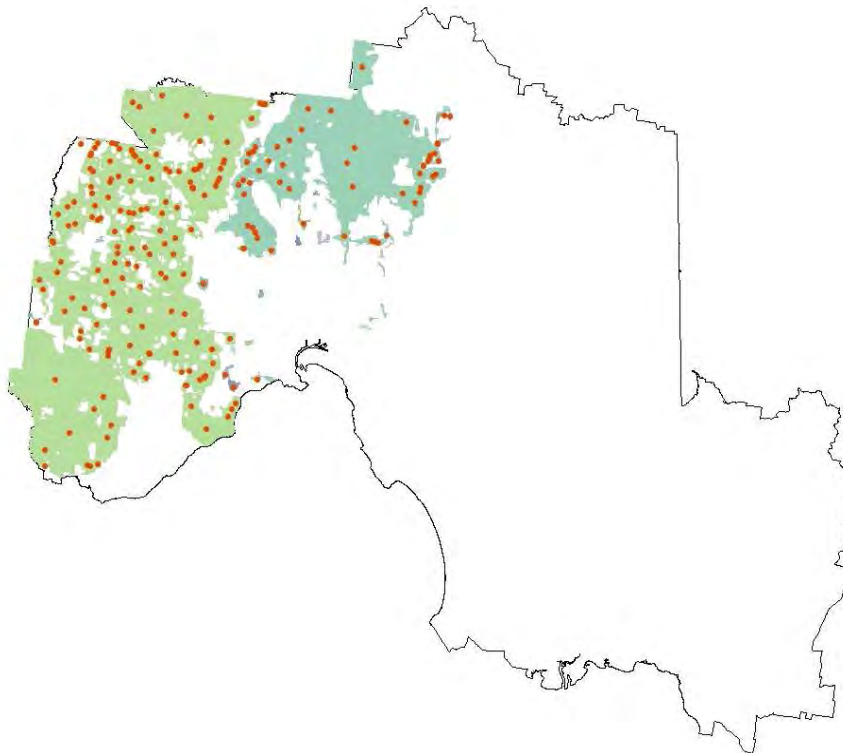


Figure 4 (b): The nodes layer for the Fat-tailed Dunnart: The number of centroids declined moderately by applying the width threshold

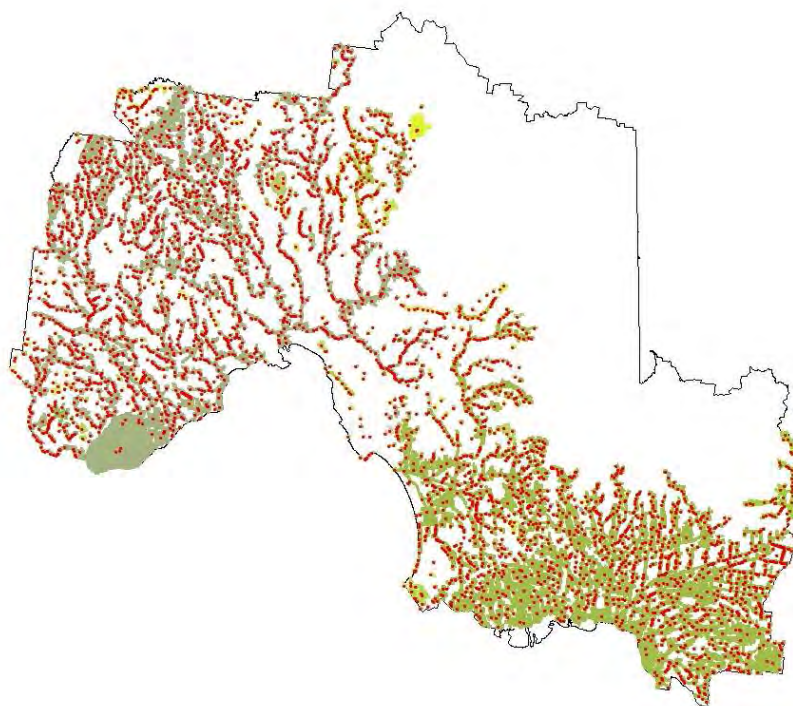


Figure 4 (c): The node layer for the Growling Grass Frog: There were almost no changes in number of centroids by applying the width threshold

Table 1 shows the number of centroids resulting after applying each algorithm. The distribution pattern of the final centroids (nodes) is different for each species. The nodes layer of the Growling Grass Frog show linear shape patches with nodes also aligning well, while nodes for the Fat-Tailed Dunnart habitat were not of a regular shape and they distribute non-linearly.

Table 1 Number of centroids after applying each algorithm

Target Species	Number of centroids after applying the Resilient Centroid Model	Number of centroids after applying the Max Distance Algorithm
Lace Monitor	7040	56
Growling Grass Frog	4535	4533
Fat-tailed Dunnart	389	221

Deriving the centroid with a shrinking polygon algorithm provides the opportunity to detect the habitat cores in the input data. Shape and orientation of habitat patches influences the functioning of the algorithm. We suggest that this algorithm is more suitable for highly perforated compact shape patches (eg. the Fat-tailed Dunnart habitat map) rather than long linear shape patches (eg. the Growling Grass Frog habitat map).

The way this algorithm eliminates cells at the edge of polygon results in fragmentation of patches into separate output patches. This assists with detecting cores in patches. These patches and the newly-formed sub-patches give a hierarchy to the landscape (Forman, 1995). The other advantage of utilising the resilient centroid model is that the resultant centroids are located inside the patches. This facilitates assigning the value of a centroid to a patch. Other methods such as geographic centroid can result in centroids outside a patch.

The results for the three species differ in terms of the number of nodes and their distribution pattern. The number of nodes is subject to a suitable width threshold (equation 1). In addition the shape and orientation of patches influences the number of centroids, where the more linear or perforated a patch is, the more centroids will be derived. The distribution of the nodes is also influenced by the shape and perforation of a patch.

5 CONCLUSION

Graph theory is an effective way to represent habitat networks into two classifications: habitat or non-habitat. In this study we aimed to simulate habitat networks for three threatened species using the concept of graph theory. We present the application of two algorithms for developing the node layer of habitat network for the target species. The two-step process of the resilient centroid model and maximum distance to edge algorithm, is an effective way for deriving centroids of complex habitat polygons, so that each polygon is represented by a number of centroids with a maximum distance from the polygon edges. These centroids represent the core regions of homogeneity within a patch that are suitable for long-term stay of species. The results of this study showed that the resilient centroid model is more suitable for perforated compact shape patches rather than elongated thin patches. Further to this research, there is a plan to derive the links between nodes in the habitat networks which will represent the functional connectivity of the networks.

ACKNOWLEDGEMENT

The Department of Sustainability and Environment of the Victorian Government supplied the input data for this study, and we especially thank Dr. Matt White who prepared the input habitat maps.

We are very grateful to Dr. Nick Clemann, Dr. Geoff William Heard, Dr. Peter Menkhorst, who provided valuable input on the dispersal ability and habitat requirement of the target species.

And finally, we sincerely thank Dr. Ascelin Gordon and Mr. Hossein Pourali for their valuable comments throughout this research.

REFERENCES

- Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villalba, S., Gulink, H., Matthysen, E. (2003). The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning*, 64, 233-247.
- Alberti, M. (2004). Ecological resilience in urban ecosystems: Linking urban patterns to human and ecological functions. *Urban Ecosystems*, 7, 241-265.
- Bunn, A. G., Urban, D. L., Keitt, T. H. (2000). Landscape connectivity: A conservation application of graph theory. *Journal of Environmental Management*, 59, 265-278.
- Bureau of Meteorology (2010). Australian climate zones.
- Crooks, K. R., Sanjayan, M. (2006). Connectivity conservation: maintaining connections for nature. In Crooks, K. R., Sanjayan, M. (Eds) *Connectivity Conservation*, Cambridge University Press, New York, 1-19.
- Deakin, R. E., Bird, S. C. (2002). The Centroid? Where would you like it to be? *Cartography*, 31, 2, 153-167.

- Department of Sustainability and Environment (2007a). advisory list of threatened vertebrate fauna in Victoria. Melbourne.
- Department of Sustainability and Environment (2007b). Growling Grass Frog. www.dse.vic.gov.au.
- Department of Sustainability and Environment (2009). advisory list of threatened invertebrate fauna in Victoria. Melbourne.
- Drielsma, M., Manion, G., Ferrier, S. (2007). The spatial links tool: Automated mapping of habitat linkages in variegated landscapes. *Ecological Modelling*, 200, 403-411.
- Farmer, E., Jones, S. D., Deakin, R. E. (in-press). Derivations of a resilient polygon centroid for natural resource management applications. *International Journal of Geographical Information Science*.
- Forman, R. T. T. (1995). *Land Mosaics; The ecology of landscapes and regions*. Cambridge University Press.
- Forman, R. T. T. (2008). *Urban Regions, Ecology and Planning Beyond the City*. Cambridge University Press, New York.
- Forman, R. T. T., Godron, M. (1986). *Landscape Ecology*. John Wiley & Sons.
- Forman, R. T. T., Sperling, D., Bissonette, J. A., Clevenger, A. P., Cutshall, C. D., Dale, V. H., Fahrig, L., France, R., Goldman, C. R., Heanue, K., Jones, J. A., Swanson, F. J., Turrentine, T., Winter, T. C. (2003). *Road Ecology; Science and Solutions*. Island Press, Washington.
- García-Feced, C., Saura, S., Elena-Rosselló, R. (2011). Improving landscape connectivity in forest districts: A two-stage process for prioritizing agricultural patches for reforestation. *Frontiers in Ecology and Management*, 261, 154-161.
- Gordon, A., Simondson, D., White, M., Moilanen, A., Bkessy, S. A. (2009). Integrating conservation planning and landuse planning in urban landscapes. *Landscape and Urban Planning*, 91, 183-194.
- Guarino, F. (2002). Spatial ecology of a large canivorous lizard, *Varanus varius* (Squamata: Varanidae). *Journal of Zoology*, London, 258, 449-457.
- Harary, F. (1969). *Graph Theory*. Reading, Mass : Addison-Wesley.
- Heard, G. W. (2010). *Pattern, process and the conservation of threatened amphibian metapopulations*. La Trobe University, Bundoora.

Hough, M. (2004). *Cities & Natural Process; A Basis for Sustainability* (Second edition ed.). Routledge, London.

Keitt, T. H., Urban, D. L., Milne, B. T. (1997). Detecting critical scales in fragmented landscapes. *Conservation Ecology*, 1, 1, 4.

King, D., Green, B. (1999). *Monitor: the biology of varanid lizards*. NSW University press, Sydney.

Menkhorst, P. W. (Ed.). (1995). *Mammals of Victoria*. OXFORD UNIVERSITY PRESS AUSTRALIA, Melbourne.

Menkhorst, P. W., Kenight, F. (2001). *A Field Guide to the Mammals of Australia*. Oxford University Press-Australia and New Zealand.

Merriam, G. (1990). Ecological Processes in time and Space of Farmland Mosaics. In Zonneveld, I. S., Forman, R. T. T. (Eds) *Changing Landscapes: An Ecological Perspective*, Springer-Verlag, New York.

Morton, S. (1976). *Ecological and Physiological Studies of Sminthopsis Crassicaudata*. Melbourne University, Melbourne.

Saura, S., Estreguil, C., Mouton, C., Rodríguez-Freire, M. (2011). Network analysis to assess landscape connectivity trends: Application to European forests (1990-2000). *Ecological Indicators*, 11, 407-416.

Saura, S., Rubio, L. (2010). A common currency for different ways in which patches and links can contribute to habitat availability and connectivity in the landscape. *ecography*, 33, 523-537.

Taylor, P. D., Fahrig, L., Helein, K., Merriam, G. (1993). Connectivity is a vital element of landscape structure. *Oikos*, 68, 3, 571-573.

Urban, D., Keitt, T. (2001). Landscape connectivity: a graph-theoretic perspective. *Ecology*, 82, 5, 1205-1218.

Victorian environmental Assessment Council (2010). *Metropolitan Melbourne Investigation, Discussion Paper*. Melbourne: Victorian Environmental Assessment Council.

Weavers, B. W. (1993). Home Range of Male Lace Monitors, *Varanus varius* (Reptilia: Varanidae), in South-eastern Australia. *Wildlife Research*, 20, 303-313.

Wheater, P. C. (1999). *Urban Habitats*. Routledge, London.

Wilson, S., Swan, G. (2003). Princeton Field Guides, Reptiles of Australia. Princeton University Press and Oxford.

Young, H., Jarvis, P. J. (2001). Measuring urban habitat fragmentation: an example from Black Country, UK. *Landscape Ecology*, 16, 643-658.

Zetterberg, A., Mörtberg, U. M., Balfors, B. (2010). Making graph theory operational for landscape ecological assessments, planning, and design. *Landscape and Urban Planning*, 95, 181-191.

Towards a Stream Hierarchy and Classification System for Victorian Waterways

Phillip Delaney, Joanne Poon, Zaffar Sadiq Mohamed-Ghouse, Derek Goodin, Kate Wilson
Sinclair Knight Merz
452 Flinders St, Melbourne, Victoria, Australia
pdelaney@globalskm.com

Darren McKinty
Department of Sustainability and Environment
570 Bourke St, Melbourne, Victoria, Australia
Darren.McKinty@dse.vic.gov.au

ABSTRACT

The Australian State of Victoria's stream network dataset, contained in the Vicmap Hydro product, is a fundamental data set that is employed to support the project planning and management of infrastructure and natural resources. It is used widely within the private sector, as well as within state, regional and local government agencies. Currently, each watercourse has a hierarchy attribute that can be identified as high, medium or low. While this classification is suitable for its original intention of cartographic representation on 1:25000 hardcopy maps, the lack of a more detailed hierarchy prevents users from utilising the full potential of the dataset. Further limiting the potential uses of the stream layer, the dataset does not contain attributes suitable to hydrology classification and analysis. As such, the Victorian Department of Sustainability and Environment commissioned a pilot study to identify ways to develop the stream network dataset to better meet the needs of the users from a cartographic and analysis perspective.

A pilot study tested a new method to classify hierarchy on the basis of stream metrics for the Mitchell River catchment in Victoria. This classification is defined by the total stream length upstream of each stream segment, instead of the current method of classifying by name and waterway type. This allows a systematic classification of streams based on hydrological importance. In addition, the pilot catchment has been classified using a Pfafstetter hydrological coding system to increase the ease with which the dataset can be analysed by spatial and hydrology users. The Pfafstetter system allows users with limited spatial knowledge or technical capability to understand and analyse a stream network using logical numerical attribute selection alone, without need for Digital Elevation Model (DEM) or stream-arc analysis.

The stream metric classification system has been trialled as a number of options, each of which provides a more relevant and detailed classification for both mapping and spatial analysis of streams than the current attribute. The Pfafstetter system was also successfully applied and tested in the pilot area, and will provide a suitable system for use in hydrology applications. However, due to the complex nature of this coding system, it is not suitable for inclusion in the regularly maintained stream network dataset. Instead, it is recommended that this dataset be created as an annual or biennial dataset, separate to the Vicmap Hydro product. The successful outcome of this pilot study provides a justification for applying hydrology coding and stream hierarchy attributes across the state-wide Victorian stream network.

KEYWORDS: Stream Hierarchy, Pfafstetter, Hydrology, Stream Network, Vicmap

1 INTRODUCTION

The State of Victoria's hydrographic feature dataset is stored in a database known as Vicmap Hydro. Within the database, elements such as watercourses, waterbodies, water infrastructure and navigation features are spatially represented. The stream network, including rivers, streams, channels, drains and connectors across waterbodies and underground structures, is contained within Vicmap Hydro Watercourse layer. This spatial layer is a fundamental data set and is used widely by state, regional and local government agencies, emergency services and within the private sector to support project planning, infrastructure management and natural resource management. The data is derived from 1:25000 or better mapping and has a spatial accuracy of 25 m or better (Department of Sustainability and Environment, 2010).

Currently, each watercourse is attributed with a feature type code describing water course type either as watercourse river, watercourse channel or connector; and a hierarchy attribute identified as either high, medium or low. While this classification is suitable for its original intention of cartographic representation for visualisation on 1:25000 hardcopy maps, the lack of a more detailed hierarchy prevents users from utilising the full potential of the dataset. Hydrologists lack the confidence to use this hierarchy in a meaningful way. This uncertainty exists for a number of reasons, including a lack of documentation or metadata of how the classification is defined; a lack of consistency in attribution; and there is uncertainty of what high, medium and low refers to (i.e. a high what?).

A more relevant classification method is needed to allow the data set to be used more in analysis in several application industries. Example industries that would benefit from a more detailed classification are hydrologists developing hydrologic models for similar streams to predict stream behaviour and understand catchment characteristics; biologists and geographers making comparisons between streams with similar characteristics; geologists studying the amount of sediment in a given area; or as a communication tool for natural resource managers.

Additionally, more flexibility and consistency will be a benefit for cartography, as three stream classes can limit mapping applications. Without updating the hierarchy within Vicmap Hydro Watercourses, the Department of Sustainability and Environment (DSE) is at risk of not meeting users' needs. There is a potential threat that agencies will develop their own datasets to meet their own operational requirements which may become disassociated with the fundamental data source.

National watercourse datasets in Australia have similar requirements. Geoscience Australia's Geodata watercourse dataset uses a hierarchy of major and minor and, similar to Vicmap Hydro Watercourses, makes a distinction between perennial and non-perennial stream flow (Geoscience Australia, 2011). The Bureau of Meteorology (BoM) Geofabric data similarly classifies stream information as major and minor, and perennial and non-perennial, but includes attribution for stream type and flow direction and uses a modified Pfafstetter stream ordering for stream classification in the surface catchments layer (Bureau of Meteorology 2010).

This study aims to assess a number of different stream classification methods and trial a selection of these methods over a study area within Victoria. The classifications will be assessed for usability and detail, as well as the alignment to national watercourse datasets. Recommendations will be made regarding ways to improve the hierarchy classification for Vicmap Hydro, as well as any additional developments that will improve the Vicmap Hydro product.

2 QUANTITATIVE STREAM CLASSIFICATION METHODS

The current stream classification method used in the Vicmap Hydro watercourse layer is a categorical classification system whereby the class of a stream segment is defined by the status of the name field. Streams with a name are categorised as 'High' or 'Medium' regardless of their length. More significant streams without a name were given the class 'Low'. As outlined above, a more scientific approach is required to classify these datasets to allow for better analysis, representation and use. As such, categorical classification systems were excluded early on in this study in favour of a more robust, repeatable quantitative system. The following sections detail the quantitative stream classification systems considered for this study.

2.1 Strahler

Strahler uses a number hierarchy to classify streams by assigning streams an order commencing at first order at the headwater fingertip tributaries and increasing incrementally towards the main stream (Strahler, 1957; Figure 1). The assumption for the Strahler classification method is that rivers which share a similar physiography are of the same order and show similar traits in drainage-basin area, mean channel network length, mean slope, and water discharge, and when two similar order streams join the mean discharge is doubled. Thus, two first order streams at a junction form a second order stream, two second order streams at a junction form a third order stream etc. If two streams of different order combine, there is no increase in hierarchy.

The limitation with this classification is that in practice it is difficult to make generalisations across catchments as two basins classified as fifth order may possess different characteristics as they are made up of different numbers of lower order streams and thus will differ in size. For example, a large number of minor tributaries may form a junction with a larger order but there would be no increase in stream order despite accumulating additional discharge.

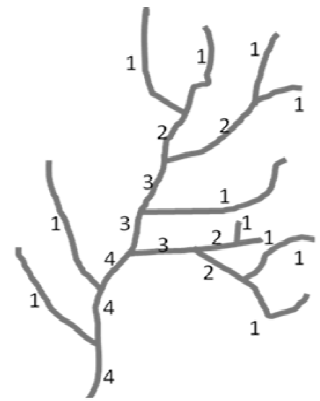


Figure 1: Stream ordering based on Strahler classification

2.2 Shreve

Shreve's classification is based on the assumption that in the absence of strong geologic controls, a natural river network is topologically random. Shreve overcomes the limitations of Strahler's approach by acknowledging the contribution of lower order tributaries and sums upstream orders (Shreve, 1966).

The advantage of this classification is that Shreve's order gives information about the magnitude of discharge in the stream. The larger the stream order, the more discharge through the system (Figure 2). However, for large systems the stream order may become a large number and becomes computationally difficult. Updates to the stream order may also affect the whole classification system requiring rework.

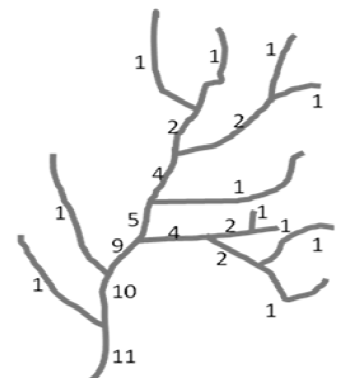


Figure 2: Stream ordering based on Shreve classification

2.3 Pfafstetter

Pfafstetter method uses numbers to code drainage networks within a basin (Figure 3). Using this hierarchy, drainage network topology within a basin and the size of the surface drainage area can be used to divide basins into small sub-catchments. Streams at level 1 indicate continental scale watersheds and increasing levels indicate finer sub-catchments (Pfafstetter, 1989). The advantage of the Pfafstetter approach is that topologic information can be inferred from the class. Therefore, a stream can be identified as upstream or downstream of a point easily due to characteristics of the numbering system. However, a limitation is that the numbering system is not user intuitive for cartographers.

Pfafstetter classification has been applied to Bureau of Meteorology's Geofabric dataset. The intended purpose of the Geofabric's surface catchment layer is for applications of hydrologic modelling and water resource mapping (Bureau of Meteorology, 2010). Developed using input from the national 9" hydrologically enforced DEM and its associated flow direction grid, the data is hierarchically ordered at the first level by 12 drainage divisions defined by topography, 191 catchment units approximating the

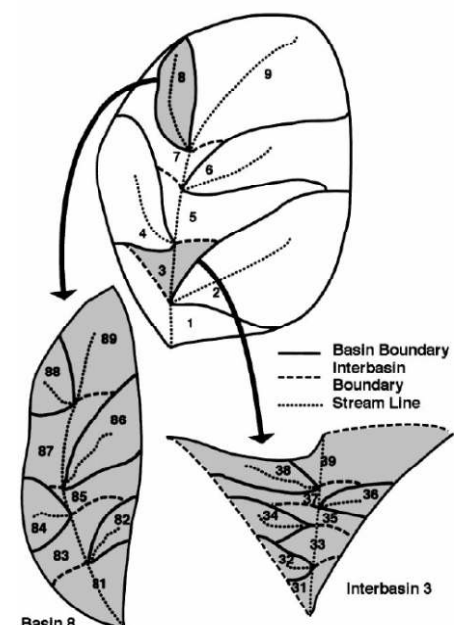


Figure 3: Stream network classified using Pfafstetter codification (Verdin and Verdin, 1999)

Australian Water Resources Council (AWRC) river basins at the second level, and the third through to the fifteenth level are ordered based on a modified version of Pfafstetter coding (Verdin and Verdin, 1999).

The stream classification is represented as a numerical integer for level 3 through to level 15 and stored in a database table. Application of a modified Pfafstetter classification within the Geofabric is beneficial to users as the codes infer topological information inherent with the numerical coding; however the limitation with the implementation of Pfafstetter within the current Geofabric dataset is that it does not adequately take into consideration a condition where more than two upstream or downstream segments connect at a junction, or node. Thus, the classification code may not represent streams with certainty where there are three or more stream segments meeting at a single junction, for example in a braided river system such as the Murray River.

2.4 Stream and Catchment References for Environmental Data (SACRED)

The Victorian Department of Water Resources (now Office of Water) in 1987 developed a reference system, SACRED, in response to a need for greater consistency and standardisation when managing Victoria's natural resources. SACRED numbers are assigned according to basin, stream, reach and type. Basins are delineated based on Australian Water Resources Council (AWRC) basins. Within each basin, streams are ordered first by the primary terminal stream (e.g. streams that flow into the sea, across a basin or state border etc), followed by terminal streams from west to east (Figure 4). Named streams are higher in the hierarchy than unnamed streams. Tributaries contributing into terminal streams are then ordered from downstream to upstream. This ordering system is also applied to minor tributaries. Appended to SACRED numbers is a feature code which attributes the type of stream (e.g. deltaic stream, reservoir, anabranh, drain, channel).

The advantage of SACRED numbering is the flexibility to incorporate suffixes to add more information about a stream, for example describing a position along a stream can be achieved by adding a reach number to the SACRED reference. It is easy to update or include new streams with the SACRED numbering system as decimal places can be used to extend the reference. However, implementing this stream classification can be subjective and therefore susceptible to inconsistencies.

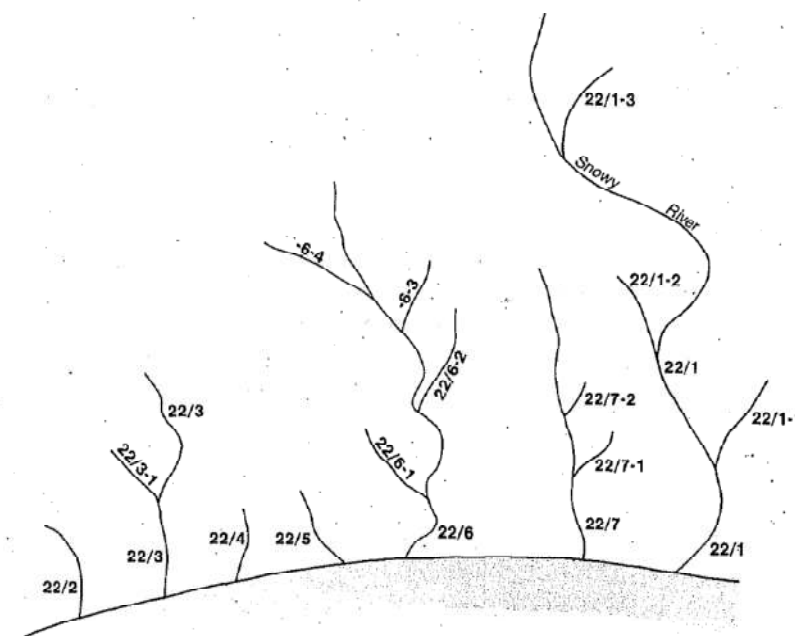


Figure 4 Stream classification using SACRED numbering (Wilson and Nason, 1991).

2.5 Stream Metric Hierarchy

The final classification method considered was developed as part of this study as a simple method to classify stream hierarchy as a metric based on the total length of stream upstream of any given stream segment. As this is not a method that has previously been studied, this paper tests various methods of total upstream length metrics, and provides a recommendation as to the suitability of this method for stream hierarchy classification. The details on this classification method can be found in Sections 4 and 5.

2.6 Design approach

An assessment of the classification systems identified that the stream metric hierarchy and Pfafstetter provided a simple, repeatable classification and as well as direct alignment with the Australian National Geofabric coding system. However, improving the hierarchy classification is not the only improvement that can be made to the Vicmap Hydro product. Applying a complex hydrological code such as Pfafstetter would not only inform the hierarchy, but also allows for complex hydrology analysis to be performed on streams without the need of a DEM or specialist software packages. Further discussion of this can be found in Section 5.

The classification techniques selected need to meet the following criteria:

- A strong scientific basis and supporting case studies
- Stream network functionality (can trace upstream and downstream from any point)
- Alignment with the Australian Bureau of Meteorology Geofabric direction
- Unique numbering to which project specific information could be joined
- Rules based approach to increase the ease of maintenance

To adequately assess the selected classification techniques, a number of data fields needed to be appended to the existing Vicmap Hydro attributes to test their applicability to form the basis of a stream hierarchy, to gain an understanding of feasibility to apply across the state, and to scope the ongoing data maintenance requirements.

The Pfafstetter stream numbering system was the best way to logically code a stream network to achieve the above outcomes. This is because while the code is not directly applicable to hierarchy, the creation of the code informs the hierarchy decisions, and it also allows for effective hydrology classification and analysis. Many of the characteristics required to determine the Pfafstetter system can be directly applied to a stream metric based hierarchy. As such, the stream metric based hierarchy attributes were also applied and tested.

The core data fields developed are:

- Pfafstetter stream hierarchy
- Metrics and ratios of the reach stream length, total upstream length and total length of stream per river basin;
- Attribution of the next downstream stream reach

3 METHODOLOGY

3.1 Pilot Area

A subset of the state-wide watercourse layer was used to test the effectiveness of each classification. The Mitchell River catchment was chosen as a catchment hydrologically complex enough to validate the methods while being an appropriate size for repeated processing. The catchment covers just less than 5000km² of area in East Gippsland.

3.2 Hydrology Classification

The Pfafstetter method uses numbers to code drainage networks within a basin. The division of basins into catchments and sub-catchments is derived on the largest changes in flow accumulation at stream junctions. This allows the basin to be divided into 9 segments. The even numbers 2, 4, 6 and 8 are given to the 4 largest diversions in flow accumulation, representing the 4

largest stream branches, and the odd numbers 1, 3, 5, 7 and 9 are given to the inter-basin areas along the main stream in a basin.

In each of these 9 stream segments, the numbering process is repeated, with each segment being divided into 9 segments, the even numbers being assigned to the 4 largest branches from the current segment, the odd numbers assigned to the inter-basin areas within the segment. This process is repeated iteratively until all stream segments within a network have been coded. These numbers are combined sequentially, and assigned to the stream segments as a unique ID.

In the Vicmap Hydro watercourse layer, there is relationship between flow accumulation and cumulative stream length such that the total stream length upstream of a stream segment is directly related to flow accumulation. As such, total upstream length (TUL) can be used as a substitute for flow accumulation in determining the division of segments by major step points. In this case, the four largest streams contributing to the major stream in a basin will be determined by identifying the largest changes in total upstream length associated with each stream segment.

Figure 5 shows the division of an example catchment in steps to assign the Pfafstetter code. 5a) shows a basin stream network. In Figure 5b) the four main upstream length steps have been identified, with the basins assigned the even numbers and the inter-basins assigned the odd numbers. 5c) shows this process being re-applied on basin 4. Figure 5d) shows the process of applying the code to an inter-basin area.

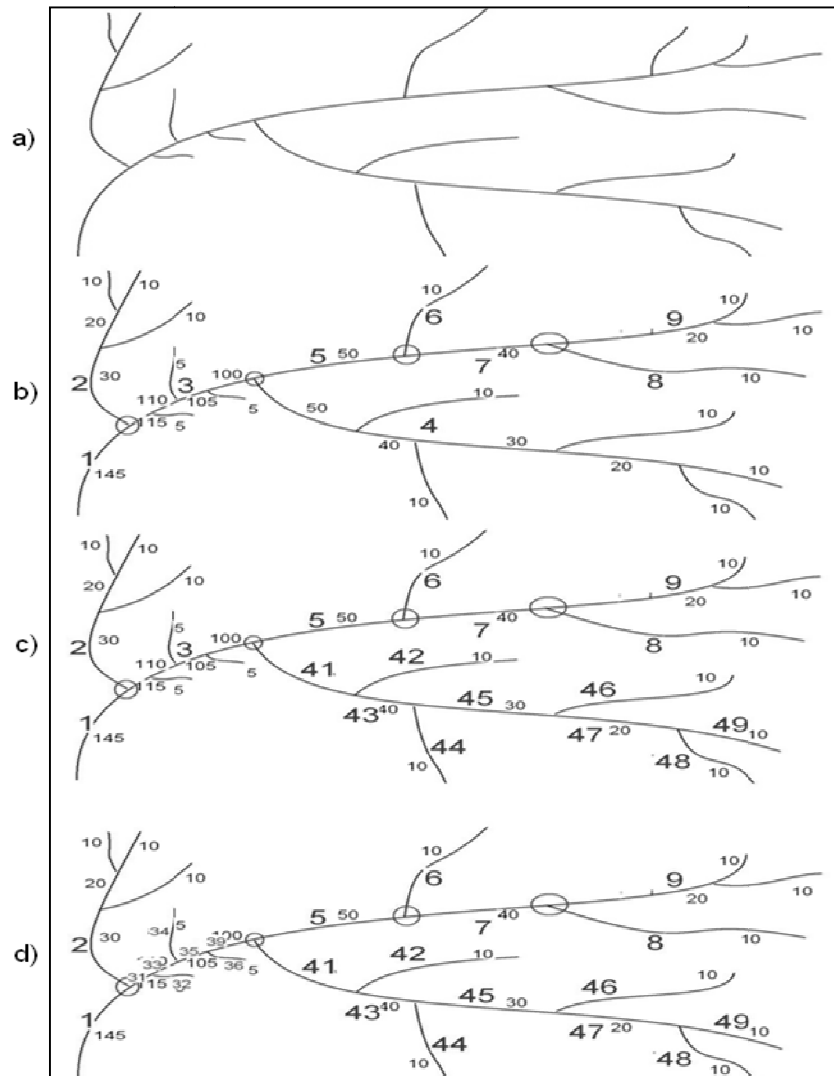


Figure 5: Division of streams using the Pfafstetter system

The existing Vicmap Hydro network contains directional streams, and the downstream Identifier (ID) has been assigned to each segment. This attribute also allows the identification of all segments between two step points, allowing the Pfafstetter code to be assigned.

To select all segments upstream of any segment, the Pfafstetter code needs to be examined to identify the last occurrence of an even number segment Pfafstetter code. The even number indicated the last point in the network where the current section deviated from the main stream. To select all the catchments upstream of a segment, an upper selection limit is defined by adding 1 to the last even number and replacing the digits after this to '0'. For example, the upper selection limit for the code 12345 is 12350 and for 123111 is 130000. This simple code selection allows analysis of a stream network using only the segment attributes.

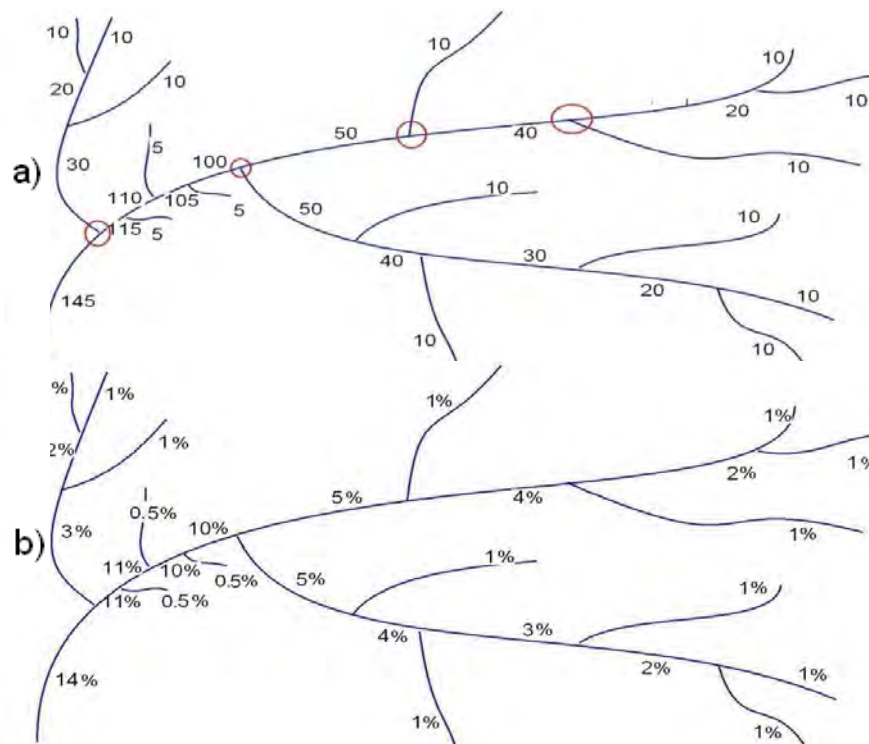
The Pfafstetter code works on a network of streams. However, within a basin there will be many networks of streams. These networks vary greatly in size, from major streams like the Mitchell River, to minor streams or creeks flowing for 3 or 4 kilometres before reaching an outlet. As such, an additional unique numerical attribute has been created breaking the Mitchell Basin into contiguous networks. The Pfafstetter code applied to the each stream segment is unique within the network only. The code will be unique within the basin by combining it with the network code. An example of this would be 123 – 456789 where 123 is the network code and 456789 is the Pfafstetter code.

3.3 Hierarchy

Pfafstetter addresses the needs of a hydrologically coded system for analysis, but does not allow for easy decoding for cartographic purposes due to the multiple stream networks that exist within any basin, as discussed in Section 5.4.1. As such, the metrics of a stream network are required to classify the importance of a stream within a basin. These stream metrics are based on the total upstream length (TUL) calculated during the Pfafstetter process, as well as the total length of stream in the basin.

The new concept introduced in this paper, TUL, was applied as a ratio of the total length of streams in the basin to provide a numerical classification of hierarchy. The TUL was used by comparing the upstream length with the total length of streams within the basin. This will produce a numeric ratio assigned to each stream segment. These ratios can then be classified to determine hierarchy.

Figure 6 shows the relationship between the stream length in a network to the total length of stream in a network. In this small example, the ratio classes could be: above 5% - 'High', 2-5% - 'Medium', below 1% - 'Low'.



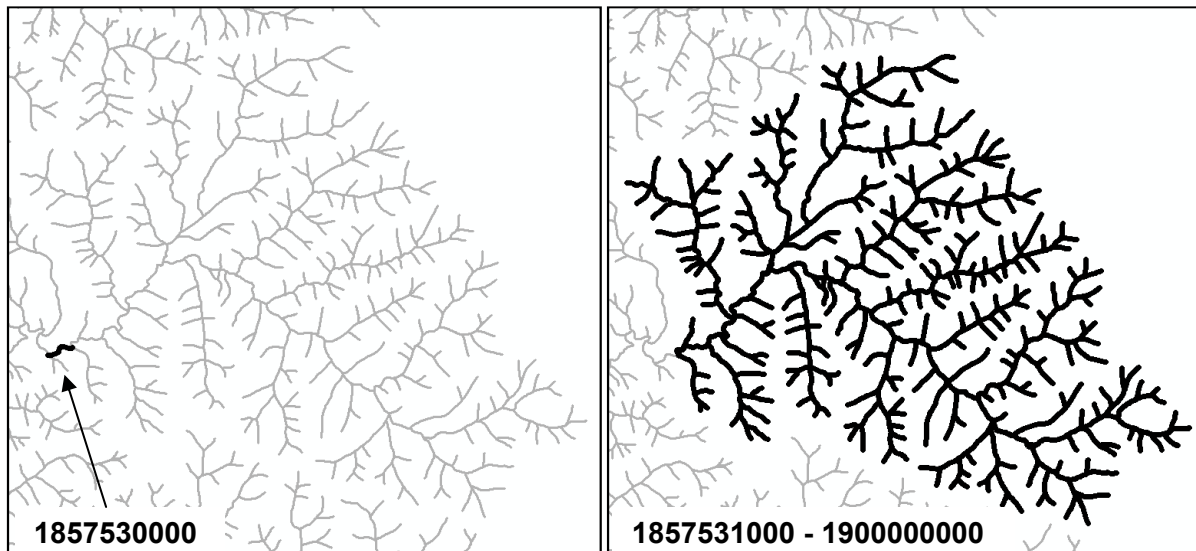


Figure 7: An example upstream selection using the Pfafstetter code

4.2 Hierarchy Results

Using the TUL calculated for the Pfafstetter Code attribution, three metrics were calculated to examine the viability of using this metric for stream classification. The different metric calculations are:

- The ratio of TUL and total network stream length
- The product of TUL and total network stream length
- The ratio of TUL and total basin stream length

Each of these classification methods provided a numerical result that can be classified. Each method produced slightly different stream classification, and a discussion as the most appropriate method can be found in Section 5.

This study does not provide a recommendation as to how to classify the results of the metric calculations in the Vicmap Hydro product. However, an example classification as applied in the Mitchell Basin can be seen in Figure 8. This figure shows a comparison of the existing 3 class hierarchy against an 8 class hierarchy using the TUL basin length ratio. This comparison makes it clear that additional classes based from stream metrics make for a clearer and more detailed representation of a stream network than the existing 3 classes.

To provide an example of classification limits, the limits of the 8 classes that were defined in Figure 8 have been documented in Table 1. These are rounded versions of a natural break (Jenks) classification within ESRI ArcMap software. This study does not provide recommendations as to appropriate classification limits, this would be addressed in a future study.

CLASS	TUL-BASIN RATIO LIMITS*
1	100.00
2	8.23
3	5.11
4	2.36
5	1.21
6	0.49
7	0.18
8	< 0.05

Table 1: Example Stream Metric Classes

VICMAP STREAM NETWORK HIERARCHY CLASSIFICATION

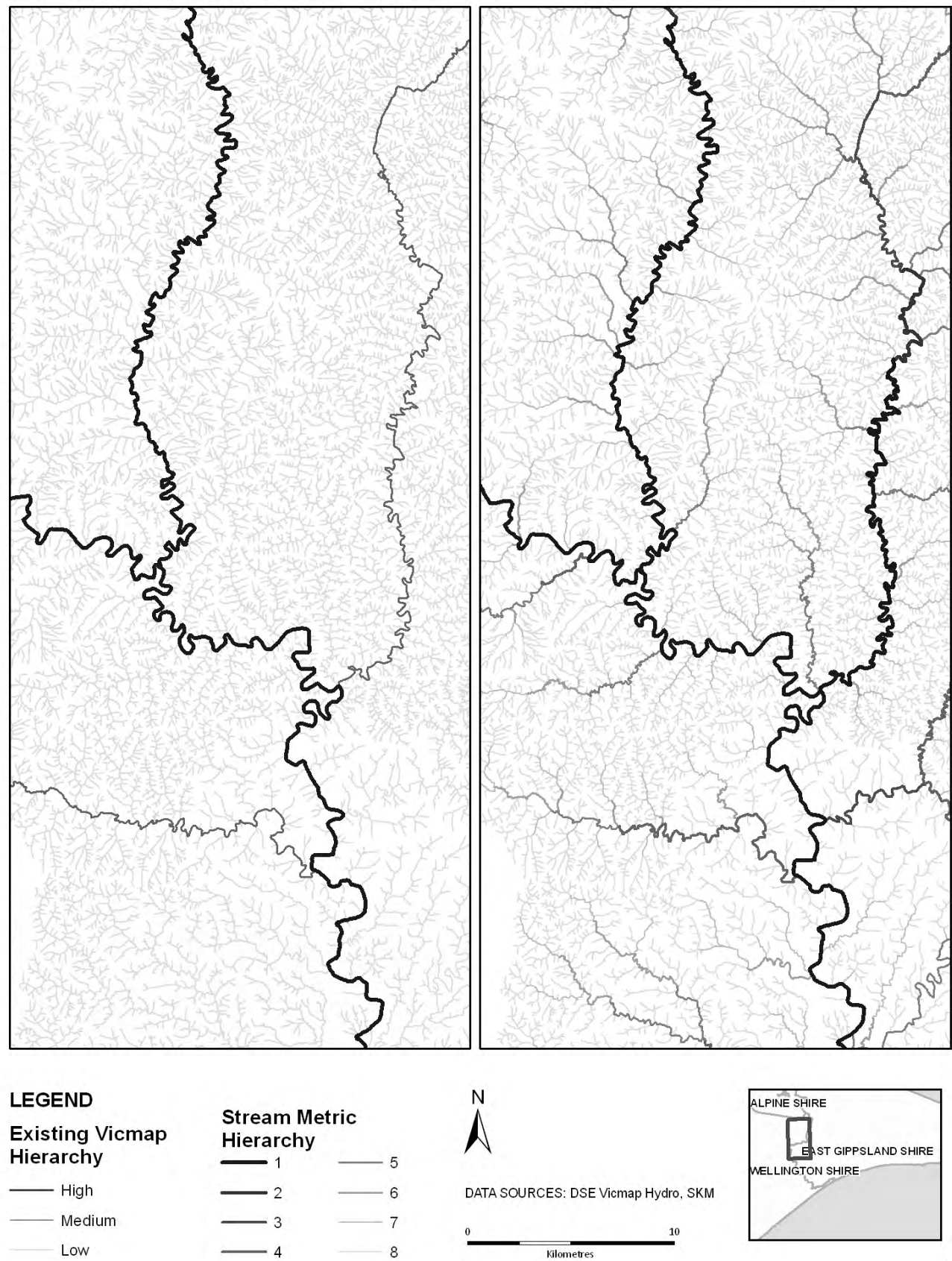


Figure 8: A comparison of existing Vicmap hierarchy and stream metric hierarchy

5 DISCUSSION

5.1 Pfafstetter

A preliminary automated GIS analysis has been developed to implement the Pfafstetter code for the Mitchell River basin. As this approach has been developed to apply to the Mitchell Basin only, it may need to be modified to appropriately apply the Pfafstetter code in each basin. This is a result of the varied topology of each basin in Victoria. Expected variations to the approach include allowing for intermittent streams (e.g. in the Wimmera), a large increase in the number of small networks (e.g. around the Otways), or the inclusion of braided stream segments (e.g. the Murray River).

The best use of a Pfafstetter coded stream network is with an associated catchment network. This would allow not only the selection of all streams upstream of any point in a basin, but also an accurate estimate of catchment area from these points. This is the method used in the Australian Bureau of Meteorology's National Geofabric dataset. The recommendation from this pilot study is that a catchment network be created along with the Pfafstetter coded dataset. In the same way that the Geofabric has become the standard catchment analysis layer for the country, this catchment network could be used as a standard catchment basis for Victoria, allowing the development of many more applications across the state.

5.2 Hierarchy

5.2.1 Ratio of TUL and Total Network Length

The ratio of TUL to network length classified streams effectively in the larger stream networks within a catchment. However, this method over represented small stream networks. Using this method, a segment with a small TUL in a large network would be given a higher classification than a segment with the same TUL in a small network. For example, a stream segment near the top of the Mitchell River may have a TUL of 100m, in a network of 1,000,000m, giving it a hierarchy ratio of 0.0001, while a stream with a TUL of 100m in a network of 10,000m would have a hierarchy ratio of 0.01. As this over-representation would affect the assigned hierarchy, this method is not a suitable classification method.

5.2.2 Product of TUL and network length

This method was applied as a way to recognise that large networks have more importance in a basin than small networks. This would allow streams in a large network to generally have a higher classification than those in a small network. This method produced the expected results, where a small segment in a large network scored higher than the same segment in a small network.

However, upon review of this method, there was no clear reason as to why a TUL should not have the same weight in a small network as in a large network. As the TUL is an approximation of flow accumulation, any TUL represents the potential flow capacity of the stream segment. In other words, a stream with 1,000m upstream in the Mitchell River is likely to carry a similar volume of water as a stream with 1,000m upstream in a small creek network. As such, the application of the product method is limited to river systems which do not contain different sized stream networks.

5.2.3 Ratio of TUL and Total Basin Length

The method of classifying hierarchy as a ratio of TUL of a stream segment and the total length of streams within a basin addresses the problems encountered in the first two classification metrics. Using TUL with the total basin length results in a classification ratio where stream segments with the same capacity to carry flow will have the same classification. This is an appropriate method to apply over a single catchment, assuming the climate is likely to be similar over this area.

However, due to the varying climate across the state, the classification method should not be applied either just by TUL or a ratio of TUL with the total length of streams in Victoria. While any

stream with TUL of 'x' will have the same capacity to carry flow, the likelihood of this stream carrying the same volume needs to be taken into account. This will mean that the ratio classes defined for the Mitchell Catchment will not be relevant in the Wimmera catchment, so a number of different ratio classes will need to be applied across the state.

5.3 Maintenance

The Vicmap Hydro product is maintained by a service provider who updates stream locations and connections as new aerial imagery or elevation models become available. Along with the new stream segments created, attributes must also be maintained. The Pfafstetter code can be applied in a logical way on the Vicmap Hydro dataset to provide a consistently coded dataset across the state. However, this code will not be able to be maintained in an efficient manner, as the Pfafstetter code divides up a stream network into smaller and smaller sub-catchments, the addition of new stream segments would require the whole code to be reassigned. As such, it is likely that the Pfafstetter code would be an annual or biennial product, separate to the watercourse layer.

The hierarchy and downstream identifier (ID) attributes will be easier to maintain than the Pfafstetter system. This is due to the fact that these attributes can be calculated once for each segment, then manually maintained throughout the year. The downstream ID will simply be entered during the editing phase, and a stream hierarchy can be assigned on the basis of the hierarchy of surrounding streams. Both of these attributes could be verified during the update of the Pfafstetter code.

5.4 Limitations

5.4.1 Pfafstetter

While the Pfafstetter system is a logical, effective hydrology numbering system, is not an intuitive system to understand. However, the logic statements needed to understand and use the system are quite simple, and follow a very small number of rules, as discussed in Section 4.1. As such, there would need to be supporting information with the Pfafstetter dataset to allow users to understand and use the data.

There are two practical obvious methods for this:

- Create and supply a plain language user guide
- Develop a simple plug-in tool

As one of the aims of the Pfafstetter system is to allow users with limited spatial knowledge or technology limitations to analyse and use this coding system, the user guide is the most broadly applicable information that can be supplied with the system to aid understanding. However, if it becomes clear that users of this system mostly rely on one or two software products, a tool would also be easy to create and supply with the data.

5.4.2 Disconnected Streams

Another issue for the successful application of Pfafstetter is the incomplete stream connections within the Vicmap Hydro watercourse layer. A stream that physically flows into the main network, but is not connected in the watercourse layer will not be classified the correct Pfafstetter code. As these streams exist as a separated network, they will be assigned their own Pfafstetter code group, applicable only within this isolated network. This will affect the upstream selections as these disconnected networks will not be able to be selected with the simple attribute selection statement. The system would be more applicable if the streams were reviewed to introduce the required connections. Some examples of separated stream networks which should be connected to the main stream network can be seen in Figure 9.

5.4.3 Hierarchy

The metric based hierarchy classification is suited to classification based on equivalent flow potential. However, this classification technique will mean that a named watercourse can have multiple classifications. For example, Crooked Creek has a TUL:Basin Length ratio varying from

0.002 to 5.4, which covers 7 of the 8 classification classes used in this study. While this is an accurate representation of the stream metric classification, it may not be as useful cartographically. A potential fix to this would be to assign any names stream the maximum class for that name. However, this would contradict the scientific basis behind a stream metric driven classification system, as not all streams that carry the same volume would be classified equally.

VICMAP STREAM NETWORK ERRORS

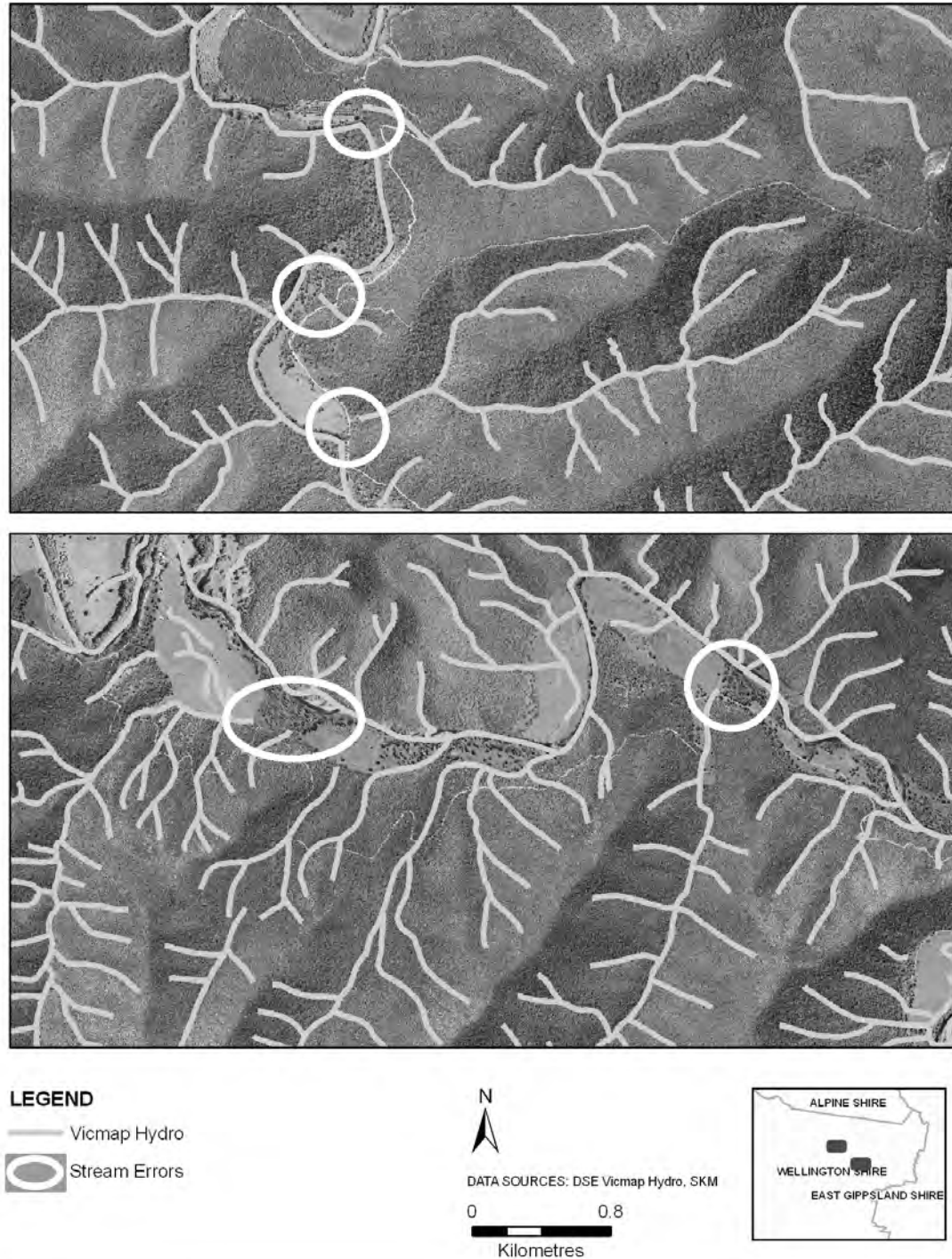


Figure 9: Examples of stream networks that should connect to the Mitchell River, but where no such connection arc exists

6 CONCLUSION

The application of both the Pfafstetter and Stream Hierarchy attributes to the Vicmap Hydro product in the Mitchell River increases both the ease of use and number of application areas of the Hydro datasets as a product. As such, it is recommended that new hierarchy attributes be created for the Vicmap Hydro dataset to allow for both detailed cartographic representation and basic hydrology analysis.

In addition, an annual or biennial Vicmap Hydrology dataset containing Pfafstetter coded stream segments should be created. This should include a layer containing catchments for each stream segment to gain the most benefit from the Pfafstetter system. This product of stream segments and catchments would be supplied with attributes detailing the Pfafstetter Code, the Basin Name and number, the Network Number, Total Upstream Length and Downstream Segment ID.

This study has successfully demonstrated two methods for improving the use and application of the Vicmap Hydro product. The metric based hierarchy classification is an effective method of assigning watercourses with the same hydrological capacity to the same class, and this classification is applicable state-wide. In addition, the creation of a Pfafstetter coded stream and catchment layer would allow the analysis of a hydrologic network without additional data or additional investment in software packages. In the same way that the Australian Bureau of Meteorology's Geofabric has become the standard catchment analysis layer for the country, this catchment network could be used as a standard catchment basis for Victoria, allowing the development of many more applications across the state.

7 ACKNOWLEDGMENTS

We would like to acknowledge Darren McKinty and the Department of Sustainability and Environment for identifying the need for this hierarchy update, and funding the investigation into the appropriate methods discussed in this report.

REFERENCES

Bureau of Meteorology (2010). Australian Hydrological Geospatial Fabric (Geofabric) Data Dictionary – Surface Catchments, Version 1.0, Commonwealth of Australia. October 2010

Department of Sustainability and Environment (2010). Vicmap Hydro: Mapping Business Intelligence. Spatial Information Infrastructure, Victorian Government, June 2010.

Department of Sustainability and Environment (2011). Vicmap Hydro Product Description version 5.2. Spatial Information Infrastructure, Victorian Government, February 2011.

Geoscience Australia (2011). GEODATA TOPO 250K Series 3 - Online and customised (including MapConnect).<http://www.ga.gov.au/meta/ANZCW0703008969.html> Commonwealth of Australia. Last accessed: 24 March 2011.

Pfafstetter, O. (1989). Classification of hydrographic basins: coding methodology. Departamento Nacional de Obras de Saneamento (DNOS), 19 p

Wilson, P. and Nason, S. (1991). SACRED Stream and catchment references for environmental data, a stream numbering system for Victoria. Department of Conservation and Environmental, Victorian Government, June 1991.

Shreve, R.L. (1966). Statistical law of stream numbers. *Journal of Geology*. 74, 17-37.

Strahler, A.N. (1957). Quantitative analysis of watershed geomorphology. *Transactions – American Geophysical Union*, 38, 913-920.

Verdin, K. L. and Verdin, J. P. (1999). A topological system for delineation and codification of the Earth's river basins. *Journal of Hydrology*, 218(1–2), 1–12

The Victorian Land Use Information System (VLUIS): A new method for creating land use data for Victoria, Australia

Elizabeth Morse-McNabb

Department of Primary Industries
PO Box 3100, Bendigo Delivery Centre, Bendigo, Victoria, 3554
Elizabeth.Morse-McNabb@dpi.vic.gov.au

ABSTRACT

Detailed land use mapping began in Victoria in 1996 and the first full coverage of the state was completed in 2005 (stored as a digital spatial data layer: landuse100). The process was extremely labour intensive, both in the office and in the field, and the method required large amounts of data, including aerial photography and satellite imagery. Until 2009 the landuse100 dataset was the most current land use information available, however, it was not consistent across the state and had lost currency. Recreating a landuse100 dataset using the same method would take many years and resources. A new way of generating land use data was needed and developed in the Victorian Land Use Information System (VLUIS) project. This method is both practical and efficient and consequently for the first time had enabled the creation of a land use time-series information for Victoria. It was critical at first to identify, by user needs analysis, the elements that needed to be mapped, how often and at what scale. VLUIS maps land tenure, land use and land cover across the whole state, for the first time on an annual basis, using the cadastral parcel as the spatial unit. This provides the what, where and when of each land information component. Secondly and most importantly a hierarchical link has been established between land tenure, land use and land cover. Classifying these three land information sets separately but joining them together in the one spatial unit base has established greater context and increased the utility of the information. This process required the integration of three independent data sets (two pre-existing and one created for the VLUIS) with the state cadastral parcel map base. The first dataset using this method was created in 2009.

KEYWORDS: land use, land cover, land tenure, state-wide

1 INTRODUCTION

The Victorian Land Use Information System (VLUIS) is a new method for creating land information for the state of Victoria. The process described in this paper began in July 2007 and was developed over three years to create the first VLUIS product in 2009. The methodology is currently being repeated and refined. By 2014 the method will have created four new fully validated datasets for the years 2009, 2010, 2011 and 2012 and an historical annualised dataset from 2004 to 2008.

The land use information available in Victoria prior to the development of the VLUIS product was a conglomeration of a jigsaw of maps created from 1996 to 2005. The process was extremely labour intensive, both in the office and in the field, and the method required large amounts of data, including aerial photography and satellite imagery. Until 2009 the landuse100 dataset was the most current land use information available, however, it was not consistent across the state and had lost currency. Nevertheless, landuse100 has proved an extremely valuable resource for many primary industries research projects, catchment hydrology modelling projects, land use planning, and land use policy processes. The number of end users and types of applications for land use

information is unparalleled by any other dataset (Morse-McNabb 2008). However, the cost and time taken to record and create the land use map was extensive taking nine years to complete across the state. Even though the use of the data was so high, it was difficult to justify repeating the process, especially as the time taken to produce new information meant it was often outdated before the final product was available to users. The latency in the production of the land use information by that method prohibits the generation of a meaningful time series. These were the drivers for the process of designing the new method that began in VLUIS in 2007.

There are very few examples of state or national scale land use mapping campaigns that map both land use and land cover information and are repeatable. Most land use and land cover mapping processes are focused on small regions (Cumming and Barnes 2007), only map land use or land cover (Running et al. 1995; Giri et al. 2005; Oñate-Valdivieso and Bosque Sendra 2010; Sulla-Menashe et al. 2011), or are a one-off snapshot (Wulder et al. 2007). Some methods attempt to produce temporally consistent mapping approaches by linking to previous maps, however; there are often many years between the maps produced (Sommer et al. 1998; Tapiador and Casanova 2003; Wickham et al. 2004; Lowry et al. 2007; Centre for Ecology & Hydrology 2011; European Environment Agency 2011). A classification was created by the National Land Use Database in the United Kingdom to map land use and land cover that identifies the very important principle of classifying each land component, such as land use and land cover, separately. However there is no further evidence that this has been put into practice (LandInform Ltd 2006). There are no readily accessible examples of land use or land cover mapping procedures that map large areas regularly and certainly none that map land tenure, land use and land cover annually. Very few methods are designed to be regularly repeated, and consequently little consideration is given to class definitions, level of detail required, map extent and the cost of production.

Creating a land use dataset is not just a surveying process, data integration process or a classification process; it is a combination of many aspects of spatial sciences, statistics and land classification. To create a new approach to the generation of land information we had to be cognisant of all three aspects as well as understand the likely user base and potential applications of this information. Most users of land use information have some spatial skills as a minimum; however, all understand their area of application very well. A Land Use Summit was held at the beginning of the VLUIS project (Morse-McNabb 2008). We invited users and creators of land use information in the state of Victoria to explain their current use and the ideal dataset required for their application. All 48 attendees from 28 organisations used land use data in a different way but there was a consistent message about the fundamental dataset requirements. The data had to have a consistent spatial footprint, be created regularly and provide a clear and consistent classification. Without these three elements it would not be possible to detect change, which has been a data gap and was expressed as an ideal requirement by the summit attendees.

Following the user summit, we concluded that it is not possible to create a dataset that matches all the requirements of every user's particular application; but it is possible to create a dataset that is consistent and repeatable and meets the bulk of user needs. Our aim was to create a dataset that covered the whole state on an annual basis, was linked to the state cadastre for spatial consistency and used set, consistent classifications.

This paper describes the process, inputs and outputs of the 2009 dataset, which has been in use by various end users for approximately 12 months.

2 METHOD

2.1 Spatial Unit

The state cadastral parcel layer (PARCEL_MP) is the base for VLUIS. PARCEL_MP records the smallest unit of land ownership across the state. Land ownership has a critical impact on the management and use of land and therefore provides a useful spatial context to land information data (SII 2008). The state cadastre has several key spatial data characteristics that influenced our choice of this dataset as a basis for the VLUIS project:

- It has a unique state-wide parcel identifier that can be traced to a particular location.
- It is linked with local government areas and is consistent to their boundaries, and has well defined tolerances.
- It can be linked to the Australian Bureau of Statistics census data via a generalising 'mesh block' number (ABS 2006).

The VLUIS process takes the Parcel Permanent Feature Identifier (PARCEL_PFI) and the local government authority (LGA) field from the dataset. PARCEL_PFI is the field that all other datasets link to. LGA is included as a contextual filter for data users.

2.2 Spatial Extent

The VLUIS product is a state-wide land based dataset. It does not include any mapping of coastal marine parks. The initial development work was limited to seven LGA across Victoria, which were chosen as a representative sample. This was done to make iteration or changes in the process quicker while the methodology was under development. Piloting a process in a small representative area is useful if the final product is very large, as in this case. However; there were many software limitations that were not discovered until the method was applied state-wide. One example is a zonal statistics function in ESRI ArcGIS (ESRI 2008) that had a limit of 100,000 parcels. There is only one LGA in Victoria that has more than 100,000 parcels and this was not one of our 7 pilot areas, therefore the issue was only discovered when all 3.1 million parcels were used during state-wide processing.

2.3 Temporal Extent

The VLUIS dataset is created annually and represents a yearly snapshot of the land tenure, land use and land cover of the state of Victoria. Even though all three land information components can change within a year, only one dominant category for each component is identified for each year. Ground truth data, used to classify land cover type, is collected during spring and early summer to assist in the identification of annual crops. Land tenure information from state-wide land management data sets is constantly updated and can therefore be extracted at any time during the year. Land use information is available from the Valuer General Victoria on a biennial basis.

2.4 Data used

Local government rates for each property in the state are calculated using a state-wide general revaluation every two years. This process is governed by the state government Valuer General Victoria (VGV) and undertaken by valuation officers contracted by local governments. The valuation data records extremely detailed information about the property type and its buildings; infrastructure and location in order to calculate the value. The property type is synonymous with land use. A small proportion of the information collected by the valuations officers is available as 'publicly releasable information' (Victorian Parliament 2010). The "property type" fits into this category and provides an accessible, consistent, detailed classification of land use.

The property type (land use) data is provided as a table with an assessment number as an LGA based identifier. The assessment number is linked to a state-wide property number (PROP_NUM) and this is then linked to the unique property PFI (PROP_PFI) number. A parcel to property linking table is then used to associate land use codes with PARCEL_PFI.

Another state-wide dataset that is embedded in the state cadastre is land tenure. Public land is managed by the Department of Sustainability and Environment (DSE) and they maintain many datasets that detail the lease hold and management arrangements on public land. When we began the VLUIS project a number of datasets were required to get a full picture of public land management across the state but in 2010, a new single dataset was created called PLM25 (DSE 2010). PLM25 is used as the single source of land tenure information since 2010. Unlike the old datasets, PLM25 links directly to the PARCEL_PFI in the cadastre.

There was no state-wide land cover data available for Victoria in 2007. Land cover is a critical component of land use information because land cover changes can indicate conversion from one cover type to another or changes to underlying land characteristics (Martínez and Gilabert 2009). Accurate and timely land cover information is needed because land cover characteristics are the result of complex human interventions and ecological and geophysical processes (Munroe and Müller 2007; Wardlow et al. 2007). Therefore it was important to find a way to create land cover data.

After a review of the literature, a remote sensing approach was developed using time series analysis of MODIS imagery to derive annual land cover (Zhang et al. 2003; Sakamoto et al. 2005; Sakamoto et al. 2006; Clark et al. 2010). MODIS imagery was chosen because it covers the whole state in one image, it is free, and it has many well documented, consistent products.

The standard MODIS product MOD13Q1 was used (NASA 2011). MOD13Q1 is a 16 day composite of daily MODIS imagery calculated as vegetation indices. Using a composite product limits cloud contamination and improves image calibration and quality. The MOD13Q1 product has two vegetation indices, the Normalised Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). We have used EVI for our time series analysis because:

- EVI measures vegetation independently of canopy and soil background effects,
- EVI is more sensitive to vegetation structure, which improves association with leaf area index (LAI), plant area index (PAI), plant leaf structure and canopy structure, and
- EVI has reduced sensitivity to atmospheric aerosols (Huete et al. 2002; Huete et al. 2006).

Each year between October and December, ground-based calibration and validation information is collected across the whole state using a hierarchical stratified random sampling approach. The ground control is primarily used to calibrate and validate the land cover dataset.

2.5 Classification used

The primary use of PLM25 data was to provide a state-wide view of parcel ownership and this is simply represented as either public or private in the VLUIS data (Table 1(a)). The PLM25 land tenure data has also been used to provide a land use for the public land areas. However, PLM25 descriptions are in terms of land management. We have created a small translation (with help from DSE) to convert the classification from land management to land use on public land. To classify land use on private land (equivalent to property type) we used the Australian Valuation Property Classification Code (AVPCC) developed by the Valuer General Victoria, which meant there was no translation necessary for land use data (Department of Sustainability and Environment 2008). This scheme uses nine primary classes and many secondary, tertiary and quaternary classes (Table 1(b)). A land cover classification system was originally based on the FAO Land Cover Classification System (Di Gregorio and Jansen 2005) but it was difficult to interrogate and so a simplified set of land cover classes have been set for Victoria and mapped each year (Table 1(c)).

Table 1: Three separate classification systems are used for the three separate datasets.

(a) Land Tenure	
2 classes	Private Public
(b) Land Use (AVPCC)	
Primary Land Use	Residential
9 classes	Commercial
	Industrial
	Extractive Industries
	Primary Production
	Infrastructure and Utilities
	Community Services
	Sport Heritage and Culture
	Conservation
Secondary Land Use	71 classes
Tertiary Land Use	578 classes
Quaternary Land Use	167 classes
(c) Land Cover	
Dominant Land Cover	Water,
12 classes and an Unknown category.	Softwood Plantation
	Hardwood Plantation
Land Cover classifications for parcels smaller than 12.5 ha are set as Not Applicable	Native Woody Cover
	Evergreen Woody Horticulture
	Deciduous Woody Horticulture
	Non-woody Horticulture
	Pasture / Grassland
	Cereals
	Legumes
	Oilseeds
	Bare ground and non-photosynthetic vegetation (NPV)
Secondary Land Cover	Number of classes based on observations

2.6 Overall method

Python scripting language has been used to automate the VLUI output using ESRI's ArcObjects through the geo-processing framework.

The VLUI product is based on the state cadastral parcel layer (PARCEL_MP). All other datasets join to the PARCEL_PFI number extracted from the PARCEL_MP layer. Land tenure data from PLM25 is joined based on a direct link to PARCEL_PFI. Land use information from the VGV data requires two table joins; PROP_NUM to PROP_PFI and then PROP_PFI to PARCEL_PFI. The land cover data is linked to PARCEL_PFI using the 'zonal statistics as table' function in ESRI ArcGIS toolbox. CMA regions are linked to PARCEL_PFI based on the region at the centroid of a parcel (Figure 1).

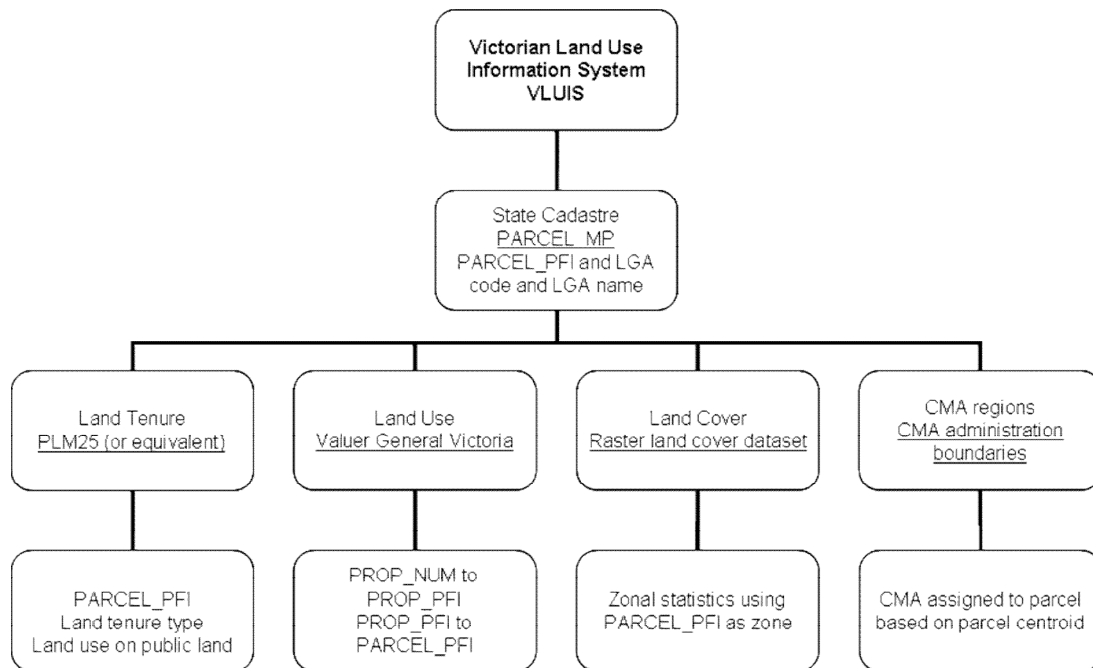


Figure 1: The VLUIS method uses the state cadastre as the base dataset and joins all other datasets using the PARCEL_PFI number

3 RESULTS

The data is combined in one ESRI feature dataset. There are 17 fields in the final dataset, described in Table 2. Parcel_SPI has been included in the final dataset for future linkage to irrigation water use licence information.

Table 2: Field names and descriptions found in the final VLUIS product

Field Name	Field Description
Parcel_PFI	The Parcel Persistent Feature Identifier is a unique number across the state that represents a particular parcel.
Parcel_SPI	The Parcel Standard Parcel Identifier is another unique identifier created from the concatenation of the 'Lot' number and 'Plan' number
Propnum	The LGA Property Number
Prop_PFI	The Property Persistent Feature Identifier is a unique number across the state for each property
LGACODE	Each LGA across the state has a code
LGA	LGA name
TENURE	Classed as either public or private
LU_CODE	A four digit code described in the AVPCC
LU_DESC	A full text land use description that links with the LU_CODE
LU_1	A single digit number that identifies the primary AVPCC class
LU_2	A double digit code that identifies the secondary AVPCC class
LU_3	A three digit code that identifies the tertiary AVPCC class
LC_CODE	A double digit code for each land cover type
LC_DESC	A full text description of each dominant land cover type
CMA	CMA name
Shape_Length	The parcel shape length
Shape_Area	The parcel shape area in m ²

3.1 Accuracy

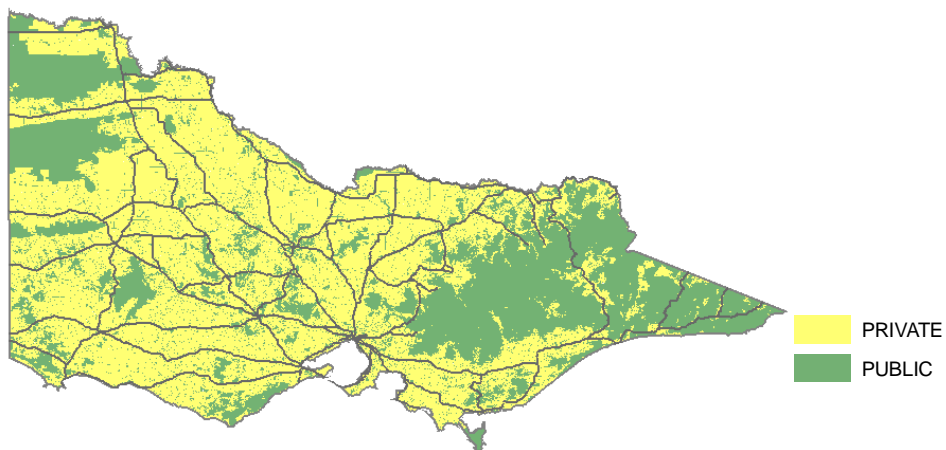
The accuracy of the three VLUIS components was assessed individually. In 2009, 1193 field sites were sampled using a hierarchical stratified random sampling method across the state of Victoria. A 100-fold cross validation of the land cover classes was undertaken using C5.0 software (Rulequest 2009). This process uses different subsets of the data to generate rules and assess the accuracy of the results, giving a better estimation of the true predictive accuracy of the rules set. The 2009 land cover dataset has a predictive accuracy of 76.1%.

The VGV land use data set cannot be assessed for accuracy in the same way that a land cover data set can because it is not always possible to make assumptions about the land use simply by a visual observation. Therefore, other datasets that could be considered 'truth', which had equivalent (or part) land use classification to AVPCC, were sourced across the state. Only two datasets were readily available and of high enough quality to use. The first is a census of perennial horticulture in the northern irrigation region of Victoria, where the land use descriptions agreed 84% of the time from a sample of 654. The second was an industry based collection of agroforestry plantation types and locations across the south west of Victoria, where the land use descriptions agreed 85% of the time from 6336 samples. We intend to find other sources and methods of validation in the future.

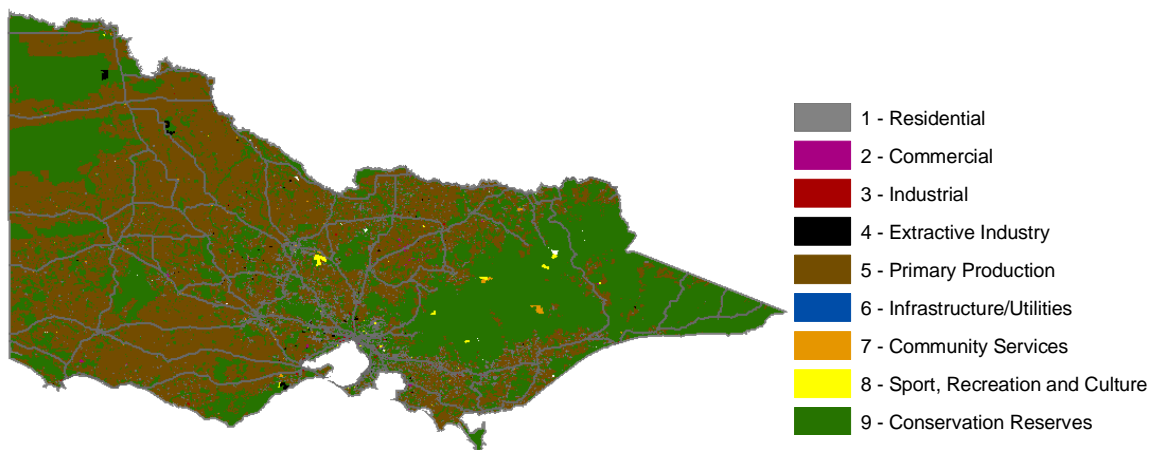
The land tenure data integrated in the VLUIS product may have some omission errors; however; due to the nature of its creation (a limited set of vector data) there are no commission errors. No assessment of the accuracy of this dataset has been undertaken primarily because it is very difficult to validate land tenure. Furthermore, the data is obtained from DSE, the land tenure custodian, and is regarded to be accurate for this purpose.

The resultant information set can be viewed as three separate maps showing land tenure, use or cover (Figure 2).

(a) Land Tenure



(b) Land Use (Primary Classification)



(c) Land Cover

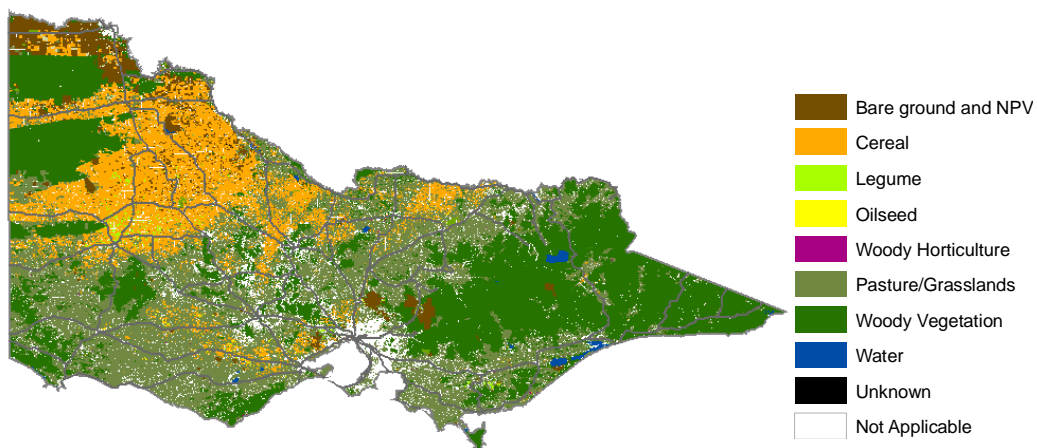


Figure 2: Victorian land Use Information System Data displayed at all three levels of land information across the state of Victoria.

4 DISCUSSION

The VLUIS approach uses three independent datasets, which lead to the following three key outcomes.

1. All three datasets could be used without loss of information in the final product and linked to create a spatially consistent product
2. There is an intrinsic hierarchy between tenure, use and cover information
3. Information accessibility and utility is improved through the VLUIS product compared with using three datasets independently.

4.1 Creation of a spatially and temporally consistent product

One of the main requirements of our potential users is the ability to track change over time. In order to do that it is critical to have a consistent spatial base for the information to be linked to. There are only two options in terms of spatially representing data; a raster base or a vector base. Each has its advantages. For example, a raster dataset is often smaller (in terms of computer memory requirements) and therefore more manageable than a vector dataset. For many of our modelling users it is the format of choice. However; vector datasets have the ability to store more contextual information more efficiently and can identify linear features such as road and rivers at any spatial resolution. The state cadastre is a vector dataset that defines the smallest unit of land ownership across the whole state. We used the cadastre as our base because it is a well maintained dataset that has links to local government and state legislation. The business of maintaining the cadastre and associated information is embedded in many state and local government processes.

After identifying two excellent state-wide datasets (VGV and PLM25) that could be integrated across the states, using the cadastre as the base, the next requirement was to settle on a consistent classification for use now and into the future. The landuse100 dataset used the Australian Land Use Mapping (ALUM) Classification. This is a classification that has been developed by all states and territories across Australia and continues to be revised using a successful collaborative approach (Australian Bureau of Agricultural and Resource Economics and Sciences 2011). We aimed to create one land use dataset from the two state-wide datasets (VGV & PLM25) using the ALUM classification. The translation from the two datasets to the one ALUM code required some assumptions and expert knowledge. We found that reclassifying data into the ALUM classification resulted in the loss of some detailed information, and that in some cases there was not enough information given to allocate an ALUM classification to a parcel. The ALUM classification was difficult to support using the two datasets available because it is a mix of land tenure; land use and land cover information. This introduces confusion because these three information types each describe different facets of the land that can each be separately discriminated (LandInform Ltd 2006). Land tenure relates to the ownership, land use describes the function for which the land is used, and land cover describes the physical surface of the earth (ABARES 2011). A data loss-less situation is achieved by not attempting to reclassify to a single scheme such as ALUM. This enabled us to leave the datasets as they were and simply use the unique parcel identifier to provide the common link to the cadastre. Therefore it was decided to not reclassify the information to fit within the ALUM classification. A mapping to the ALUM scheme can be developed if this is required by users in the future.

4.2 Hierarchy of land information

Even though land tenure, use and cover describe different components of the land and should be separately classified to avoid confusion, they are also intrinsically linked. Land tenure is defined as; 'the form of an interest in land'(ABARES 2011). In VLUIS, tenure is simply classified as either public or private and in this sense land tenure is controlled by legislation and planning policy and therefore limits the possible land uses. Land use is defined as; 'the purpose to which the land cover is committed' (ABARES 2011). This definition noticeably links to land cover type but unfortunately gives no indication of the influence of land tenure. Land cover is defined as; 'the

physical visible surface of the earth' (ABARES 2011). Land cover type does not include any factors below the surface and does not give any indication of condition. Land cover and land use are fundamentally interlinked and influence each other. Therefore, land cover, land use and land tenure is not independent of one another. Land cover type is the result of ownership and management decisions that can be defined by the tenure of the land and the purpose for which the land is used. Land cover type is often constrained by the size of the land unit and whether it is one of many parts of a singular unit.

The hierarchy between land tenure and land use remains constant. The possible range of land use types will always be dependent on the land tenure. Land tenure is a legislative function and cannot be influenced by the land use and will not change based on a land use change. However; the hierarchy between land use and land cover is not always that clear. As the definition of land use above suggests, land use type can be based on the collection of land covers in any one area. If the land use classification is based on only a snapshot of the land covers, an inappropriate land use type may be defined. VLUIIS aims to remove this issue by using the AVPCC property type classification rather than one based on land cover at any one time. AVPCC takes into account property infrastructure and longer term land cover influences (DSE 2008). For example, a small area of a beef-grazing property newly planted to wine grapes does not make the property a vineyard but over the long term this land cover change may create a land use change and certainly highlights a potential to change. VLUIIS is also produced consistently (spatially and in classification) each year, which also limits erroneous land use classification based on a land cover snapshot. This also allows for real land cover change to be tracked and where appropriate be used to indicate a potential land use change.

By identifying the intrinsic link between land tenure, use and cover and establishing a stable hierarchy, VLUIIS can provide a stable base for change detection and analysis. Accurate and timely information is critical to the assessment of the impacts of human activity on our communities and the environment.

4.3 Added-value information

The final VLUIIS dataset has many levels of information that have been demonstrated to link and conform to a hierarchy. Separately classifying each land type means that each individual facet can be described clearly, succinctly and accurately repeated on an annual basis. One common criticism of the landuse100 dataset was the level of classification. The class, 3.2.0, defined as 'grazing modified pastures' gave no indication of what was grazing the pasture or how often and whether it was part of a rotation or permanently in place (ABARES 2011). The VLUIIS data classifies pasture in the land cover dataset and has a similar amount of information as the landuse100 product. However, in VLUIIS the land cover type is also associated with land use and land tenure, which together provides contextual information about land cover. For example, in the case of pasture it may be present in a dairy property, a sheep property, a beef property or a conservation area. This level of information is only available by separately classifying each land component. This information can be used to determine the range of land cover types on dairy farms across the state or to understand the potential grazing regime and therefore evapotranspiration rate of pasture across the state. Linking each of the datasets to the one spatial unit provides extra contextual information to the user by defining social, economic and environmental factors around each land component.

5 CONCLUSION

The VLUIIS is a new method for creating accurate, repeatable and consistent land component information across large areas. The process began with an extensive review of all land use data users and creators. The consultation process identified that there are many applications for land use information in planning policy, natural resource management, emergency management and environmental modelling. The consultation identified the main components of land information required by all users were consistency in spatial unit, classification and time.

The VLUIS method produces annual datasets that classify the land tenure; land use and land cover for each cadastral parcel across the state of Victoria. Land tenure information is taken from the state-wide PLM25 dataset, land use from the VGV general valuation dataset and land cover information is created for the state using time series analysis of MODIS satellite imagery. Each data set is classified using a separate classification. By creating a dataset in this way the VLUIS process has taken advantage of the intrinsic link and hierarchy between the three land components. Linking three individual datasets in one spatial unit has also provided a greater level of contextual information than would have been gained by combining all three datasets in one classification.

The method described in this paper is specific to Victoria. However, the principles used to create a repeatable land use information product remain regardless of geographic area. Land use mapping should be a continuous process, not a haphazard practice undertaken to suit the immediate need of one aim. It is a fundamental dataset that underpins the basis of many natural resource management programs, policy and research projects. If time is taken to understand the required elements (how often, at what scale, what classification) of a land use information set at the outset it is possible to create a method can be consistently repeated year after year as we have demonstrated here with the VLUIS 2009 dataset and method.

REFERENCES

- ABARES (2011). Guidelines for land use mapping in Australia: principles, procedures and definitions. Australian Government. Canberra.
- ABS (2006). "MESH BLOCKS DIGITAL BOUNDARIES, AUSTRALIA " Retrieved First, from <http://www.abs.gov.au/ausstats/abs@.nsf/productsbytitle/73058E7FD2A1342BCA2573FD0015BFB1?OpenDocument>.
- Australian Bureau of Agricultural and Resource Economics and Sciences (2011). Guidelines for land use mapping in Australia: principles, procedures and definitions. Australian Government. Canberra.
- Centre for Ecology & Hydrology (2011). "LCM2007." *Land Cover Mapping* Retrieved 11th August 2011, from <http://www.ceh.ac.uk/LandCoverMap2007.html>.
- Clark, M.L., Aide, T.M., Grau, H.R. and Riner, G. (2010). "A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America." *Remote Sensing of Environment* 114(11): 2816-2832.
- Cumming, G.S. and Barnes, G. (2007). "Characterizing land tenure dynamics by comparing spatial and temporal variation at multiple scales." *Landscape and Urban Planning* 83(4): 219-227.
- Department of Sustainability and Environment (2008). Valuation Best Practice, 2010 Specifications Guidelines.
- Di Gregorio, A. and Jansen, L.J.M. (2005). *Land Cover Classification System (LCCS), version 2: Classification Concepts and User Manual*. Rome, FAO.
- DSE (2008). Valuation Best Practice, 2010 Specifications Guidelines. Department of Sustainability and Environment.
- DSE (2010). "PLM25." Retrieved 11th August 2011, from <http://www.giconnections.vic.gov.au/content/vicgdd/record/ANZVI0803003978.htm>.
- ESRI (2008). ArcGIS 9.3. Redlands CA, ESRI.
- European Environment Agency (2011). "CORINE Land Cover." Retrieved 11th August 2011, from <http://www.eea.europa.eu/publications/COR0-landcover>.
- Giri, C., Zhu, Z. and Reed, B. (2005). "A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets." *Remote Sensing of Environment* 94(1): 123-132.

- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. and Ferreira, L.G. (2002). "Overview of the radiometric and biophysical performance of the MODIS vegetation indices." *Remote Sensing of Environment* 83(1-2): 195-213.
- Huete, A.R., Huemmrich, K.F., Miura, T., Xiao, X., Didan, K., Van Leeuwen, W., Hall, F. and Tucker, C.J. (2006). Vegetation Index greenness global data set. *White Paper for NASA ESDR/CDR*.
- LandInform Ltd (2006). National Land Use Database: Land Use and Land Cover Classification. O. o. t. D. P. Minister. London.
- Lowry, J., Ramsey, R.D., Thomas, K., Schrupp, D., Sajwaj, T., Kirby, J., Waller, E., Schrader, S., Falzarano, S., Langs, L., Manis, G., Wallace, C., Schulz, K., Comer, P., Pohs, K., Rieth, W., Velasquez, C., Wolk, B., Kepner, W., Boykin, K., O'Brien, L., Bradford, D., Thompson, B. and Prior-Magee, J. (2007). "Mapping moderate-scale land-cover over very large geographic areas within a collaborative framework: A case study of the Southwest Regional Gap Analysis Project (SWReGAP)." *Remote Sensing of Environment* 108(1): 59-73.
- Martínez, B. and Gilabert, M.A. (2009). "Vegetation dynamics from NDVI time series analysis using the wavelet transform." *Remote Sensing of Environment* 113(9): 1823-1842.
- Morse-McNabb, E.M. (2008). Land Use Summit. *Victorian Land Use Information System Project*.
- Munroe, D.K. and Müller, D. (2007). "Issues in spatially explicit statistical land-use/cover change (LUCC) models: Examples from western Honduras and the Central Highlands of Vietnam." *Land Use Policy* 24(3): 521-530.
- NASA (2011). "MODIS Land, Vegetation Indices." Retrieved 14/04/2011, from <http://modis-land.gsfc.nasa.gov/vi.htm>.
- Oñate-Valdivieso, F. and Bosque Sendra, J. (2010). "Application of GIS and remote sensing techniques in generation of land use scenarios for hydrological modeling." *Journal of Hydrology* 395(3-4): 256-263.
- Rulequest, R. (2009). See5. St Ives NSW, Australia.
- Running, S.W., Loveland, T.R., Pierce, L.L., Nemani, R.R. and Hunt, E.R. (1995). "A remote sensing based vegetation classification logic for global land cover analysis." *Remote Sensing of Environment* 51(1): 39-48.
- Sakamoto, T., Van Nguyen, N., Ohno, H., Ishitsuka, N. and Yokozawa, M. (2006). "Spatio-temporal distribution of rice phenology and cropping systems in the Mekong Delta with special reference to the seasonal water flow of the Mekong and Bassac rivers." *Remote Sensing of Environment* 100(1): 1-16.
- Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N. and Ohno, H. (2005). "A crop phenology detection method using time-series MODIS data." *Remote Sensing of Environment* 96(3-4): 366-374.
- SII (2008). Product Description - VICMAP PROPERTY Version 5.4. DSE. Melbourne, Spatial Information Infrastructure,.
- Sommer, S., Hill, J. and Mégier, J. (1998). "The potential of remote sensing for monitoring rural land use changes and their effects on soil conditions." *Agriculture, Ecosystems & Environment* 67(2-3): 197-209.
- Spatial Information Infrastructure (2008). Product Description - VICMAP PROPERTY Version 5.4. DSE. Melbourne.
- Sulla-Menashe, D., Friedl, M.A., Krankina, O.N., Baccini, A., Woodcock, C.E., Sibley, A., Sun, G., Kharuk, V. and Elsakov, V. (2011). "Hierarchical mapping of Northern Eurasian land cover using MODIS data." *Remote Sensing of Environment* 115(2): 392-403.
- Tapiador, F.J. and Casanova, J.L. (2003). "Land use mapping methodology using remote sensing for the regional planning directives in Segovia, Spain." *Landscape and Urban Planning* 62(2): 103-115.

- Victorian Parliament (2010). Valuation of Land Act 1960. No. 6653 of 1960. Version No. 121.
- Wardlow, B.D., Egbert, S.L. and Kastens, J.H. (2007). "Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains." *Remote Sensing of Environment* 108(3): 290-310.
- Wickham, J.D., Stehman, S.V., Smith, J.H. and Yang, L. (2004). "Thematic accuracy of the 1992 National Land-Cover Data for the western United States." *Remote Sensing of Environment* 91(3-4): 452-468.
- Wulder, M.A., White, J.C., Magnussen, S. and McDonald, S. (2007). "Validation of a large area land cover product using purpose-acquired airborne video." *Remote Sensing of Environment* 106(4): 480-491.
- Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C. and Huete, A. (2003). "Monitoring vegetation phenology using MODIS." *Remote Sensing of Environment* 84(3): 471-475.

Issues related to the use of non-traditional, digital data sources for enhancing park management data

Monique Elsley, William Cartwright

RMIT University, School of Mathematical and Geospatial Sciences
GPO Box 2476, Melbourne VIC 3000, Australia
monique.elsley@student.rmit.edu.au, w.cartwright@rmit.edu.au

ABSTRACT

The Web provides access to digital archives containing information made available by news organisations, governments, libraries and the like. In the New Web era, data providers now also include the general public who can contribute information using collaborative and participatory tools, available via the Web or mobile devices and collectively known as Web 2.0. With the arrival of geospatial platforms, mapping tools and devices that are location aware, a growing amount of these data is geo-referenced.

A research project is investigating if the existing data archive of a park management organisation can be enhanced through amalgamation with these digital data repositories available via the Web - alternative data sources that traditionally have not always been considered. The project is developing a methodology for a 'geo-knowledge tool' to access data based in part on their geographic attributes. The methodology comprises a conceptual model and demonstration prototype.

This paper addresses two issues associated with the data and theoretical methodologies devised to address these. The first issue addressed concerns the variation in the georeferencing of data including different methods like coordinates and geotagging, and a variation in geographic detail or precision. A basic geographic framework is described comprising a fixed, formal component and a dynamic part that combined recognise official, conventional and personal georeferences thus allowing for the broadest possible range of relevant data to be presented.

A second issue identified relates to the 'accuracy' and 'usefulness' of alternative data. It is argued that setting clear guidelines for these two terms is not straightforward, and their importance may vary depending on the purpose and opinion of the user of the data. A methodology is proposed that uses a number of data attributes to derive confidence ratings. These can be attached to the data to assist users in assessing whether the data will be suitable for their purpose. To address the complexity of deriving a single rating, one solution would be to make the system a user driven rating alongside a 'formal' rating guide. The second solution proposes to create confidence ratings for individual data attributes with users to decide which attributes rank higher for their purpose. In both cases it is envisaged that the confidence ratings model should become an evolving model with users attaching ratings as the data are being used.

KEYWORDS: geo-knowledge tool, Web 2.0, alternative data sources, user contributed information, data quality, georeferencing

1 INTRODUCTION

The information age has seen an increase in available information (Brown and Duguid, 2000; Dykes and Mountain, 2003). Furthermore, in the New Web (Tapscott and Williams, 2008) era, information is not provided by experts alone. Instead, information is provided via the Web by experts and non-experts alike, with digital data repositories made available by organisations such as governments, news companies and libraries as well as the general public and other non-

traditional data providers. People are able to contribute information using participatory and collaborative Web tools and applications like blogs, wikis and media sharing sites that are made available via the Web or mobile devices, in effect creating digital repositories of user contributed photos, videos, opinions or observations.

A current research project at RMIT University is investigating if the existing data archive of a state-managed park organisation can be enhanced through amalgamating it with these existing Web sources that traditionally have not always been considered. A methodology is being developed using the concept of a 'geo-knowledge tool' (GKT) to provide data access based in part on the data's geographic attributes (Elsley and Cartwright, 2011). The methodology comprises a conceptual model and a demonstration prototype using a case study to focus the research.

This paper addresses some of the issues associated with the data that have been encountered and the theoretical methodologies devised as part of the conceptual GKT being developed for this research project. After a brief overview of the research project that includes the background to the project and its objective and methodology, the issue of variation in georeferencing of data is discussed. This variation is found within the existing park management data archive as well as the non-traditional data sources. A basic geographic framework is described to address the issue. Secondly, an assessment of the accuracy or usefulness of the digital data repositories is covered, incorporating a theoretical methodology to derive confidence ratings or quality indicators to attach to the data using a number of data attributes. A conclusion completes the paper.

2 RESEARCH PROJECT

Parks Victoria (PV), the body that manages parks on behalf of the Victorian Government in Australia, is a collaborator on the research project. The organisation is geographically distributed with a central head office, regional management offices and close to 3000 parks and reserves to manage that are spread out over the State. PV possesses a vast amount of data collected over the years, which are inherently georeferenced – linked to geographic locations such as the whole park network, an individual park or a location within a park. The data are stored in different formats and various storage mediums, and it is hence suggested that the data are not used as effectively as they potentially could, and are not always readily accessible to support decision making.

The research project is developing a methodology for accessing PV's existing data using the concept of a GKT as a data interface. The GKT is regarded to be a digital information or knowledge system that in part utilises the geographic data attributes to present the data. In addition to PV's data, the tool will also give access to rich data available on the Web to potentially enhance the park management data archive.

Contemporary Web concepts are being considered for the GKT's design. These concepts comprise the participatory and collaborative aspects of the New Web - also known as the Social Web (Kamel Boulos and Wheeler, 2007) or Web 2.0 (O'Reilly, 2005). A key notion of Web 2.0 is the suggestion that information contributed by users of participatory tools - user contributed information (UCI) - has the potential to be useful (O'Reilly, 2005; Surowiecki, 2005) and to enhance existing information (Vickery and Wunsch-Vincent, 2007). This research project is assessing this for an existing park management data archive. Aggregating information from individuals through crowdsourcing is one way to capture Surowiecki's (2005) 'wisdom of the crowd' – combining the knowledge of individuals to create a collective intelligence.

Geospatial Web developments are also being considered. Geospatial platforms such as *Google Earth* and *Google Maps* and mapping tools are now freely available via the Web, and allow people to access and visualise georeferenced information in a variety of ways. There is also a geographic element in many Web 2.0 applications, described by Turner and Forrest (2008) as *Where 2.0*. For example, information contributed by users of participatory tools is increasingly georeferenced in nature – so-called volunteered geographic information (Goodchild, 2007), collaboratively contributed geographic information (Bishr and Mantelas, 2008) and user generated geo-content (Das and Kraak, 2011). This is due to the ready availability of location aware consumer devices like GPS enabled mobile phones and cameras.

2.1 Objective

The primary objective of the research project is to develop a methodology for making better use of PV's existing knowledge resource through enhanced data access. The question that is being addressed is whether applying contemporary Web concepts can benefit the methodology, and result in a useful GKT. The tool will not just provide access to this existing archive but also to information available via the Web in the form of digital data repositories provided by organisations and through participatory tools and applications. The research will investigate whether UCI and Web repositories can enhance PV's existing data, and improve knowledge or fill information gaps. It will thus be assessed whether non-traditional data sources can be amalgamated with traditional data to form part of an effective knowledge base to potentially aid decision making and support park management.

For this research, users are defined as being not just Web users but are viewed from a broad perspective and include staff, visitors to the park, other organisations and the general public for example. Although UCI is generally associated with users of participatory tools, in the context of this research project UCI is regarded as information contributed by users in the same broad sense, and includes existing data repositories on the Web.

2.2 Methodology

The initial exploratory phase of the research, now completed, gathered background information on topics such as Web 2.0, knowledge and knowledge systems and examined projects that applied Web 2.0 concepts like crowdsourcing and geographical Web tools. The current phase is using this background research to assist the primary goal of developing a methodology for a GKT. A case study is helping to focus the needs – the study area is Wilsons Promontory National Park (WPNP), one of the parks managed by PV, whilst the data focus relates to fire management and prescribed burning in parks. A number of activities were undertaken as part of the case study including a park visitor survey at WPNP (Elsley and Cartwright, 2011).

The development of a methodology for the GKT comprises a conceptual model and a demonstration prototype. The conceptual model's four broad components address 1) 'data in' (the potential data sources the tool can draw on), 2) the users, 3) the functionality, and 4) 'data out' which includes topics such as data classification and visualisation. The conceptual model shows how individual parts of these four components are connected as well as issues to be considered. These issues were generally identified during the explanatory phase of the research project and would need to be addressed if the GKT is ever built. Issues encountered include privacy, copyright and data ownership (particularly with regards to UCI), equity and accessibility, variation in geographic referencing of data as well as data quality or accuracy. These latter ones – variation in georeferencing and data quality – are discussed in the remainder of this paper.

3 VARIATION IN GEOREFERENCING OF DATA

One of the capabilities of the proposed GKT will be to access data based on their geographic attributes or 'geo-knowledge', the information or knowledge relating to geographic places. The data need to be georeferenced or tagged with the geographic attributes of place so that these attributes can be used to retrieve and present the data accordingly. This is arguably not straightforward when considering the variation in georeferencing of PV's data and the alternative data.

3.1 Observations for existing park management data

Let us firstly consider PV's existing data. Key data that are used regularly and are accessible via various information systems are classified and stored in a hierarchical manner, so that they are relatively easily found. Two initial categories generally applied when accessing PV's information are 'teams' (departments in the organisation) and management areas such as natural values management and fire and emergency management. These two categories are not geographic *per se*, although the physical location of a team is geographic. But like many organisations that are

spread out geographically, a team's location may not necessarily be a contiguous area. 'Regions', a subdivision of teams, is mainly geographic – the Central, West, East and Melbourne Regions have geographic boundaries - but other team subdivisions do not have such boundaries or else to a lesser extent. Regions are divided into districts which in turn encompass individual parks. PV's structure of regions and districts are regarded as the two broadest geographic attributes applied.

PV uses the regions and districts as initial geographic divisions for its map based internal information system. This system has GIS data at its core and can display data from various management areas. For a general query, searchable geographic attributes include management areas (requiring regions and districts as input), park names, work centres, place names and geographic coordinates. Specific queries can utilise non-PV specific geographic divisions such as bioregions when completing a flora search.

Official data records are maintained within an appropriate record and information management system organised by work areas or units and topic. Although some work units and topics may have a clear geographic connection – the regions and districts or specific programs or projects associated with a geographic area like Nine Mile Creek – these are accidental rather than deliberate georeferencing. These geographic descriptors would also be broad in nature, for example, research in a small area of a park in the Alpine District would only have Alpine District as its geographic attribute. Additional information can be attached to the record using free text but it is not mandatory to include a geographic reference here (J Wotton 2011, pers. comm. 10 August).

Geographic coordinates are generally applied to data covering smaller areas and point locations, or data within parks – some of these locations may not have geographic names and require descriptions instead. New data now being collected within parks have coordinates attached (J Whelan 2011, pers. comm. 23 February) using GPS devices, in line with a current initiative to distribute mobile GPS and GIS devices to park staff for field data collection. These geographic coordinates provide the most accurate geographic location, whilst various other geographic scales still apply (park name, region and so on).

Unlike the data in the internal map based system, other PV data do not necessarily have a clear geographic attribute associated with them. As aforementioned, this is partly due to the non-geographic nature of two main organisational divisions (teams and management areas). There is also not a clear and appropriate geographic attribute in use for data at the broadest level, that is, data that apply to the whole park network or to PV as an organisation. Any data that apply statewide are generally regarded as corporate, and may be associated with the organisation's head office that has a geographic location (a building in Melbourne). However, such a point location may not be a very useful geographic representation for data applicable to the whole park network. Although some corporate data may only be applicable to staff or departments at head office, statewide documents that apply to all parks and staff should be tagged accordingly so that they can be represented on a statewide map.

Other formal and non-PV geographical divisions in place are also applicable to PV's data, yet the borders of some of these do not necessarily align with the borders of PV's divisions. Bioregions were mentioned earlier, and fire management related data used for the demonstration prototype for example also fit within a second geographical structure. PV undertakes its fire management and prescribed burning activities in close collaboration with the Victorian Government's Department of Sustainability and Environment (DSE). It subsequently follows DSE's geographic division for fire area and fire regions that differ geographically from PV's own divisions. WPNP thus sits in PV's East Region and Wilsons Promontory District management structure, but for fire management the park falls under the Central fire region and South Gippsland fire district. Other parks in PV's East Region however can fall under different DSE fire regions and fire districts.

Lastly, there can also be a variation in the detail or precision of the geographic attributes of data from different eras or sources; some data provide generalised geographic information that are broad in nature, whereas other data give more exact and precise geographic details. As an example, a database and GIS file were recently created at WPNP that contain details of research papers and reports held in filing cabinets about projects conducted in the park over the years. The attributes attached to individual records include a reference to the geographic location involved, and obtained from geographic information in the textual documents. It became clear that the accuracy of those geographic descriptions and details varied from specific to general (K Bennetts

2011, pers. comm. 02 March) resulting in a map that shows where research has taken place with varying certainty and precision.

3.2 Observations for alternative data sources

The alternative data sources the GKT will provide access to comprise existing digital data repositories and UCI. The location information of these data is very variable, including differences in methods and formats, accuracy and precision, or even if there is a geographic reference. The Australian Broadcast Corporation (ABC) (www.abc.net.au) for example, applies geotags to their articles in a standard format – melbourne-3000, vic, australia and so on. Although geographic information contributed to crowdsourced projects like OpenStreetMap (www.openstreetmap.org) most likely displays specified geographic attributes, with UCI in general specifying location is up to the users hence a greater variation in georeferencing exists. A feature of Web 2.0 is tagging – attaching descriptive keywords to digital data (Anderson, 2007) – whereas geotagging attaches geographic information or metadata (Luo et al., 2008). This can be done through visible textual tags as found on *Flickr* (www.flickr.com) or the ABC website, or through the use of coordinates found in metadata. Images taken with GPS enabled cameras have such metadata attached and if uploaded onto a photosharing site like *Flickr*, these geographic coordinates can show where the picture was taken.

There are a number of issues that need to be recognised with regards to geotagging of some of these alternative data sources and UCI in particular. Firstly, textual geotags that refer to the same place can vary widely. This is because of the ‘personalisation’ aspect of Web 2.0 (Lackie and Terrio, 2007; Shirky, 2006) that lets users attach tags and geotags that make sense to them. Some may use official geographic place names, whereas others may use variations or abbreviations of such names or use different terminology altogether - what Goodchild (cited in Gravois, 2010) refers to as the democratisation of geography. Looking for example at the photos on *Flickr* that relate to WPNP, the most common tag used is ‘Wilsons Promontory’. There are already variations in this place name: it can be a single tag with space in between the two words, one tag but no space (wilsonspromontory) or two tags ‘Wilsons’ and ‘Promontory’. The term Wilsons has variations too – ‘Wilson’s’ being the most common. Also spelling mistakes are made, with Promontory sometimes spelled without the ‘n’ (promotory) and Wilsons missing the final ‘s’ (Wilson). Then there are the more colloquial versions: Wilsons Prom, The Prom and Prom, with one or two tags and space/no space variations. The official geographical name Wilsons Promontory National Park is also used with multiple tags, as one tag with or without spaces (for example wilsonspromontorynationalpark), with National Park abbreviated to NP or written as one word or with the same spelling errors as before. Locations in WPNP are tagged with similar variations and errors. Jones and Purves (2008, p. 221) identify the issue of “vague geographic terminology” when it comes to geographic information retrieval, which comprise these personalised and colloquial geographical terms. Whereas the latter may give a clue as to the real location, the former can include references that may not. Consider places like ‘Mum’s favourite beach’ or ‘Lisa’s hut’ that indicate a geographic location to the person that attached the information but will not make sense to most others. It is possible that a more exact location could be inferred by looking at other details like other data uploaded by that user or their own geographic location, but that will not always be the case.

A second issue is the variation in accuracy and precision of the georeferencing, similar to the issue encountered during the creation of the research database described earlier in section 3.1. Apart from the correctness of the actual geographic location, there is the precision of the information. Geotagging an image with ‘Wilsons Prom’ is not incorrect *per se* but if the image was taken at a specific beach, the ‘Wilsons Prom’ geotag is not very precise. Similarly, the only town in WPNP is Tidal River where most people stay. Using this as a geotag for a picture taken somewhere else in the park may remind the user of their stay at Tidal River, but it is not geographically correct.

If digital data is also geotagged with geographic coordinates, this can alleviate the ambiguity of the textual geographic reference. Although only data generated with GPS enabled mobile devices are likely to have such accurate coordinates, these devices are growing in use, capturing an increasing variety of information apart from images. What does need to be taken into consideration in this instance is the difference between ‘source location’ and ‘target location’

(Amitay et al., 2004). That is, someone tweeting about their trip to WPNP (geographic feature) from a café in Melbourne afterwards and uploading images from their trip on their home computer means the geographic coordinates associated with the place of uploading (geographic source) can be ignored, and attention should be paid instead to the content and tags (geographic feature).

Finally, another issue identified by Jones and Purves (2008) with regards to UCI is also applicable to WPNP and PV. The fact that some geographic place names have non-geographical meanings means that search results using the tags 'Wilsons' and 'Prom' on *Flickr* also produce images of a 'prom' (American school dance) from people called 'Wilson'. And the non-geographical meaning of 'tidal' and 'river' means that *Flickr* search results for the tag 'Tidal River' in the first instance do not give any reference to the campsite in WPNP but instead show Washington DC and Burton in Wales as the top geographic locations associated with the tag.

3.3 Discussion

It is argued that PV's data have a range of geographic scales or levels inherent within them although these are not always obvious or made clear, but can be inferred instead. This geographical hierarchy ranges from data applicable to the whole organisation down to an area or location within a park defined by geographic coordinates (additionally there is a larger extent beyond PV that encompasses Australia and beyond). Being in the business of park management, the vast majority of PV's data are associated with one or more parks, an overarching district or region or else staff working in those areas, meaning PV's geographic attribute 'work centre' is applicable. If that is not the case, the broadest geographic scale should apply - data associated with PV as an organisation and thus the whole park network. Even data from geographically dispersed departments can be assigned a geographic area, albeit not necessarily a contiguous one. Only very limited data, if any, seems not to be geographically linked at all.

Geographic coordinates attached to data generally provide accurate locations associated with those data. The increased use of geographic coordinates for field data at PV will provide similar accuracies and improve geographic attributes. Geographic coordinates attached to UCI and other Web data, particularly if generated by GPS enabled devices, can be a useful means to position these alternative data sources and link them to appropriate PV locations. However, these coordinates are not generally regarded as being useful when it comes to searching for data by location. This is simply because many people would not know exact coordinates and even knowing approximate ones may not necessarily lead to the correct location. Geographic coordinates are therefore useful to associate data with exact locations (give or take the GPS devices' error margins), but additional geographic attributes using keywords or tags are required to make them more easily searchable.

3.4 A basic geographic framework

In order for the GKT to provide access to relevant data based on their geographic connections, a basic framework for geographic referencing is required. This geographic framework would house the various geographic scales and should recognise the different methods for geotagging that exist in both PV's own data and the alternative data sources. As a starting point, a geographic scaling method for PV's existing data should be designed. This would incorporate geographic attributes already applied, as observed in their map based information system, and should consider referencing methods used in systems like the official record management system as well as other geographic divisions in place such as DSE's fire management structure. The system should combine official geographic names and PV geographic terminology if they differ. This would be the fixed or formal component of the framework with standards developed for georeferencing data with respect to geographic scales and names. All data should be georeferenced based on this framework and multiple geotags should be applied if applicable so that data are being presented at all relevant scales.

In addition to this formal part, the system should also comprise a dynamic component. This part would recognise alternative and informal georeferencing methods, and the range of variations with respect to formats and precision that is being applied to alternative data accessible via the Web. The dynamic component should subsequently link to the fixed framework by associating

these informal geographic references with an official geographic location at the applicable level. Part of the process is for the alternative data to be tagged and geotagged appropriately so they are available when needed. The dynamic aspect of the system should include the ability by all users to add personalised tags to both existing and new data. When new data is being made available via the tool, the formal georeferencing method needs to be applied, but informal geotags should be encouraged to enhance the user experience and effectiveness. This could be in the form of free format text, similar to the free format text that can be added to PV's official record management system, or it can be by attaching keywords like *Flickr* and *Google* applications allow users to do. Figure 1 illustrates the various considerations for the fixed and dynamic components of the basic geographic framework.

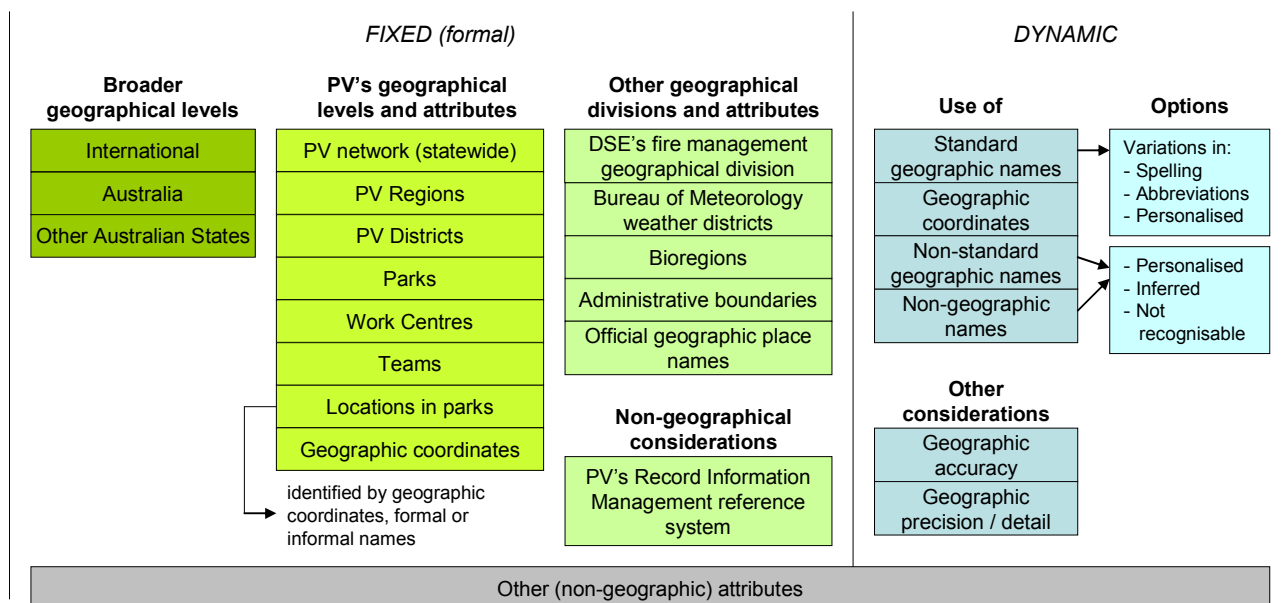


Figure 1: Basic geographic framework with fixed and dynamic components

It is suggested that the broadest PV level that encompasses the whole organisation should not be classed as 'statewide' as it may clash with the overarching geographical area of the State of Victoria. Geographically speaking, PV's network does not entail the full State of Victoria but only parts thereof. Its partner in fire management, DSE, could on the other hand use 'statewide' or 'Victoria' as a geographic reference. The various levels that comprise PV's geographical structure in Figure 1 are not necessarily complete but can be used as a starting point.

Devising a technical solution is outside the scope of the research project. How such a system would work in real life, how it can be developed or if technology exists that can do what is suggested is therefore not discussed. However, considering the growth of research into geographic information retrieval (Jones and Purves, 2008; Silva et al., 2006), no doubt a solution for building such a framework could be found should the GKT ever be built.

4 ACCURACY AND USEFULNESS OF ALTERNATIVE DATA SOURCES

With the arrival of Web 2.0 and UCI, the amount of available information may be larger but an immeasurable amount of it is also of poor quality (Keen, 2007; Shirky, 2008). This is because there is no system in place to filter "good from mediocre" (Shirky, 2008, p. 81), with many Web applications lacking quality control and leaving the editing and filtering in essence with users (Keen, 2007). Projects like Wikipedia (en.wikipedia.org) and OpenStreetMap ensure the quality of the data contributed by their users is sufficiently accurate for the system to work - the former through a community of 'Wikipedian' administrators (Weinberger, 2007) and the latter by providing guidelines and standards to which data collectors must adhere (Hudson-Smith et al., 2009).

A distinction also needs to be made between information that is of poor quality or merely silly, and content that is downright incorrect. Inaccurate geographic information from a park management perspective, for example, could result in burning an area of a park that houses

sensitive fauna species or slashing a rare orchid during weeding activities because they were thought to be located somewhere else.

Another point to consider is to view the UCI from the intended audience's perspective. Information put on the Web, although accessible to everyone, still has a specific target audience. So if users disagree with the usefulness or relevance of certain content, they are probably not the anticipated audience (Shirky, 2008). Considering what the original purpose and appropriate usage of the data are may make a difference as to their perceived usefulness.

As the GKT provides access to alternative data sources including UCI, these issues need to be acknowledged and addressed if possible. A theoretical methodology has therefore been developed that uses confidence ratings to assess the potential quality and usefulness of alternative data. When applied in practice, this will make the GKT more effective and assist users when considering whether to use the accessible data in their decision making.

4.1 A theoretical methodology for deriving confidence ratings

Assessing quality or accuracy and usefulness is not straight forward, and arguably a difficult task. What is this based on? Who decides whether information is 'quality' information or not? Although some data can clearly be accurate or inaccurate, there are also different levels of accuracy each of which may be acceptable for different purposes. Assessing usefulness is in part dependent on the user. What is the user's purpose or what are their intentions with the data? Resources – time, money and manpower – could of course be used to analyse the data and thus determine which data are 'acceptable' and which are not. High and low quality data could be filtered out in such a way, but for the data in-between, the data of medium quality, the distinction is not so clear as they could be useful for some purposes but not for others.

Because quality and usefulness are not straightforward to assess, the research project proposes that an appropriate technique is to provide indicators instead. These would show what quality can be expected or what the likely usefulness of the data will be. These indicators are not 'absolute' *per se* and their relevance can vary depending on how the data are used. Data quality may appear to be a more dominant issue, usefulness is equally considered in part because the two are linked. Furthermore, 'quality' within the realm of this methodology is related to characteristics such as 'accuracy' and 'credibility'. The term 'quality indicators' was therefore regarded to be too narrow, and the research project is developing a theoretical methodology for a scale of confidence or 'confidence ratings' instead in an attempt to encompass the multiple considerations.

A part of the conceptual GKT considers the classification of data – according to both existing conventional categories used by PV and user contributed keywords – to enable data access using these data attributes. Certain data attributes may not only help assess the usefulness of that data, some of the attributes may also affect or influence whether the data are regarded as accurate and trustworthy or not. Most of the data attributes described in this paper were identified during an investigation into existing digital data repositories and PV's data deemed relevant to the case study. These data attributes are being used as the basis for deriving the confidence ratings methodology. One thing to bear in mind is that the idea of 'confidence' can involve multiple concerns, such as confidence in the quality of the data, confidence in relevance to the topic and confidence in the accuracy of the geographic location. For the purpose of this research project 'confidence' is considered as a broad concept, however, the system will allow for personal annotations to be attached to data to describe specific concerns.

4.2 Data attributes

The initial three attributes that can assist in assessing the potential quality and usefulness of existing data are 1) the data source, 2) the data contributor and 3) the level of control from PV or WPNP in the data generation. To explain:

1. The data source - This is arguably the key indicator for a quality assessment; where do the data come from? Can the source be regarded as 'reputable' or not? Are providers experts in their field? Are they a professional organisation or it is a personal blog? Is the source likely to be biased? And if there is a potential for bias, can the data still be regarded as reputable because the provider is a professional or supports their view, or is it just someone's personal

opinion without any such credentials? Examples of potentially relevant data sources with seemingly varying 'quality' or 'usefulness' ratings are the International Union for Conservation of Nature, Melbourne's newspaper *The Age*, the Victorian National Parks Association (VNPA), a personal bushwalking blog, and an apparent standalone blog on the history of WPNP with hyperlinks to VNPA and the Friends of the Prom community volunteer organisation.

2. The data contributor – Although related to the data source, the difference is perhaps best explained by considering websites that gives access to data contributed by different people. For example, the ABC website (www.abc.net.au) on one hand comprises news articles written by professional journalists as well as opinion pieces - some written by authors with clear agendas that are linked to or supported by certain organisations and whose credentials arguably vary. Additionally, there are reader comments, written by the general public. They are arguably the 'non-experts' who can add comments to existing stories, although, like the opinion writers, some of them can contribute valid and accurate points and may have hidden credentials (or none at all). If we take the archives of the State Library of Victoria (SLoV) and the National Library of Australia (NLA) on the other hand, these in essence give access to data from others and it is perhaps the origin of these original data creators that should be considered rather than the reputable SLoV itself who is more of an access enabler here. The data contributor, like the data source, can refer to organisations, groups or individuals and their expertise or reputation, ranging from expert and professional to the non-expert general public.
3. Level of control from PV or WPNP involved in data generation - This particularly applies to newly generated UCI that is specifically asked for versus any UCI. For example, if WPNP involved volunteers to monitor specific wildlife (as seen in a special project between the Conservation Volunteers organisation and WPNP for instance (Conservation Volunteers, 2010)), there is a high likelihood that they will obtain useful data at a desired quality. This is because there is a high level of control involved that asks the participants for specific details, visit specific areas, look for specific species, and record specific observations. Alternatively, encouraging park visitors to contribute to PV's website using participatory tools without detailed instructions may result in UCI having varying levels of both quality and usefulness. Furthermore, data analysis will most likely be required to ascertain the UCI's potential level of quality and usefulness.

Additional attributes that have been identified as having the potential to assist with assessing data quality and usefulness are:

4. Data access - Can the data be viewed online in its complete form or are they merely digital records of the data's existence? Would accessing actual documents require lodging a request and going to a particular repository or location? It may be necessary to register in order to see complete documents and if so, is this a free and relatively simple exercise or is a paid subscription required?
5. Search options - How good are the search facilities of the site? Does it give advanced or detailed search options? Do they give the required results or is it a matter of sifting through many search results? And are the data available in their full extent online, or are they merely records providing limited information only?
6. Level of detail - Is the information generalised and on a broad scale or does it provide precise details? This also relates to the data source, as detailed data from an expert source could be regarded as useful and accurate, whereas detailed data that are more likely to be biased would be less so.
7. Data purpose - Is the information factual or is it opinion? Do the data provide background information or are they meant to be combined with other data forming new information?
8. Year of data creation - How old are the data? If data were created some time ago, are they still accurate or have they been superseded by newer versions? Can the information provide a historic perspective or is it not relevant any more?

4.3 Classifying data attributes

Considering the primary indicator, the data source, deciding which data providers are reputable or biased is potentially a subjective issue in itself. An organisation such as PV may have views about what sources are considered reputable or credible and which are not. Usefulness is

similarly a subjective matter. Apart from the purpose for using the data, user attributes like available time and resources can make the difference between some users being able to visit a library or searching through many results whereas others can not. Hence the attributes listed in section 4.2 will be viewed differently by those users

One way to deal with this subjectivity is to apply general guidelines or indicators. This can be achieved by classifying and subsequently ranking the data attributes. Data sources for example can be classified using terms like media organisations, government and natural environment organisations, academic or educational sites, and not for profit organisations. These immediately give an indication to the user as to the perceived usefulness or potential quality. Ranking these with regards to reputability or reliability however is again less clear cut. For instance, has information from a small local newspaper owned by a multinational corporation the same 'quality' or 'objectivity' as the information from a major national broadcast corporation? What about a blogger with links to a well-known volunteer conservation organisation, a conservationist blogger without such connections or a blogger who used to work for a reputable newspaper? Consequently, developing a formula to establish a confidence rating system using weighted expressions to differentiate between the relevance of some of the attributes, such as the data source, is regarded as difficult. This is because in addition to ranking the different attributes, a hierarchy also exists within some attributes, which in turn can vary depending on the user purpose – in other words there are multiple sets of variables.

To clarify in more detail, Table 1 shows the same data attributes previously listed with possible categories for each attribute. Some examples are provided for clarification purposes and it should be noted that the order of these categories does not imply a ranking *per se*.

Table 1: Data attributes and examples of potential categories within those attributes for developing a confidence rating system.

1. Level of PV control	<ul style="list-style-type: none"> - PV project or site where users contribute specific data (for example a special crowdsourcing project organised by WPNP asking visitors to collect specific data); - PV project or site collecting any user generated data (for example general public contributing photos or comments to PV's website using participatory tools); - Non-PV project or site where users contribute specific data (for example OpenStreetMap where users contribute geographic data); - Non-PV project or site where users contribute any data (for example user comments found in response to an opinion piece on the ABC website about prescribed burning and ecological fire management);
2. Data source	<ul style="list-style-type: none"> - Government organisations; - Media organisations; - Academic or educational organisations (for example Directory of Open Access Journals - http://www.doaj.org); - Not for profit / volunteer organisations (for example Victoria Naturally Alliance - http://victorianaturally.org.au/index.php); - Businesses (for example Australian Human Resources Institute Centre of Excellence - http://www.ahri.com.au); - Special interest / community groups (for example Atlas of Living Australia - http://www.ala.org.au); - Social media sites with associations or credentials (for example blog written by Andrew McAfee, the author who coined the term Enterprise 2.0 – http://andrewmcafee.org); - Social media sites without obvious associations or credentials (for example <i>Flickr</i>);
3. Data contributor	<ul style="list-style-type: none"> - Expert (for example Bushfire Cooperative Research Centre or its researchers - www.bushfirecrc.com); - Professional (for example Melbourne newspaper <i>The Age</i> or journalists writing for the paper - www.theage.com.au); - Commercial (for example O'Reilly Media, the online media company often associated with the term Web 2.0, or its employees – www.oreilly.com); - Volunteer (for example contributors to the Atlas of Living Australia – www.ala.org.au); - Amateur or non-expert (for example a blog written by a bushwalking enthusiast, or reader comments on the ABC's website);

4. Data access
<ul style="list-style-type: none"> - Full data available on line, and no need to join up to access; - Partly available online but need to join for full access – free and easy to join; - Partly available online but need to join for full access – restrictions on joining; - Online record providing limited details only – data are available offline in full and for free; - Online record providing limited details only – data are available offline with restrictions; - Online record providing limited details only – data are not available offline;
5. Search Options
<ul style="list-style-type: none"> - Advanced search options; - Search within results; - Multiple keyword search; - Keyword search; - Limited search options, and need to manually search through information or Web pages;
6. Level of detail
<ul style="list-style-type: none"> - Ranging from broad to in-depth detail;
7. Data purpose
<ul style="list-style-type: none"> - Factual; - Opinion; - Background; - Input data (for example to be used for further analysis, or as input for tools or models);
8. Year of creation
<ul style="list-style-type: none"> - A series of time frames can be devised suitable or relevant to PV.

There are many variables that can change in importance for different people according to the different purposes. Turning these into one ranking system would require either discounting issues or else result in many alternatives of the 'if..., else...' variety, both of which could render the system impractical. Developing such a confidence ratings methodology in practice is therefore a complex task that requires time and input from organisations, as well as expert knowledge to assist with the multivariate algorithms. An interactive model, where users tick which attributes are most important before a rating is provided may be a possible option. However, this would be less practical as it would require user input and thought and thus make the system, and tool, less effective. It would also not necessarily be less complicated because of the variations within each attributes. For example, not all blogs or media websites are of the same quality and the videosharing site *YouTube* hosts videos that are of dubious quality alongside videos that are useful and of high quality.

4.4 Two proposed solutions

It is clear that the quality and usefulness of data are difficult to assess objectively and transform into a single 'rating' that can be attached to individual data. Two possible solutions are proposed to assist with this.

Firstly, take the notion posed in section 4.1 that data attributes are indicators of data quality. One option is to attach these attributes to the data as tags and leave it up to users to decide what kind of data suits their purpose and where they want to look first. Someone writing a background story may need data that are old, the author of a policy paper may need to review community opinions, whereas the production of a research paper may be aided by information obtained from governmental, educational or academic sources. Although a formal ranking originating from an authoritative source (in this case PV) is useful as a guide, it ultimately depends on individual users' opinions as to which data are suitable. The users in turn should also be able to add personal suggestions or observations about the quality and usefulness of the data as they are using it. This solution at first would not be particularly effective from a users' perspective. Apart from a range of keywords that organises the data to assist their search, users would have to virtually make all assessments without any guide. Once the system is in place however and users have started to add suggestions and ratings, it would gradually turn into a more user friendly and effective system.

The second option is to look at individual attributes only, and define a rating system for each, where possible. Again, the onus would be on the user to decide which attribute is important to them and undertake a search accordingly. So rather than combining all attributes, resulting in one confidence rating system, a series of ratings can be attached for individual attributes. Figure 2 is a

sample matrix showing potential confidence indicators for the four classifications under the ‘level of PV control’ attribute.

Confidence indicators:	<i>PV project – specific UCI</i>	<i>PV project – any UCI</i>	<i>Non-PV project – specific UCI</i>	<i>Non-PV project – any UCI</i>
High (5)				
High to medium (4)				
Medium (3)				
Medium to low (2)				
Low (2)				

Figure 2: Potential confidence indicators for four categories of the ‘level of PV control’ attribute.

There are several issues that arise out of this approach. Firstly, it would be beneficial to reduce the number of attributes or else prioritise them to make the system more workable. The first three attributes of Table 1, data source, data contributor and level of PV control in the data creation, are regarded as the key attributes that influence data quality and usefulness. Attributes 4 and 5, data access and search options, can be rated relatively easily if focusing on usefulness – fully and freely accessible data would rate higher than limited and restricted data access for example. The remaining three attributes 6, 7 and 8 are perhaps less able to be rated with respect to quality or usefulness as it depends on the user as to the level of detail they require, what kind of data they need, and how old or recent the data should be. The classification of these three attributes could therefore be attached as tags and would not be considered for the quality and usefulness assessment. Figure 3 shows the resulting model: the first five data attributes are each divided into hypothetical classes with each class assessed and provided with a confidence rating. The user can select what confidence rating they require for different attributes depending on what they think is important, and further refine the search by tags that include the remaining three attributes.

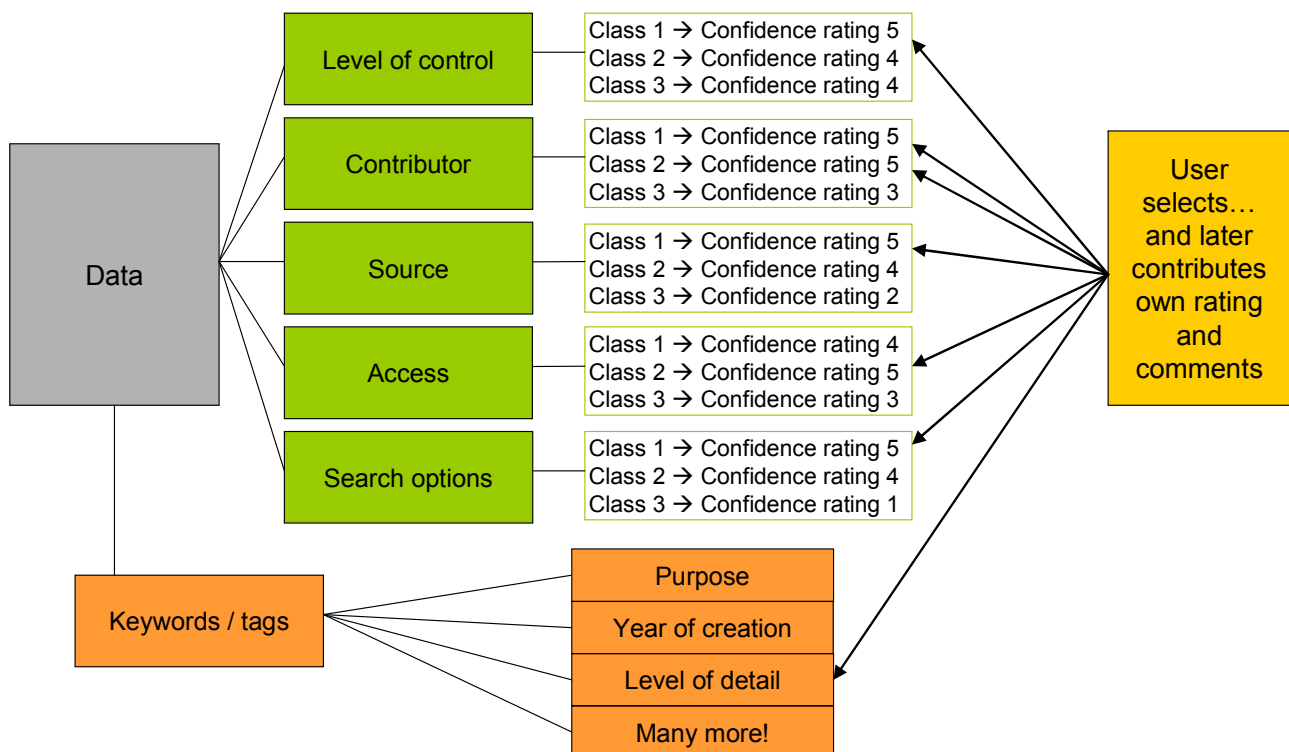


Figure 3: Solution 2 – rating hypothetical classes of key data attributes and attaching remainder as tags.

The second issue is the aforementioned requirement for a hierarchy or formal rating defined by PV, which would require input and time. A PV representative would have to be involved in the creation of such as rating and decide which data sources were considered to be better or more trustworthy and which ones were less so. This would have to be done for all attributes. But even these ratings should be used as a guide only as to what quality and usefulness can be expected of the source rather than becoming an absolute measure. To further enhance the system, users of the data should be able to rate the data and attach tags or comments. As with the first solution, the system would become more effective once users contributed to the system as they were using it.

The end result is therefore regarded to be an evolving assessment model; as the data or site are being used, users can give their personal confidence ratings based on quality or usefulness of the data for other users to consider. This is in line with the aforementioned personalisation of information and the bottom-up approach associated with Web 2.0. It is feasible that data will end up being merely tagged as 'additional data'. This would mean the system will not be as effective as it could be, as it would be up to each new user to assess the data for themselves without guidance. However, the data would still be on hand if required and easily available for users interested in further information. In the end, even without a clear confidence rating system in place, classifying the data and attaching a variety of attributes to data will already make a range of information more readily available to users to assist them in assessing the data's suitability for their needs, thus, in essence, already making use of the GKT more effective.

Because the confidence ratings methodology is only theoretical, it will not be a working component of the demonstration prototype. However, the prototype will illustrate that such a methodology is being developed, and also has ratings attached to the limited data to which it provides access. These ratings are 'top down' ratings to be supplied by a WPNP staff member associated with the research project. It is envisaged that the addition of a slider bar would allow users to set a minimum confidence rating and subsequently only data having those ratings would be presented. Figure 4 shows a screenshot of a portion of the demonstration prototype currently being designed that incorporates the hypothetical confidence ratings. Figure 4 also shows that the current design uses a classification of the data source attribute condensed to social media, media, government, research and organisations.

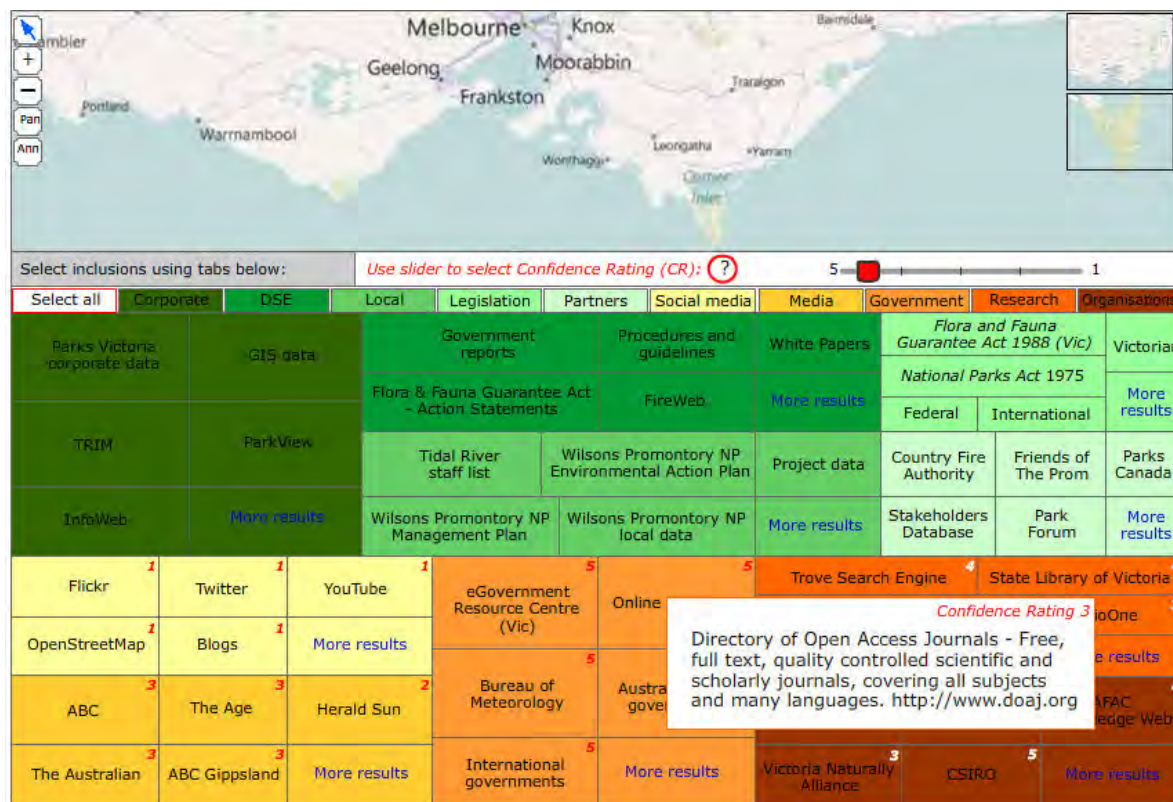


Figure 4: Portion of demonstration prototype currently being designed showing a sliding bar for users to set confidence ratings.

5 CONCLUSION

The essence of the proposed GKT is to enable data access using geographic attributes, which means that data must be georeferenced to facilitate the access. Despite an inherent geographic structure in place, appropriate geographic attributes are not always applied to PV's data. Georeferenced UCI can display an inconsistency in geotagging and a variation in depth and accuracy of the geographic information. The proposed basic geographic framework comprises a formal component that can accommodate PV's data based on a geographic structure that uses existing geographical divisions. An informal component handles non-traditional data and enables linking this informal geographic information to the formal geographic structure in place.

With regards to quality and usefulness of alternative data, the proposed methodology uses a variety of data attributes to derive confidence ratings that can assist in assessing data quality and usefulness. Because the importance of the attributes can vary depending on user opinions, they can not be ranked and rated easily to produce a single confidence rating. The proposed solutions address this by suggesting that attributes be rated individually so that users can decide which attributes are most suitable to them and rank them accordingly. An evolving system that lets users add ratings and information as they are using data means that the system should become more effective in time.

In order for the GKT to work, it needs to be accepted that UCI generally takes a less formal and less precise approach to georeferencing. The Web 2.0 notion that UCI can enhance existing data and thus that data from informal sources can be valuable also needs to be recognised, whilst the differences in quality and usefulness need to be understood and dealt with. The solutions provided in this paper are aimed at addressing these issues.

This research project proposes to amalgamate traditional data with data not traditionally relied upon - building on the suggestion that UCI can enhance existing data - to assess if these alternative data can benefit an existing park management data archive. However, the issues identified are not unique to PV or the park management sector. This means that the theoretical methodologies put forward in this paper could be applied to other organisations or to other fields, provided there is an inherent geographical structure. Information management and knowledge management can deal with the basic premise of this research, that existing data is not used as

effectively as it potentially could, whereas research in the field of Geographic Information Retrieval focuses on geospatial information specifically. This research project fits within these realms and the results will add new ideas than can be built upon.

The methodology being developed for this research differs from PV's traditional approach to data collection because of the inclusion of alternative data sources and applying concepts of Web 2.0. The idea that involving users – clients, visitors, staff – through participation and collaboration can potentially be useful is growing, and Web 2.0 concepts are increasingly being used by organisations including libraries (e.g. Byrne, 2008), education facilities (e.g. Trinidad and Broadley, 2008) and others under the term Enterprise 2.0 (McAfee, 2006). PV itself has utilised Web 2.0 tools for the drafting of two park management plan (see www.weplan.parks.vic.gov.au for details), and has recently created a presence for itself on four popular social media sites (*Facebook*, *Twitter*, *YouTube* and *Flickr*). It is still relatively early stages of the adoption of social media into organisations with varying uptake in Australia (Williams, 2011). This research subsequently can assist in providing further insight into the opportunities and benefits that a participatory approach may bring.

And finally, it is argued that if the amalgamation of traditional and non-traditional data can result in an effective GKT, this should benefit all users of the tool and, in this case, assist park management and thus ultimately users of those parks through, for example, enhanced visitor services. When adopted by other organisations or different fields, a similar flow of benefits should apply to organisations and users alike.

6 ACKNOWLEDGEMENTS

The PhD research project described in this paper falls under the Affective Atlas Project being undertaken under the umbrella of the Geo-Placed Knowledge Research Initiative, part of RMIT's Design Research Institute. The research project is supported in part by an Australian Research Council linkage grant (LP0883291) and an RMIT postgraduate award.

Parks Victoria is a collaborator on both the research project described in this paper and the overarching Affective Atlas project, and provides both financial and in-kind support including use of data, facilities and the expert knowledge and assistance of its staff members.

REFERENCES

Amitay, E., Har'El, N., Sivan, R., Soffer, A. (2004). Web-a-where: geotagging web content. Proceedings of 27th annual international ACM SIGIR conference on Research and development in information retrieval, Sheffield, United Kingdom, 25-29 July 2004, 273-280.

Anderson, P. (2007). What is Web 2.0? Ideas, technologies and implications for education: JISC Technologies & Standards Watch.

Bishr, M., Mantelas, L. (2008). A trust and reputation model for filtering and classifying knowledge about urban growth. *GeoJournal*, 72, 3, 229-237.

Brown, J. S., Duguid, P. (2000). *The Social Life of Information*. Harvard Business School Press, Boston.

Byrne, A. (2008). Web 2.0 strategy in libraries and information services. *The Australian Library Journal*, 57, 4, 365-376.

Conservation Volunteers (2010). Wilsons Promontory National Park Wildlife Monitoring. <http://www.conservationvolunteers.com.au/volunteer/WilsonsPromontory.htm>. Accessed 29 July 2011.

Das, T., Kraak, M. J. (2011). Does Neography Need Designed Maps? Proceedings of 25th International Cartographic Conference, Paris, France, 3-8 July 2011.

- Dykes, J. A., Mountain, D. M. (2003). Seeking structure in records of spatio-temporal behaviour: Visualization issues, efforts and applications. *Computational Statistics and Data Analysis*, 43, 4, 581-603.
- Elsley, M., Cartwright, W. (2011). Contemporary and Collaborative Web Concepts as part of a Geo-Knowledge Tool to Assist Park Management. In Ruas, A. (Ed) *Advances in Cartography and GIScience*, 1, Springer, Berlin, Heidelberg, 261-277.
- Goodchild, M. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69, 4, 211-221.
- Gravois, J. (2010). The Agnostic Cartographer [Electronic Version]. *Washington Monthly*, July/August. <http://www.washingtonmonthly.com/features/2010/1007.gravois.html>. Accessed 04 April 2011.
- Hudson-Smith, A., Batty, M., Crooks, A., Milton, R. (2009). Mapping for the Masses - Accessing Web 2.0 Through Crowdsourcing. *Social Science Computer Review*, 27, 524-538.
- Jones, C. B., Purves, R. S. (2008). Geographical information retrieval. *International Journal of Geographical Information Science*, 22, 3, 219-228.
- Kamel Boulos, M. N., Wheeler, S. (2007). The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. *Health Information & Libraries Journal*, 24, 1, 2-23.
- Keen, A. (2007). *The Cult of the Amateur - How Today's Internet is Killing Our Culture and Assaulting Our Economy*. Nicholas Brealey Publishing, London.
- Lackie, R. J., Terrio, R. D. (2007). Mashups and other new or improved collaborative social software tools. *Multimedia and Internet@Schools*, 14, 4, 12-16.
- Luo, X., Xu, H., Song, M., Song, J. (2008). Research on SOA-based platform to enable mobile mashups. *Proceedings of 11th International Conference on Communication Technology, ICCT*, 607-610.
- McAfee, A. (2006). *Enterprise 2.0: The Dawn of Emergent Collaboration*. MIT Sloan Management Review, 47, 3, 21.
- O'Reilly, T. (2005). What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1>. Accessed 24 March 2009.
- Shirky, C. (2006). Ontology is Overrated: Categories, Links, and Tags. *Clay Shirky's Writings About the Internet*. http://www.shirky.com/writings/ontology_overrated.html. Accessed 22 May 2009.
- Shirky, C. (2008). *Here Comes Everybody: The Power of Organizing without Organizations*. The Penguin Press, New York.
- Silva, M. J., Martins, B., Chaves, M., Afonso, A. P., Cardoso, N. (2006). Adding geographic scopes to web resources. *Computers, Environment and Urban Systems*, 30, 4, 378-399.
- Surowiecki, J. (2005). *The wisdom of crowds: why the many are smarter than the few*. Abacus, London.
- Tapscott, D., Williams, A. D. (2008). *Wikinomics*. Penguin Group, New York.

- Trinidad, S., Broadley, T. (2008). Using Web2.0 applications to close the digital divide in Western Australia [online]. *Education in Rural Australia*, 18, 1, 3-11.
- Turner, A., Forrest, B. (2008). *Where 2.0: The State of the Geospatial Web*: O'Reilly Radar.
- Vickery, G., Wunsch-Vincent, S. (2007). *Participative Web and User-created Content: Web 2.0, Wikis and Social Networking*. Organisation for Economic Co-operation and Development Publishing, [Paris].
- Weinberger, D. (2007). *Everything is Miscellaneous*. Times Books, New York.
- Williams, P. (2011, Monday 01 August). *Head in the cloud*. *The Age*, 15.

Understanding the relationship between spatial information, property markets and macroeconomic policy

Nilofer Tambuwala, Rohan Bennett, Abbas Rajabifard, Ian Williamson

Centre for Spatial Data Infrastructures and Land Administration, Department of Infrastructure Engineering, University of Melbourne
Parkville, VIC 3010
nilofert@unimelb.edu.au, rohanb@unimelb.edu.au, abbas.r@unimelb.edu.au, ianpw@unimelb.edu.au

ABSTRACT

An important relationship exists between property markets, macroeconomic policy and spatial information. This relationship is often constrained by institutions, policy and the lack of technology. Increasingly intergovernmental cooperation and advancement in spatial technologies enables spatial enablement, offering the potential to include authoritative spatial information generated by a country's land administration system, in government macroeconomic decisions. A review of literature shows both disciplines of land administration and macroeconomics advocating the role of secure property rights as key to economic progress. However the link between the agencies that support secure market transactions and generate land information, and those that use the information to manage the economy is not made. As a result, suboptimal governmental fiscal and monetary decisions can occur.

Two case studies of the Australian context, with a focus on managing land taxes and administering interest on debt financing, illustrate this. Results of the case studies are used to derive a simplified empirical model that aims to articulate and promote the role of government spatial information in policy and decision-making. Practical implementation will require determination of the legal, institutional, and technical requirements of the model.

KEYWORDS: spatial information, property markets, macroeconomic policy, land administration

1 INTRODUCTION

Spatial information is increasingly important for evidence-based macroeconomic policy. Information about land, one of three factors of production in classical economic theory, is collected and maintained by land administration agencies. In many federated countries economic management and land administrative responsibilities are divided between different levels of government. Both disciplines operate independently and their interdependence is not reflected in literature.

In the United States for instance, the financial collapse of the late 2000s that resulted in the global financial crisis (GFC) is suggested to have emanated in part from a lack of integration between the information processes of land administration and macroeconomic policy making (Roberge and Kjellson, 2009 and Buhler and Cowen, 2010).

In Australia, states and territories run successful Torrens land registration systems that generate secure titles to parcels. The transferability of rights in land underpins an active and secure land market that plays a key role in the country's economic progress. Australia's spatial information systems that record information about transactions in land are supplied

by independent land administration agencies. The link between these agencies and those that are responsible for economic management is unclear.

The capacity of spatial information to support macroeconomic policy making is in need of new evaluation. To achieve this, existing theory and practices within the disciplines of macroeconomics and land administration are presented. The short comings of these theories in practice are demonstrated by presenting the results of two case studies from the Australian context. The outcomes of the case studies lead to the articulation of a simplified empirical model that links macroeconomics, land administration and spatial information about land market transactions. The model is aimed at policy and decision makers in higher levels of national government. It aims to emphasize the necessity for seamless land administration and macroeconomic processes.

2 BACKGROUND

According to classical economic theory, there are three means of generating wealth in an economy – capital, labour and land (Dale and McLaughlin, 1999). Transactions in land form a significant part of the GDP of any market economy. Formal land transactions are only possible through the existence of land administration infrastructures that allow for private land ownership to be registered, land values to be established, and rights in land to be exchanged in a market environment. As such, these administrative structures have a critical impact on the national economy as a whole.

The relationship between economic policy and its impact on local and national land or property markets has been well documented (McKenzie and Betts, 2006; Poterba, 1992 and Joshi, 2006). However, the role that spatial information about property market transactions, should play in economic policy making is not well understood. Central governments in most market economies employ principles from institutional and Keynesian economics to support government management of aggregate demand in the economy. Government macroeconomic policies are intended to combat the instabilities that a pure market structure may cause. However, the land or property market is formalised by land administration systems whose importance in macroeconomic policy is not yet clearly understood. This is particularly significant in federated nations, such as Australia, where land administration and the collection of spatial information about transactions in land is the constitutional authority of the state and territory governments, while macroeconomic policies are administered at federal level.

A large part of macroeconomics is about land. For instance, the availability of money in the economy, especially unrestricted mortgage financing, managed by monetary policy tools, has a significant impact on land or property markets. Operations in the secondary mortgage market also influence purchasing power (Carper et al., 2007). For this reason, analysis of how property markets work is predominantly carried out in the language of economics. However in a land or property market the product is formalised by a land administration system that determines the type of title and land right through the tenure system, and other functions of land value, land use and land development (Enemark, 2007). The land administration system identifies the complex real property right that is the foundation of formal property markets. It also collects and maintains authoritative spatial information about ownership and sale. These core land administration functions are not necessarily reflected in economic literature that deals with property markets.

Property or real estate economics endeavours to account for the unique characteristics of property markets, by linking the actions of people to their effect on the value of property. This discipline tries to apply general economic theory to the realities of real estate practices. Interactions within the property market are founded in economic theory however; property markets come from a nation's capacity to create 'property rights institutions' (North and Thomas, 1973). These are often within the domain of land administration. This interaction between the two disciplines is not well reflected in literature.

Economic theory does recognise the role of land administration systems in national wealth creation. For instance North and Thomas (1973), Deininger and Binswanger (1999) and DeSoto (2000) argued for the importance of the capacity of a nation to build land and property rights. Some land administration theory too promotes the importance of this discipline to the creation of wealth. Wallace and Williamson (2006) suggested that in a property system the rights are the commodities, not just the land itself, and successful land markets derive strength from creating and marketing land rights and complex commodities. Williamson et al. (2010) and Wallace et al. (2010) convey the importance of land administration processes that are influenced by national land policy and economic systems. They emphasise that infrastructures that manage land data should allow access, interoperability and multipurpose use of that data.

However much of this land administration literature deals with developing nations, with a focus on establishing land administration systems and recording land parcels, for the purposes of basic taxation and the construction of formal land markets. In most federated countries the land administration systems and consequently the property markets are already well established, relying on a complex set of interrelated institutions, formal and informal, to promote information flow. In Australia for instance, land registration and property valuations department forms part of the state governments. Often property valuation methods can be as varied as the property laws in various jurisdictions. For instance, Queensland, New South Wales and some Victorian councils value and tax unimproved land. Other states tax land and buildings based on capital improved values. Valuations data is collected either by in-house or private valuers hired by the state or local governments. Data relating to ownership and value of properties is generally stored in multiple jurisdiction-based databases.

Consequently, the land or property market is impeded by information asymmetries (Garmaise and Moskowitz, 2004). One function of market-supporting institutions is to ensure that information flows smoothly (McMillan, 2003). Of this, spatial information about land transactions flowing up to the central government is essential in the management of national wealth. However there is a lack of understanding at higher levels of government about the role that spatial information collected and maintained by established state-based land administration systems should play in macroeconomic policy making.

The short comings of economic and land administration theory in supporting the link between property markets, spatial information and macroeconomic policy is evident within the literature. The following case studies further explore this link in practice, by looking at two macroeconomic processes, taxation of land and setting of interest rates, in the context of property markets in Australia. The information flows between these functions and the land administration processes of tenure and value that directly underpin property markets (Enemark, 2007) are studied, for this federated country. The property object approach, introduced by Bennett et al. (2008), is used to distinguish each tax or complex commodity derived from a land parcel as a separate attribute of the land. This permits an in-depth analysis of the information flows to the levying authority of each instrument or attribute of the land parcel. The case study approach gives a clearer understanding of the spatial information needs of federal policy makers in a federated market economy.

3 RESULTS OF CASE STUDY 1: MONETARY POLICY IN AUSTRALIA – INTEREST RATES

Macroeconomic policy tools such as taxation and the setting of interest rates assist in the management of national wealth. Monetary policy is generally controlled by a country's central bank. Monetary policy decisions require authoritative information about transactions in capital, labour and land to effectively judge the status of the economy. Information about the land market is collected and maintained by land administration agencies.

Official interest rates, set by the Reserve Bank of Australia (RBA), depend upon how the economy is functioning at a certain time. Monetary policy decisions are expressed in terms of a target for the cash rate, which is the benchmark overnight rate for bank lending. Most banks charge a separate, slightly higher rate for debt financing.

The property or land market in Australia plays a significant role in the national economy. Of this, the housing market is the largest (West, 2010). Due to the size of the investment, most property purchases require debt financing. Kohler and Rossiter, (2005) found that an important consideration for property ownership is the ability to make financial commitments towards purchasing property and to meet any repayment obligations if a loan is taken out to purchase the property. Since the interest rates set by banks generally follow the official RBA cash rate, changes to the cash rate affect affordability and hence investment in the property market. This consequently impacts on supply and demand. Since the land or property market, especially the housing market, contributes significantly to national wealth in Australia, timely and accurate spatial information about property market transactions is essential to macroeconomic policy decisions that aim to effectively manage national wealth. In Australia authoritative spatial information about the property market, such as ownership, value and property sale information is maintained by the state and territory land administration agencies. There is no central database that records property and prices (West, 2010).

The situation in Australia with regard to the access to authoritative land market information by monetary policy makers is illustrated by Figure 1.

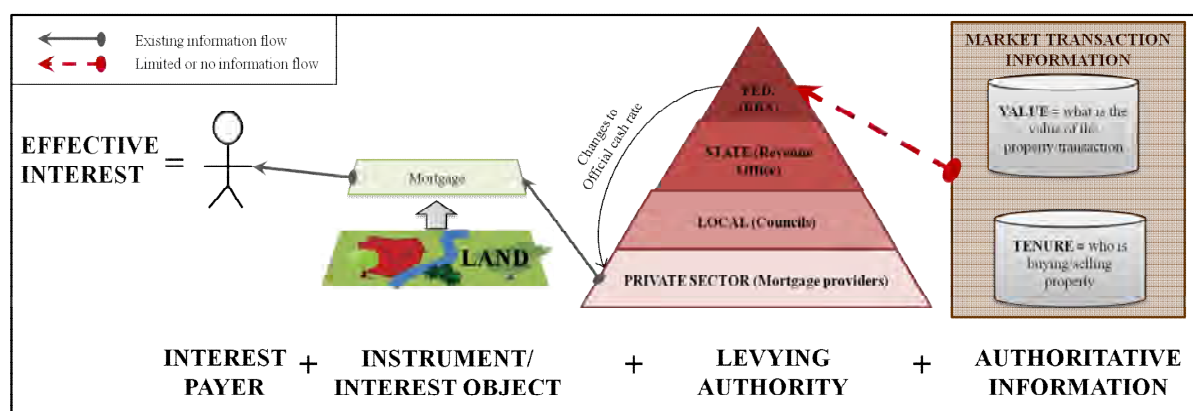


Figure 1: The relationship between spatial information about market transactions and interest rates in Australia

Using the property object approach mentioned earlier, effective interest requires an interest payer, an instrument or interest object, a levying authority and authoritative information. In the figure above, the interest object attached to the land is a mortgage levied by a mortgage provider. The mortgage provider adjusts its interest rate on debt financing to closely follow the Reserve Bank's official cash rate. Information flows between these three entities, the individual, mortgage provider and RBA, are already established.

Currently in Australia, though spatial information about market transactions is eventually available at federal government level, it is held in separate state databases. As Figure 1 shows, there is currently limited or no access to these authoritative data stores at federal level. Connectivity needs to be established between the government creators and users of market information.

Authoritative implies publically sourced, timely and accurate data. As the RBA pointed out, 'data timeliness' is a major problem with access to housing price data (RBA, 2004 and 2005). This is attributed to the lack of consistency in transaction reporting requirements between the states. In most states, there is an absence of reporting requirement at the time of sale. Reporting generally occurs when documents are lodged for registration which is after settlement and can be up to three months after the contract of sale. From the perspective of

efficient economic policy, it is desirable for market analysis on house price data to be based on the period in which the price was determined, rather than when the transaction was later settled (RBA, 2005). Due to insufficient and untimely information flows and poor data integration at a national level, the RBA collects sale and transaction data from the private sector. For instance the RBA collects information about the commercial property sector, including vacancy rates, property prices and rents from the Australian Bureau of Statistics and other organisations such as Jones Lang LaSalle, the Property Council of Australia and Savills Research (RBA, 2009). For sales transactions, the Bank also relies primarily on the Australian Property Monitors (APM). Private sector sourced information is neither authoritative nor assured.

In summary, the flow of accurate, authoritative and assured land information from state government to macroeconomic policy decision makers is impeded resulting in policy decisions being based on less than optimal datasets. This may ultimately result in fiscal policies being potentially out of kilter with the fiscal reality of the jurisdiction.

4 RESULTS OF CASE STUDY 2: FISCAL POLICY IN AUSTRALIA – LAND TAXATION

The setting of tax rates form part of a nation's fiscal policy that manage national wealth. Higher transaction taxes may cause otherwise affordable transactions, to become unaffordable. Similarly, higher taxation on the holding of property increases the cost of ownership and consequently increases the incentive to sell. These taxes in turn affect supply and demand in the market place.

In a market economy, anything that is tradable or disposable is taxable. Taxation involves taking processes that have value and extracting part of that value for government. The ownership and sale of property make up the main processes within a land market. In the context of this study, we can define taxes on land to be *government charges on the transactions and holdings of property that form part of a land market* (derived from FIG, 1995 and CCH Editors, 2010) As with any other form of tax, taxes on land are compulsory contributions levied by the state on a taxpayer (individual or legal entity). The main taxes on land in Australia are: Capital Gains Tax (CGT), stamp duty, land tax, Goods and Service Tax (GST) and rates.

2009-10 values from the Australian Bureau of Statistics indicate that GST accounts for about 27% of total tax revenue for all levels of government. Taxes on property make up about 10% of total tax revenue. Of this, taxes on immovable property account for 6% and taxes on financials and capital transactions about 4%. Taxes on property were the largest source of taxation revenue (37%) for state governments in the same time period and were also the sole source of income for local governments (ABS, 2011).

In Australia, property is taxed at different levels of government. According to the earlier definition of a tax on land, an effective taxation system needs a taxpayer, an instrument being taxed, a levying authority and authoritative information about the taxpayer and the value of the instrument being taxed. For effective fiscal policy concerning land, the governments at each jurisdictional level need authoritative information about the ownership and value of property: Figure 2.

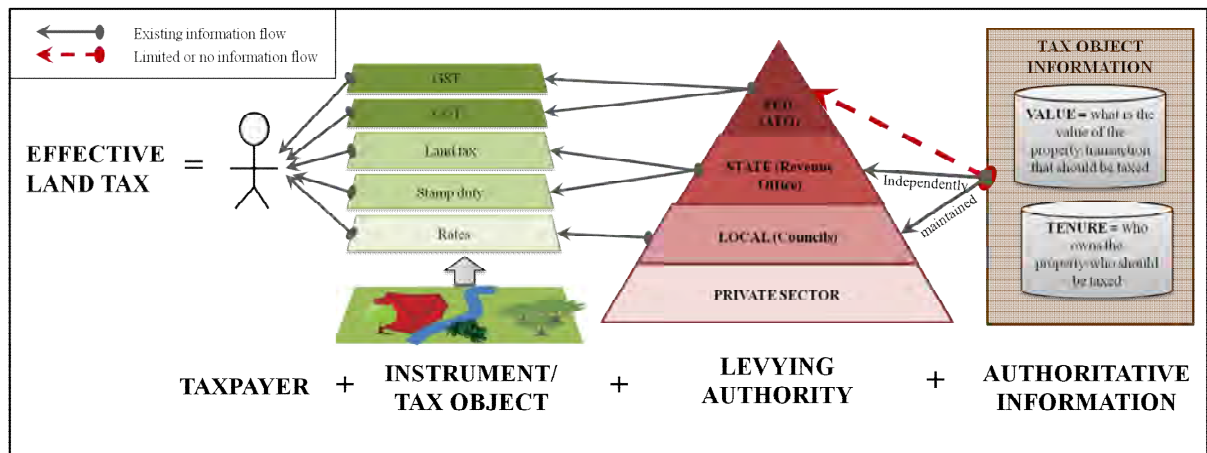


Figure 2: The relationship between spatial information about taxable transactions and taxation in Australia's land market

The figure above shows five tax objects levied by the various governments in Australia. Within the respective state governments, the revenue offices collect land tax and stamp duty, while the land administration agencies maintain the data stores of spatial information about the land. In Australia, though publically sourced spatial information about property ownership and value is available to the state taxation offices and local councils, it is generally maintained in separate databases by independent levying authorities. Duplication is evident. Additionally, as the dotted line in Figure 2 shows, the Australian Taxation Office (ATO) has limited or, in some cases, no access to the authoritative data stores of tenure and value information. It relies on information declared in tax returns and on data purchased from the private sector in order to collect capital gains tax and GST on real property. This brings the reliability and accuracy of this information into question.

There is clearly information asymmetry in operation here. This relates to the gap between information available within land administration agencies and what is actually shared with the Australian Taxation Office. The problems associated with information asymmetries within property markets have been well studied (Clapp et al., 1995, Dolde and Tirtiroglu, 1997, Milgrom and Stokey, 1982 and Garmaise and Moskowitz, 2004). However, much of this literature deals with horizontal information asymmetries between agents, brokers, buyers and sellers, or between neighbourhoods or over time. However literature does not adequately account for the problems associated with information asymmetries between different levels of government within a federated economy, in the context of managing national wealth.

What policy makers need is national spatial information about taxable objects related to land in order to meet broad policy needs such as assessing tax revenue capacities and meeting economic productivity challenges. This includes land tenure information, particularly ownership of taxable properties, and the value of the property or transaction to be taxed. For better management of national wealth in Australia, the large federal departments and agencies such as the ATO and RBA need authoritative, publically sourced property market information, collected and maintained by state land administration agencies.

The following section presents a simple, empirical model that links spatial information, land administration processes and macroeconomic functions. Triangulated (c.f. Golafshani, 2003) from the results of the case studies, the model is aimed at policy makers in higher levels of government, to initiate a better understanding of the need for seamlessness between the two disciplines, and a holistic spatial information management approach for a better informed government.

5 THE PROPERTY MARKET TREE

The need for Evidence-based Policy (EBP) is gaining a strong hold, particularly in Australia:

“The primary goal is to improve the reliability of advice concerning the efficiency and effectiveness of policy settings and possible alternatives” – Head, 2009.

Good data or ‘high-quality information bases’ is one key component of this (Head, 2009). To achieve this within a property market context requires links between the institutions that manage national wealth and those that support land market transactions to be established. If the information available to national agencies about the real state of the market is inaccurate, their ability to make sensible policy interventions is jeopardised.

The case studies clearly illustrate the missing link between authoritative land market information and elements of monetary and fiscal policies in Australia. The situation in other federated market economies such as the United States and India is similar, and potentially worse given the size and complexity of their respective federations. Where fiscal and monetary policies are used to manage a country’s wealth, policy makers need authoritative spatial information about market transactions to judge the state of the nation’s economy and make evidence-based policy decisions.

In countries with a tiered government structure where constitutions divide responsibilities, there is a need to initiate a paradigm shift regarding the importance of better information flows between the government creators and users of spatial information. The Property Market Tree represents a new model to illustrate better links between spatial information generated by land administration processes, property markets and macroeconomic functions, for better informed government: Figure 3.

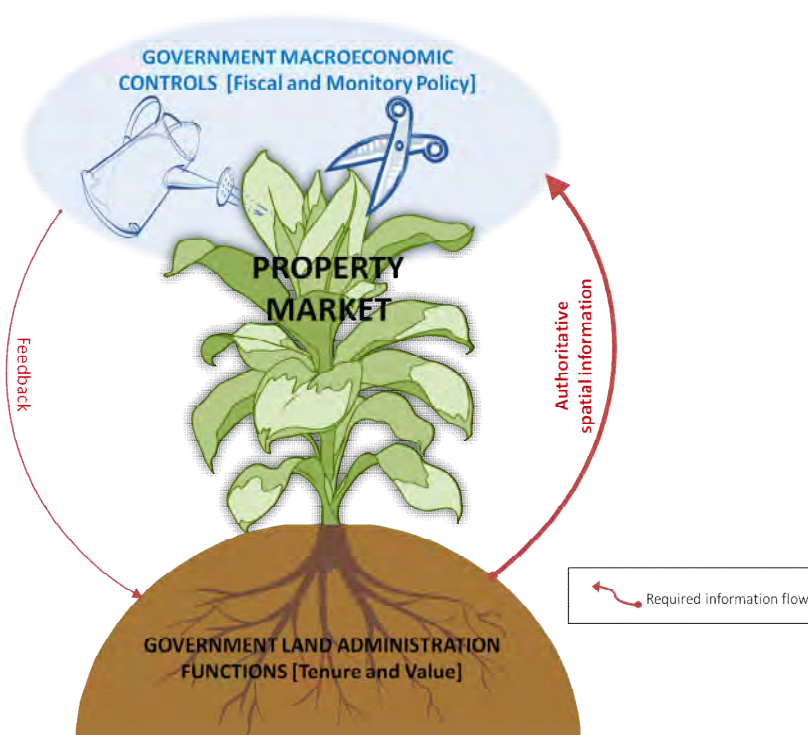


Figure 3: The Property Market Tree

The Property Market Tree illustrates the need for adequate information flows between the government land administration and policy institutions, in order to sustain a healthy land market. A healthy market can be growing or retracting, as circumstances suit. Currently there is little coordination or aggregation of spatial information within land administration in

many federated countries. Co-ordination and integration require continuous, constructive action.

The Property Market Tree aligns well with Australia's Spatial Data Infrastructure initiatives by ANZLIC since 1996, and the National Government Information Sharing Strategy (AGIMO, 2009) which promotes information sharing between government agencies in Australia. The strategy envisions that

“timely, reliable, and appropriate information sharing is the foundation for good government and has the capacity to deliver a better way of life for all Australians.”

Benefits to government agencies such as improved capacity for evidence-based policy and decision making and greater confidence in data quality and accessibility, are expected to ensue from agencies sharing information with each other (AGIMO, 2009).

Emerging practical approaches also support the paradigm shift recommended by the Property Market Tree. In Europe the INSPIRE Directive has led to the development of legally established ‘key registers’ of addresses and buildings in Netherlands use of which is mandatory for all public agencies in the Netherlands. Sedunary (1984), in his nodal approach to land database configuration, showed the need for high levels of communication between the primary nodes of legal/fiscal and geographic or land information. Additionally the popular and well-known concept of the multi-purpose cadastre to integrate land related data from individual land administration sectors was explained in the early 1980's (NRC, 1980; Kaufmann and Steudler, 1998). More recently, Roberge and Kjellson (2009) and Buhler and Cowen (2010) showed how the absence of a reliable property rights infrastructure and national cadastre in the United States contributed to the collapse of its land market.

Like the United States, in Australia too, spatial information datasets maintained by state land administration agencies cater primarily for internal information needs within individual agencies, and co-ordination at a national level is very limited. Land is an integral component of national wealth. Government policies formulated to manage national wealth should be based on authoritative information, a key component of which is spatial information about transactions in land. What federated market-based nations need is a better recognition of the role of accurate, authoritative and assured property market transaction information in macroeconomic policy making, and a seamless approach to land information access and delivery at higher levels of government.

6 CONCLUSION

Current economic and land administration literature does not adequately exemplify the importance of a seamless approach between property markets, spatial information and macroeconomic policy. In practice, the information flows between the tools of macroeconomic policy and market-supporting land administration functions are also problematic. This is particularly the case in many federated mixed capitalist nations where property market information is collected and maintained by state agencies, and macroeconomic policies are made at central government level. This paper presents a starting point to emphasise the importance of a holistic approach to land information management, at higher levels of government. The Property Market Tree illustrates that the tools for management of national wealth need authoritative information flows between land administration and federal policy making institutions. A national property market information infrastructure is the next step to achieving this goal. Future research should take a more detailed look at individual states within a federated country to identify the areas where access to spatial information by higher levels of government is impeded. Research is also required on the technical, institutional and legal requirements of national property information

infrastructure. While technical issues will also play a part, the governance, coordination and cooperation required for data sharing and integration is likely to be most challenging.

7 ACKNOWLEDGEMENTS

The authors are grateful for the assistance of our colleagues in the Centre for Spatial Data Infrastructures and Land Administration, Department of Infrastructure Engineering at The University of Melbourne in preparing this article. We would also like to thank our research partners, Land and Property Management Authority, Land Victoria, Landgate, PSMA Australia Limited and the Australian Research Council for their support. The views expressed in this article are those of the authors and may not represent those of the research partner organisations. Additionally, the authors would like to acknowledge that an extended version of this paper has been accepted for publication by the journal of *Environment and Planning C: Government and Policy*.

REFERENCES

- ABS (Australian Bureau of Statistics) (2011). Total Taxation Revenue, Australia, 2009-10. [<http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/5506.0Main%20Features12009-10?opendocument&tabname=Summary&prodno=5506.0&issue=2009-10&num=&view=>]
- Bennett R., Wallace J., Williamson I. (2008). Organising property information for sustainable land administration. *Journal of Land Use Policy*, 25(1), 126-138.
- Buhler D., Cowen D. (2010). The United States mortgage crisis and cadastral data. *Proceedings of FIG2010 Congress*, Sydney, Australia, 11-16 April.
- Carper D., West B., McKinsey J. (2007). *Understanding the Law*. Thompson.
- CCH Editors, (2010) *Australian master tax guide: tax year end edition 30 June 2010*. McPherson's Printing Group.
- Clapp M., Dolde W., Tirtiroglu D. (1995). Imperfect Information and Investor Inferences from Housing Price Dynamics. *Real Estate Economics*, 23(3), 239-269.
- Clarkson G., Jacobsen T., Batcheller A. (2007). Information asymmetry and information sharing. *Government Information Quarterly*, 24(4), 827-839.
- Dale P., McLaughlin J. (1999). *Land administration*. Oxford University Press.
- Deininger K., Binswanger H. (1999). The evolution of the World's Bank's land policy: principles, experiences, and future challenges. *World Bank Research Observer*, 14(2), 247-276.
- DeSoto H. (2000). *The Mystery of Capital: Why capitalism triumphs in the West and fails everywhere else*. Black Swan.
- Dolde W., Tirtiroglu D. (1997). Temporal and Spatial Information in Real Estate Price Changes and Variances. *Real Estate Economics*, 25, 539-565.

Enemark S. (2007). Land management in support of the global agenda. Proceedings of V International Congress GEOMATICA 2007, Havana, Cuba, 12-17 February.

FIG (International Federation of Surveyors) (1995). FIG Statement on the Cadastre. FIG Publication No. 11

Garmaise M., Moskowitz T. (2004). Confronting Information Asymmetries: Evidence from Real Estate Markets. *Review of Financial Studies*, 17(2), 405-437.

Golafshani N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, 8(4), 597-607.

Head B. (2009). Evidence-based policy: principles and requirements. Proceedings of Strengthening Evidence-based Policy in the Australian Federation: Roundtable Proceedings, Canberra, Australian, 17-18 August, 13 - 26.

Joshi H. (2006). Identifying asset price bubbles in the housing market in India - Preliminary evidence. *Reserve Bank of India Occasional Papers*, 27(1, 2).

Kaufmann J., Steudler D. (1998). Cadastre 2014: A vision for future cadastral systems. FIG.

Kohler M., Rossiter A. (2005). Property Owners in Australia: A snapshot. Research Discussion Paper 2005-03, Economic Research Department, Reserve Bank of Australia.

McKenzie D. J., Betts R. M. (2006). *Essentials of real estate economics*. Thomson South-Western.

McMillan J. (2003). *Reinventing the bazaar: a natural history of markets*. W.W. Norton.

Milgrom P., Stokey N. (1982). Information, Trade and Common Knowledge. *Journal of Economic Theory*, 26, 17-27.

North D. C., Thomas R. P. (1973). *The rise of the western world: A new economic history*. Cambridge University Press.

NRC. (1980). Need for a multipurpose cadastre. Panel on a Multipurpose Cadastre, Committee on Geodesy, Assembly of Mathematical and Physical Sciences, National Research Council. National Academy Press.

Poterba J. M. (1992). Taxation and housing: Old questions, new answers. *American Economic Review*, 82(2), 237-242.

Roberge D., Kjellson B. (2009). What Have Americans Paid (and Maybe the Rest of the World) for Not Having a Public Property Rights Infrastructure? Proceedings of FIG Working Week 2009, Eilat, Israel, 3-8 May.

Sedunary M. (1984). LOTS, and the nodal approach to total land information system. Proceedings of The Decision Maker and Land Information Systems - FIG Commission 3 International Symposium, Alberta, Canada, October.

Wallace J., Marwick B., Bennett R., Rajabifard A., Williamson I., Tambuwala N., Potts N., Agunbiade M. (2010). Spatially Enabling Land Administration: Drivers, Initiatives and Future Directions for Australia. In Rajabifard A., Crompvoets J., Kalantari M., Kok B. Eds. (2010)

Spatially Enabling Society, Research, Emerging Trends and Critical Assessment. Leuven University Press.

Wallace J., Williamson I. (2006). Building Land Markets. *Journal of Land Use Policy*, 23(2), 123-135.

Williamson I., Enemark S., Wallace J., Rajabifard A. (2010). *Land administration for sustainable development*. ESRI Press.

Review of Australian Land Use Mapping and Land Management Practice

Govinda Prasad Baral, Kevin McDougall, Albert Chong

Faculty of Engineering and Surveying & Australian Centre for Sustainable Catchments
University of Southern Queensland, Toowoomba 4350 QLD, Australia

GovindaPrasad.Baral@usq.edu.au, kevin.mcdougall@usq.edu.au, chonga@usq.edu.au

ABSTRACT:

Land use information plays a vital role in effective management of natural resources in any country. The land use and land cover mapping is always a dynamic issue in every country because of the changing nature of the land use. Australia is experiencing similar traits. Knowledge of land use change patterns has important implications for sustainable development and sustainable environmental management. It helps in the management of water, soil, nutrients, plants and animals and provides relationship between land use dynamics and economics and social condition in urban and regional area. Although Australia has a long history of land use mapping and land management practice, no systematic study of the land use mapping status and land management practice can be confirmed.

With the establishment of the Commonwealth Advisory Committee on the Environment in June 1972, later known as the Australian Advisory Committee on the Environment, the Australian government gave prime importance to land use and the environment. The committee recommended to the government that "Land use is fundamental to any consideration of the environment." The committee also found an urgent need for an efficient, co-ordinated, and comprehensive system of national and state land use planning. Subsequently, many organizations and institutions like BRS, ACLUMP, CSIRO, QDERM, and DPI started on land cover/land use mapping from national to catchment level and many Land Care groups began working on land management the local level. This study reviews the status of the land use mapping and land management practices as implemented in Australia

KEYWORDS: Land use, Land cover, ALUMP, Land Management Practice, Australia

1. BACKGROUND

On June 13th, 1992, nearly 100 heads of states, the biggest gathering of heads of states, met during "Earth Summit" in Rio de Janeiro, Brazil to discuss the environment and development. This led to Agenda 21 - a very important document to guide the development of earth in a sustainable manner (AGENDA 21 1992). This document emphasizes the necessity of research to determine the capacity of land and the interaction among various land uses and environmental processes. Agenda 21 defines land as a "physical entity as well as a system of natural resources". Land resources include the soil, minerals, water, plants and animals in all their biological and generic diversity. An integrated approach to land use will provide a broader perspective from which to make informed choices regarding safest and most efficient use of land resources.

Below are a few programs and activities related to this research, from 19 of such programs suggested by Agenda 21 (AGENDA 21 1992).

1. Improved planning and management systems require more appropriate tools for data collection and interpretation. Detailed land inventories must be undertaken to determine the actual capacities of local land area.

2. Far more research is needed regarding land resource use potential and the interactions between the various factors affecting land use. Pilot projects based on an integrated approach to land use must be developed and tested. The results of these projects must be distributed to local and community land use planning groups in order to enable them to make decisions based on the latest available information.

These two programs suggested by Agenda 21 emphasize the need for improved methods of data collection, interpretation and implementation of land use maps at the local level. This highlights the relevance of this research since this research reviews the existing status of land use mapping in Australia against the notion of Agenda 21

Knowledge of land use change pattern has important implications in sustainable development and sustainable environmental management (Cheng et al. 2008; Tiwari et al. 2010). Australia has a long history of land use mapping and land management practice (AACE 1974) but no systematic studies of the status of land use mapping and the land management practice were found.

National Land and Water Resources Audit convened a land cover workshop in Canberra on the 24th July 2007 (Audit 2007). A pre-workshop meeting was held to determine the agenda and identify who should attend. Representatives from various levels of government, who had a vested interest in land cover identified user requirements for a number of natural resource themes such as soils, water and vegetation. User needs, applications and gaps were also identified. Many speakers highlighted the necessity of improvement of the land use classification method and establishing collaboration and partnership between the stakeholders to avoid duplication of work. One of the key issues and challenges determined by the workshop was that a refined "Classification system needs to be determined/developed/endorsed. There are many land cover classification systems. An Australian standard classification system will need to be determined by the co-investors and collaborators" (Audit 2007).

2. CLARIFICATION OF TERMINOLOGY

There is often confusion among the terms 'land use' and 'land cover'. They may be applied in the same context, perhaps because of the common use of remotely sensed satellite imagery or photography for mapping. The distinction between 'land use' and 'land management practice' is also not always well understood (Lesslie 2004). A number of land related terms and classifications are often used to describe the earth's surface characteristics, the production capability of land, the uses to which land is put and the ways in which land is managed. Mixed use of terminology can cause confusion as to the specific component being described or assessed by particular initiatives (Audit 2007).

2.1. Land Cover

Land Cover is the observed biophysical cover on the earth's surface (Gregorio & Jenson 2000). Land cover refers to the observed physical surface of the earth, including various combinations of vegetation types, soils, exposed rocks, water bodies (BRS 2006).

2.2. Land Use

Land use refers to the purpose to which land is committed, including the production of goods (such as crops, timber and manufacturing) and services (such as defence, recreation, and biodiversity and natural resources protection). Some land uses, such as cropping, have a characteristic land cover pattern. These land uses frequently appear in land cover classifications. Other land uses, such as nature conservation, are not readily discriminated by a characteristic land cover pattern. For example, where the land cover is woodland, land use may be timber production or nature conservation (BRS 2006).

Figure 1 below illustrates the differences between land use and land cover and their implications

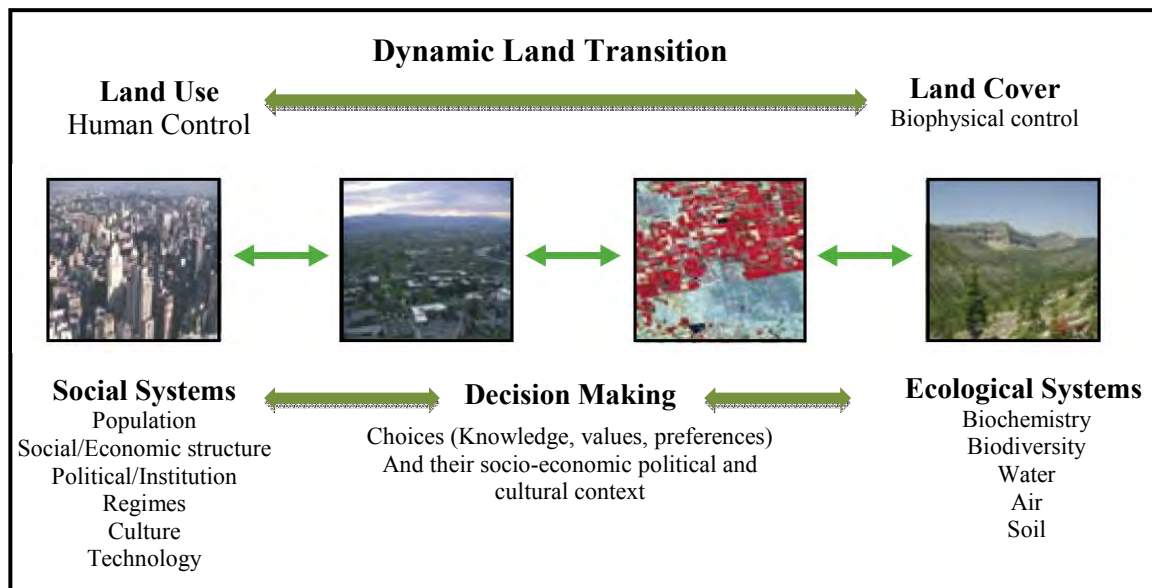


Figure 1. Land use - land cover relations (Adapted from Global Plan Report (GLP 2005))

2.3. Land management practice

Land management practice means the approach taken to achieve a land use outcome — the 'how' of land use (e.g. cultivation practices, such as minimum tillage and direct drilling). Some land management practices, such as stubble disposal practices and tillage rotation systems, can be discriminated by their characteristic land cover patterns (ACLUMP 2010a).

2.4. Land capability and land suitability

Land capability assesses the limitations to land use imposed by land characteristics and specifies management options. Land suitability (assessed as part of the process of land evaluation) is the fitness of a given type of land for a specified kind of use (ACLUMP 2010a).

3. LAND USE AND DEVELOPMENT

The development and use of land has been a fundamental human activity since the dawn of agriculture and the permanent human settlements. Hamlets, villages, towns, and cities evolved to accommodate larger populations and the developing needs of society for livelihood, security, commerce, and culture (Randolph 2004). Information on land use and land management practices is essential for the sustainable management and economic development of natural resources. Information regarding the current state of resources and how they are being managed is required to predict future states. It is also crucial to the understanding of land degradation processes, for predicting their potential outcomes as well as for the evaluation of natural resource and agricultural investment strategies (Rowland & Calvert 2000). The industrial revolution, the expansion of the railway and transportation, the advancement in communication means and so on changed the urban settlement pattern and the land development. Most people live in urban and suburban areas hence considerable attention is given to land use of such areas. John Randolph in his book *Environmental Land Use Planning and Management* (Randolph 2004) states "Rural and small-town land use and development are also important in environmental land use management for three regions. First, these greenfield areas are home to important ecological, cultural, and agricultural resources. Second, inherent use of rural land for resource production of agriculture, forestry, and mineral extraction has considerable environmental impact. And third, rural places are increasingly attractive as people grow weary of the congestion and lifestyle of the city and suburbs. As a result, sprawling patterns of rural development are impacting them at an increasing rate"

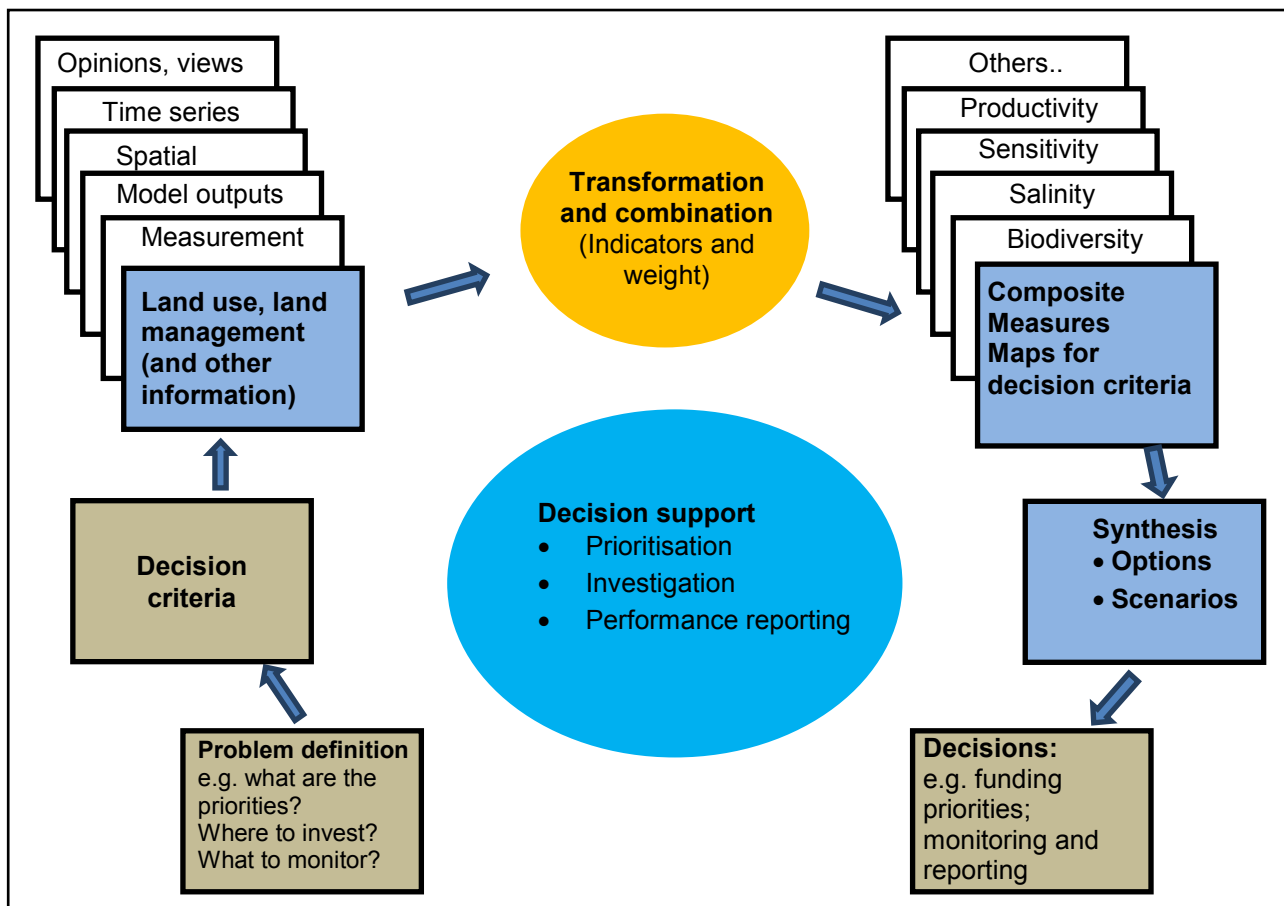


Figure 2. The contribution of land use and land management practices information to policy and program decision making (Adapted from ACULUMP(2010b))

Whatever the reason, development and land use are interdependent (Figure 2); forming the basis for sustainable land use management which is the basis for urban and rural land use planning.

4. HISTORY OF LAND USE IN AUSTRALIA

With the establishment of the Commonwealth Advisory Committee on the Environment in June 1972, later known as the Australian Advisory Committee on the Environment, the Australian government gave prime importance to land use and the environment. The committee recommended to the government that “Land use is fundamental to any consideration of the environment, and that land use management is one of the most important single elements affecting the quality of environment”(AACE 1974). The committee also found an urgent need for an efficient, co-ordinated, and comprehensive system of national and state land use planning as part of the broad spectrum of natural resources planning and management.

Similar views were also reflected in reports from the United Nations Conference on the Human Environment, in Stockholm June 1972, stressing the importance of resource management for the preservation and improvement of the environment, and including recommendations aimed at promoting the development of methods of integrated planning and management of natural resources (AACE 1974). The map and table below show land use distribution in Australia (Figure 3 and Table 1).

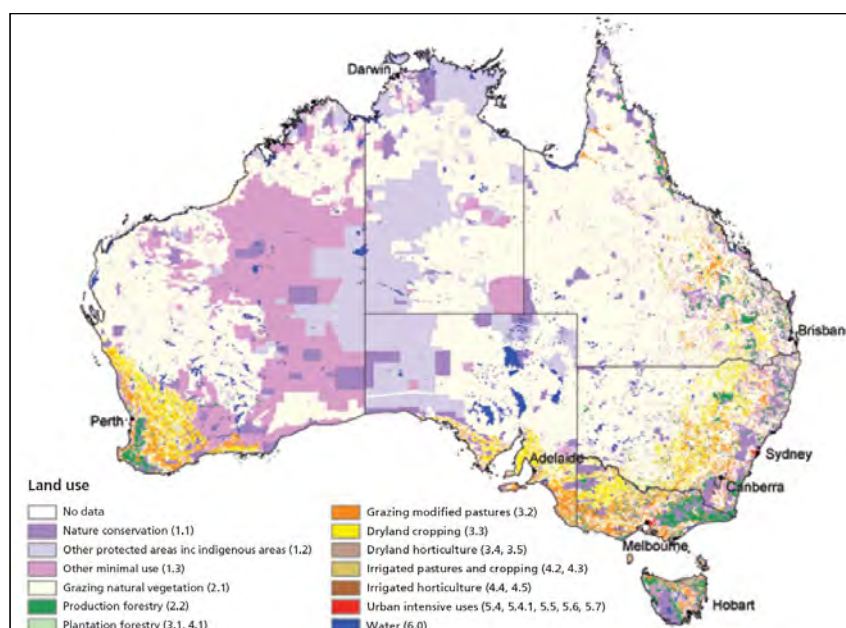


Figure 3. Land use mapping at national scale for the National Land & Water Resources Audit
ALUM Classification (v6)

Table 1. ACLUMP National scale land uses 2005–06(ACLUMP 2010b)

Land use	Area (Sq. Km)
Nature conservation	571,483
Other protected areas including Indigenous uses	1,015,359
Minimal use	1,242,715
Grazing natural vegetation	3,558,785
Production forestry	114,314
Plantation forestry	23,929
Grazing modified pastures	720,182
Dryland cropping	255,524
Dryland horticulture	1,092
Irrigated pastures	10,011
Irrigated cropping	12,863
Irrigated horticulture	3,954
Intensive animal and plant production	3,329
Intensive uses (mainly urban)	16,822
Rural residential	9,491
Waste and mining	1,676
Water	125,618
Total	7,687,147

The committee recommended that because of the necessity of a national survey of land attributes, research on ecological relationships, national resource data banks and inventories be investigated as these form the basis for land use mapping and conducting research in this field. After that many organizations and institutions like BRS, ACLUMP, CSIRO, SLATS, DPI etc. worked on land cover/land use mapping in national to catchment level and many Land Care groups started working on land management at the local level.

Before 1999 the availability of detailed mapping in Australia was limited and uncoordinated. Australian and State government agencies independently produced land use mapping at a range of scales using a variety of cartographic methods and classification systems. A collaborative national land use mapping initiative was established in 1999 by the Audit, the Department of Agriculture, Fisheries and Forestry (DAFF), the Bureau of Rural Sciences (BRS), the Murray-Darling Basin Commission and State agency partners. In 2000 DAFF accepted leadership for the national coordination of land use information and BRS took on responsibility for the development of ACLUMP. This is continued under the Audit's current national information coordination arrangements. Mapping products are now in strong demand, and there is widespread adoption of agreed standards (ACLUMP 2010d). Australian Bureau of Agriculture, Resource Economics and Sciences (ABARES) under Department of Agriculture, Fishery and Forestry is formed by merging the Bureau of Rural Sciences and the Australian Bureau of Agriculture and Resource Economics in July 2010. The merger enabled the bureau to provide an integrated research offering to allow evidence-based policy making by both science and economic analysis and advice (ABARES 2011). At the state level, Queensland Land Use Mapping Program (QLUMP) maps Queensland at catchment scale using Australian Land Use and Management (ALUM) classification scheme which has a three-tiered hierarchical structure with primary, secondary and tertiary classes (Rowland et al. 2002).

5. LAND USE MAPPING AND MANAGEMENT INITIATIVES IN AUSTRALIA

Land cover had been identified as a critical information need for national and regional reporting and decision making by the National Land & Water Resources Audit (the Audit), Geoscience Australia (GA), the Bureau of Rural Sciences (BRS), the Commonwealth Scientific Industrial Research Organisation (CSIRO) and most of the national coordinating committees (representing all jurisdictions). There have been a number of land cover related projects undertaken in Australia by Australian Government departments, including the Agricultural Land Cover Change (ALCC) Project, the National Vegetation Information System (NVIS), Australian Land Use and Management (ALUM), Vegetation Assets and Transition (VAST) and the Australian Greenhouse Office National Carbon Accounting System (NCAS) Land Cover Change Program (Audit 2007).

The major initiatives for land use/land cover mapping are presented in the following sub-sections.

5.1. ACCLUMP

The Australian Collaborative Land Use Mapping Program (ACLUMP) is a consortium of Australian and State Government partners. The program is part of the national natural resource information co-ordination arrangements established by the National Land & Water Resources Audit (the Audit). ACLUMP promotes the development of nationally consistent land use and land management practices and information for Australia (ACLUMP 2010d). ACLUMP activities are financially supported by program partners, the Natural Heritage Trust and the National Action Plan for Salinity and Water Quality.

5.1.1. Stakeholders in land use mapping and management in Australia

Listed below are the major players in land use mapping and management activities especially those who are partners of ACLUMP:-

State agency partners:

- New South Wales - Department of Infrastructure, Planning and Natural Resources
- Northern Territory - Department of Infrastructure, Planning and Environment
- Western Australia - Department of Agriculture
- Queensland - Department Environment and Resources Management

- South Australia - Department of Water, Land and Biodiversity Conservation
- Victoria - Department of Primary Industries
- Tasmania - Department of Primary Industries, Water and the Environment

Australian Government partners:

- Department of Agriculture, Fisheries and Forestry
- Australian Bureau of Agriculture, Resource Economics and Sciences (ABARES)
- National Land & Water Resources Audit
- Murray-Darling Basin Commission
- Australian Greenhouse Office

Others:

- Natural Heritage Trust
- National Action Plan for Salinity and Water Quality
- Local Authorities
- Land care groups

5.2. LUMIS

The mechanism for developing a collaborative program to record and map land management practices is a Land Use and Management Information System (LUMIS). The first of concern to LUMIS is information about actual on-ground action by resource managers. The second is that information about capacity and willingness of managers to make changes in how they manage, is highly relevant but outside the scope of LUMIS (ACLUMP 2010c).

The LUMIS development is advised by the National Committee for Land Use and Management Information (NCLUMI). Its membership includes Australian and State Government partners. Its work is sponsored by the Australian Government Department of Agriculture, Fisheries and Forestry with day-to-day management provided by the Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES). NCLUMI and its technical advisory group(s) need to ensure that LUMIS development maintains consistency with the Australian Land Use and Management (ALUM) classification.(ACLUMP 2010c)



Figure 4: LUMIS's hierarchical system from target to action (BRS 2011)

There are too many different management actions across Australia to collect and map them individually. To deal with this problem, LUMIS groups related management actions into a number of standardised categories. These categories — what we are calling 'practices' — can be

hierarchically organised to show their direct relationships to the primary 'object' that is being managed (e.g. vegetation or water or animals). LUMIS will accommodate land management practices information compiled at varying levels of detail and mapping scale. It will also include technical specifications for spatial sampling and data collection. Figure 4 shows the hierarchical system of LUMIS (BRS 2011).

5.3. Land use classification schemes

The ALUM Classification has a three-tiered hierarchical structure with primary, secondary and tertiary classes broadly structured in terms of the potential degree of modification and impact on a putative 'natural state' (essentially unmodified native land cover).

Primary and Secondary classes relate to land use - the prime use of the land defined in terms of the management objectives of the land manager. Tertiary classes can include commodity groups, commodities, land management practice, or land cover (e.g. vegetation) information. Tertiary agriculture classes have been based on Australian Bureau of Statistics commodity groups as well as dominant land management practices.

ALUM Primary classes are as given below-

- a) *Conservation and natural environments* - Land used primarily for conservation purposes, based on the maintenance of the essentially natural ecosystems present.
- b) *Production from relatively natural environments* - Land used primarily for primary production with limited change to the native vegetation.
- c) *Production from dryland agriculture and plantations* - Land used mainly for primary production, based on dryland farming systems.
- d) *Production from irrigated agriculture and plantations* - Land used mostly for primary production based on irrigated farming.
- e) *Intensive uses* - Land subject to extensive modification, generally in association with closer residential settlement, commercial or industrial uses.
- f) *Water* - Water features. Water is regarded as an essential aspect of the classification, but it is primarily a cover type.

The classification is intended to be flexible such that new land uses or management systems can be accommodated as long as there is no conflict with other existing items.

5.4. Land use classification method used in Australia

Land-use history – the number, type, and duration of previous land uses – is relevant to many questions (Kasel & Bennett 2007).

A joint Commonwealth-State workshop in February 1999 agreed that a modified version of a classification scheme developed by Baxter and Russell in 1994 would be suitable as a land use classification for Australia (BRS 2006). It would promote the creation of nationally consistent, although not necessarily uniform, land use data sets, meet a wide range of user needs, and make the best use of existing data and available resources (BRS 2002). A flow diagram showing the land use mapping method used in Australia is given below (Figure 5).

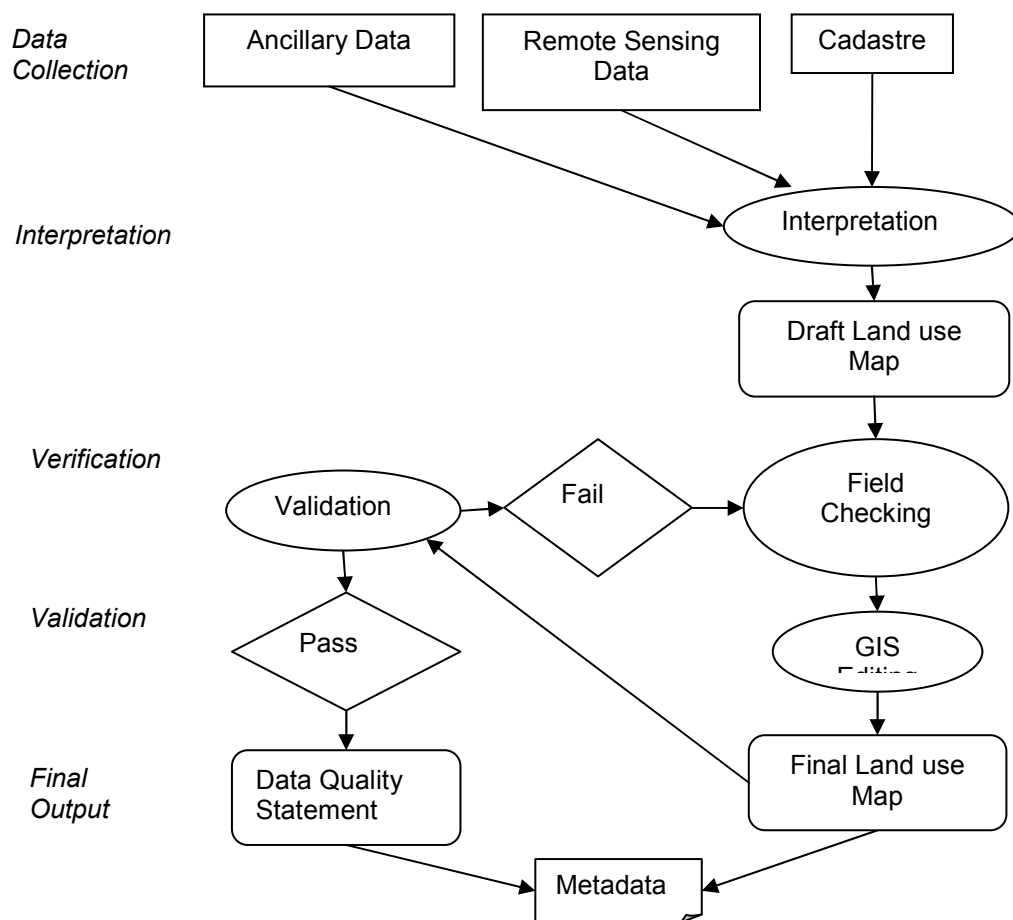


Figure 5: Land use mapping procedure (Rowland et al. 2002)

6. CHALLENGES AHEAD

Sustainable land management is a central challenge in the sustainable management of earth systems and resources. On the one hand, land management must ensure a growing supply of food and other resources to human populations, which are expected to grow for decades to come. On the other hand, management of land to procure these resources is linked with potentially negative consequences in the form of climate change, biodiversity loss and pollution. Moreover, local alteration of land use and land cover can have global consequences, requiring local and regional solutions to global problems and the cooperation of the world's policymakers, land managers, and other stakeholders in land management at local, regional and global scales (Ellis 2010).

Land use information is currently being used in Australia to manage catchment salinity, nutrient and sediment problems, measure greenhouse gas emissions and sinks, assess agricultural productivity and opportunities for agricultural diversification, for land value determination, in local and regional planning, pest and disease control and emergency response planning. Vegetation management, monitoring greenhouse gas emission from tree clearing, rural leasehold land strategy, controlling spread of woody weeds, dam safety and water use, water quality monitoring are the major drivers for land use mapping. Although there is huge requirement for up to date and precise land use data, there are still many challenges to overcome. Issues of spatial and temporal scales that are fit for the purpose of planning and development at national land regional level is to be addressed. It is essential to define/modify land use/cover classes according to the needs of the users.

Listed below are few of many challenges for Australian land use mapping against the user requirement.

- developing effective mapping tools for land use decision makers
- establishing appropriate land use mapping standards and specifications

- using existing data from land use layers to create a baseline, avoid mismatching and conflicting data
- avoiding duplication of effort between different data sets
- facilitating and coordinating land use mapping across jurisdictions
- time series land use data creation
- greater resolution land use/cover information using multi sensor images as an alternative to Landsat images
- introducing new applications like water quality, riparian vegetation.

7. CONCLUSION AND RECOMMENDATIONS

It is obvious that state and local agencies have requirements of higher scale data than Australian Government agencies. Existing National Scale (1:1 000 000 to 1:2 500 000), and Regional and Catchment scale (1:25 000 – 1:250 000) land use dataset may not be sufficient to meet the increasing demand of users. Higher scale land use datasets are the up to sub-catchment level (up to 1:10 000) is the requirement of time. This is only possible if we have a cohesive partnership between the stakeholders to avoid the duplication of work and sharing data, resources and knowledge. A consistent set of land use classes should be developed and the current land use classification scheme should be modified for the automation of classification process and to make it friendly across different satellite sensors. A common repository of time series data sets in different scale should be created. A mechanism should be developed to serve these very large data sets to meet client requirements.

This study reviewed the state-of-the-art status of Australian land use mapping. As stated before, up to date and accurate land use information is the basis of sustainable development and sustainable environment. To keep an equal footing between the development and environment such up to date land use information plays an important role.

References:

AACE 1974, *Land use in Australia*, Australian Government Publishing Service, Canberra.

ABARES 2011, viewed 2011/08/17, <www.daff.gov.au/abare-brs>.

ACLUMP 2010a, *Land Use in Australia – At a Glance*, Australian Collaborative Land Use Mapping Programme (ACLUMP).

ACLUMP 2010b, *Land Use and Land Management Information for Australia: Workplan of the Australian Collaborative Land Use and Management Program*, Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra.

ACLUMP 2010c, *Status of Land Management Practices:Activities of the Australian Collaborative Land Use and Management Program: Status report*, Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra.

ACLUMP 2010d, *Australian Collaborative Land Use Mapping Program, Information for landscape solutions*, viewed 07/04/2011, <<http://adl.brs.gov.au/mapserv/landuse/index.cfm?fa=acl.aclump&tab=aclump>>.

AGENDA 21 1992, *AGENDA 21 : the Earth Summit strategy to save our planet*, EarthPress, Boulder, Colo.

Audit, NLWR 2007, *Australian land cover mapping: Proceedings of a workshop to discuss interest in land cover mapping*

BRS 2002, *Land Use Mapping at Catchment Scale: Principles, Procedures and Definitions*, Edition 2.

BRS 2006, *Guidelines for land use mapping in Australia: principles, procedures and definitions. A technical handbook supporting the Australian Collaborative Land Use Mapping Programme*, Bureau of Rural Sciences.

BRS 2011, *Land management Practices Classifications and Mapping*, Bureau of Rural Sciences, viewed 2011/02/05, <<http://adl.brs.gov.au/mapserv/landuse/index.cfm?fa=landmage.Progreport&tab=Progreport>>.

Ellis, E 2010, *Land-use and land-cover change*, Eds. Cutler J. Cleveland (Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment). [First published in the Encyclopedia of Earth April 18, 2010; Last revised Date April 18, 2010; Retrieved February 3, 2011 <http://www.eoearth.org/article/Land-use_and_land-cover_change>.

GLP 2005, *Science Plan and Implementation Strategy*, IGBP Report No. 53/IHDP Report No. 19., IGBP Secretariat, Stockholm. 64pp.

Gregorio, AD & Jenson, LJM 2000, *Land Cover Classification System (LCCS): Classification Concepts and User Manual*, Food and Agriculture Organisation of the United Nations, Rome.

Kasel, S & Bennett, L 2007, 'Land-use history, forest conversion, and soil organic carbon in pine plantations and native forests of south eastern Australia', *Geoderma*, vol. 137, no. 3-4, pp. 401-13,

Lesslie, R 2004, *Land use and land management practices: Concepts, terms and classification principles*, Kamberra Winery, Canberra 11-12 May 2004.

Randolph, J 2004, *Environmental land use planning and management*, Island Press, Washington, DC.

Rowland, T & Calvert, M, *Land use management mapping for the Murray darling Basin, Phase 1, Report of Queensland Pilot Study in St. George*, 2000, DoNR Queensland, Brisbane.

Rowland, T, O'Donnell, T, Weller, D & Witte, C 2002, 'Catchment Scale Land Use Mapping in Queensland', paper presented to 11th Australasian Remote Sensing and Photogrammetry Conference, Brisbane, Queensland.

Developing Spatio-temporal Prediction Models for Arbovirus Activity in Northern Australia Based on Remotely Sensed Bioclimatic Variables

Bernhard Klingseisen, Robert J. Corner

Department of Spatial Sciences, Curtin University of Technology
GPO Box U1987, Perth WA 6845, Australia
bernhard.klingseisen@gmail.com, r.corner@curtin.edu.au

Mark Stevenson

EpiCentre, Massey University
Private Bag 11-222, Palmerston North 4442, New Zealand
m.stevenson@massey.ac.nz

ABSTRACT

Vector-borne diseases pose an ongoing threat to public and animal health in the north of Australia. A number of surveillance programs are in place to determine the extent of virus activity and control the risk, but these are labour- and cost intensive while producing data with large temporal and spatial gaps. Using the example of Bluetongue virus, the aim of this study was to investigate the potential of remotely sensed variables to facilitate the development of area-wide predictive models that complement traditional surveillance activities.

Bioclimatic variables were derived for the Northern Territory from MODIS and TRMM remote sensing data products covering a period of nine years. Spatial and temporal uncertainty in the surveillance data required the annual aggregation of environmental variables on a pastoral property level. Generalized Additive Models (GAM) were developed based on variables such as NDVI and land surface temperature to produce annual prediction maps of virus activity. External validation showed that the model correctly predicted 75% of the results from cattle stations tested for Bluetongue. Remaining uncertainty in the model can be mainly attributed to the spatio-temporal inconsistency of the available surveillance data.

This case study has developed a cost-effective approach based on a set of robust environmental predictors that facilitate the generation of arbovirus prediction maps soon after the peak of risk for infection. While this research focused on Bluetongue Virus, we see a large potential to expand the method to other areas and viruses particularly in view of the increasing populations in Northern Australia.

KEYWORDS: arbovirus, epidemiology, remote sensing, Northern Australia, spatio-temporal modelling

1 INTRODUCTION

Australia hosts more than 75 arthropod-borne (arbo-) viruses, which are a group of viruses maintained in nature by transmission between vertebrate hosts (e.g. humans, livestock, feral animals) and bloodsucking insects, such as mosquitoes, ticks or biting midges. Twelve of these viruses are of concern for human health, including Dengue, Ross River or Murray Valley Encephalitis. Others, such as Bluetongue, Akabane or Bovine Ephemeral Fever are affecting animal health and impact the economics of the livestock industry. Each component of the transmission cycle is influenced by the interplay of underpinning environmental variables such as climatic conditions, vegetation, terrain and soil properties, which characterise the vector and host habitat structure. This paper presents the results of a case study that investigated the utilisation of remotely sensed variables for the area wide prediction of Bluetongue Virus (BTV) to support

ground based monitoring efforts in the remote areas of Northern Australia. For full details see Klingseisen (2010).

The first evidence of Bluetongue virus in Australia was provided by the isolation of serotype 20 from *Culicoides* collected at Beatrice Hill, Northern Territory, in March 1975. Following this discovery, the existing sentinel herd system, established in the late 1960's in Northern Australia, was immediately expanded to other regions. In subsequent years, further serotypes, namely 1, 2, 3, 7, 9, 15, 16, 21 and 23, were isolated (Animal Health Australia 2001, 2008, 2009). Of those, serotypes 1 and 21 are widely distributed across northern and eastern coastal regions of Australia, including the northern parts of Western Australia (WA), the Northern Territory (NT) and Queensland (QLD) as well as eastern Queensland and north eastern New South Wales (NSW) (Kirkland 2004). This corresponds well with the observed limits of the biting midge *Culicoides brevitarsis* (Murray and Nix 1987) which, due to its abundance, is considered the most important vector species for BTV transmission in Australia, although it is not the most competent vector.

More competent vectors and pathogenic BTV serotypes have as yet been confined to the far North of the continent, far removed from susceptible commercial sheep populations. Although disease has been observed in small groups of susceptible sheep when they were moved to the tropical part of the NT (Kirkland 2004), no signs of clinical disease have been found in sheep raised in the area. However, due to factors such as climate change and increases in the frequency and distance of livestock movements, the risk of a southward spread of the virus cannot be ignored. The presence of BTV resulted in stringent export restrictions being applied to both cattle and sheep after its first isolation. Even if trade is not prevented, it becomes very expensive due to serological tests and other measures necessary to reduce the perceived BTV risk (Oliver 2004).

Since 1992, surveillance for BTV has been conducted as part of the National Arbovirus Monitoring Program (NAMP). The system consists of sentinel herds in areas where commercial livestock are raised with sites selected representatively to allow mapping of the distribution of infection (Melville 2004). Most herds are positioned along the border between expected infected and uninfected areas, or where infection occurs irregularly. As part of irregular opportunistic serological surveys, herds within the affected areas are also tested to assess the seasonal intensity of infection. Supplementary areas, expected to be uninfected, are monitored to verify their BTV-free status (Animal Health Australia 2009). Sentinel herds usually consist of 10 to 25 young cattle, which have initially been negatively tested for BTV antibodies. Cattle in sentinel herds are replaced annually or after seroconversion has occurred. Sentinel herds are bled at regular intervals, with the frequency being approximately proportional to the observed arbovirus activity. Seroconversion information is supplemented by vector trapping and the quantification of *Culicoides* species at sentinel herd sites and numerous other strategic locations (Cameron 2004).

Under the NAMP and its predecessor programs, many years of monitoring data have been collected and detailed knowledge of virus ecology and the linkages between environmental factors and virus presence has been accumulated. The ongoing operation of the NAMP is expensive and sampling is often impractical in the remote areas of Northern Australia, particularly during and after the wet season. Due to these constraints it is useful to investigate how accumulated data can be used to inform future surveillance and to determine the utility of alternative information sources, such as remotely sensed data, to predict both the place and time of seroconversion.

Models predicting the spatial and temporal distribution of BTV and other diseases as well as their vectors have been developed for the Mediterranean and North Africa (Tatem et al. 2003), Spain (Calvete et al. 2009), Sicily (Purse et al. 2004) and Corsica (De La Rocque et al. 2004; Guis et al. 2007). In these models, environmental variables derived from remote sensing satellites have been used as surrogates for the bioclimatic factors that define vector and host habitats. Amongst the variables tested, temperature minima and maxima, and the Normalised Difference Vegetation Index (NDVI) have shown the most significant correlations with BTV vector occurrence (Baylis et al. 1998; Calvete et al. 2009; Purse et al. 2007; Wittmann, Mellor, and Baylis 2001).

Remote sensing satellites are able to provide area-wide coverage of environmental conditions over long periods and hence facilitate monitoring of vector presence as it changes over time. The Moderate Resolution Imaging Spectroradiometers (MODIS) on NASA's earth observation satellites Terra and Aqua are able to measure geophysical parameters at a spatial resolution between 250 and 1000 m four times daily. This configuration aids in obtaining cloud-free images in tropical areas, where cloud cover is a major issue and optical remote sensing data may

be only occasionally available during the wet season. A number of freely available pre-processed data products, including Land Surface Temperature, and Vegetation Indices, make MODIS a valuable resource for epidemiological applications (Tatem, Goetz, and Hay 2004). Data and derived products from the Tropical Rainfall Measuring Mission (TRMM), a joint mission by the US and Japanese Space Agencies NASA and JAXA, respectively, are also freely available as 3-hourly and daily rainfall estimates in tropical regions around the equator (Ebert, Janowiak, and Kidd 2007; Hay 2000). Both earth observation systems are operational at the time of writing and follow-up missions are planned to ensure the ongoing supply of data to build long term time series.

Despite the increasing availability and application of remote sensing data in spatial epidemiology (Ostfeld, Glass, and Keesing 2005) and although a number of models for BTV have been developed for Australia based on environmental variables (Cameron 2000; Murray 1995; Murray and Nix 1987; Ward and Thurmond 1995), no attempt has been made to use satellite images for the prediction of BTV seroconversion risk. In this study, we investigate the use of remotely sensed data as input for BTV distribution models as a means for providing better focus for BTV surveillance, i.e. identify areas of the country with a higher risk for BTV seroconversion. Using data for the seasons 2000/2001 to 2007/2008 we tested a range of environmental variables as predictors within a multivariate regression model. We then assessed the model's predictive capability using data from the 2008/2009 season. In contrast to other studies that compiled seasonality variables over a number of years to model general habitat suitability (Hay, Graham, and Rogers 2006; Scharlemann et al. 2008), prediction maps for virus presence are generated for each individual year. This approach facilitates the generation of annual prediction maps and the analysis of drivers for the inter-annual variations in virus distribution.

2 METHODOLOGY

2.1 Study Area

The analysis focused on the Northern Territory of Australia, where Bluetongue is endemic north of 13° latitude, and epidemic south of that line, with occurrences in Katherine and the Victoria River District in some years. The peak of virus activity is typically between January and May, but late seroconversions have occurred between August and September (Animal Health Australia 2001). The climate of the Northern Territory is under the influence of the north-west monsoon with a summer wet season and a winter dry-season. From north to south the climate changes from Köppen-Geiger zones Tropical Savannah to Hot Arid Steppe and Hot Arid Desert (Peel, Finlayson, and McMahon 2007), with annual rainfall ranging between 250-500 mm in the South and 1000-1700 mm in the North. Most rain falls between November and April. While it is warm in the North throughout the year, frost can occur in the arid interior as far north as Alice Springs (Wilson et al. 1990).

2.2 Arbovirus Monitoring Data

Surveillance data for BTV were obtained from the NAMP information system (Cameron 2004) on a pastoral property level for the years 2000 – 2009. This includes data on the seroconversion of animals within sentinel herds as well as data on serological status of animals derived from periodic cross sectional surveys. Due to the irregularity of sampling intervals at most localities, the true date of seroconversion is unknown. Therefore, data on BTV status were aggregated on an annual basis, using samples collected from 1 November to 31 October the following year. This period incorporates seroconversions at the peak of transmission during and shortly after the wet season between late summer and autumn (Animal Health Australia 2001) and the months thereafter to the onset of the new season. Figure 1 provides an overview of monitoring sites that were included in this study and a summary of the seroconversion data is provided in Table 1.

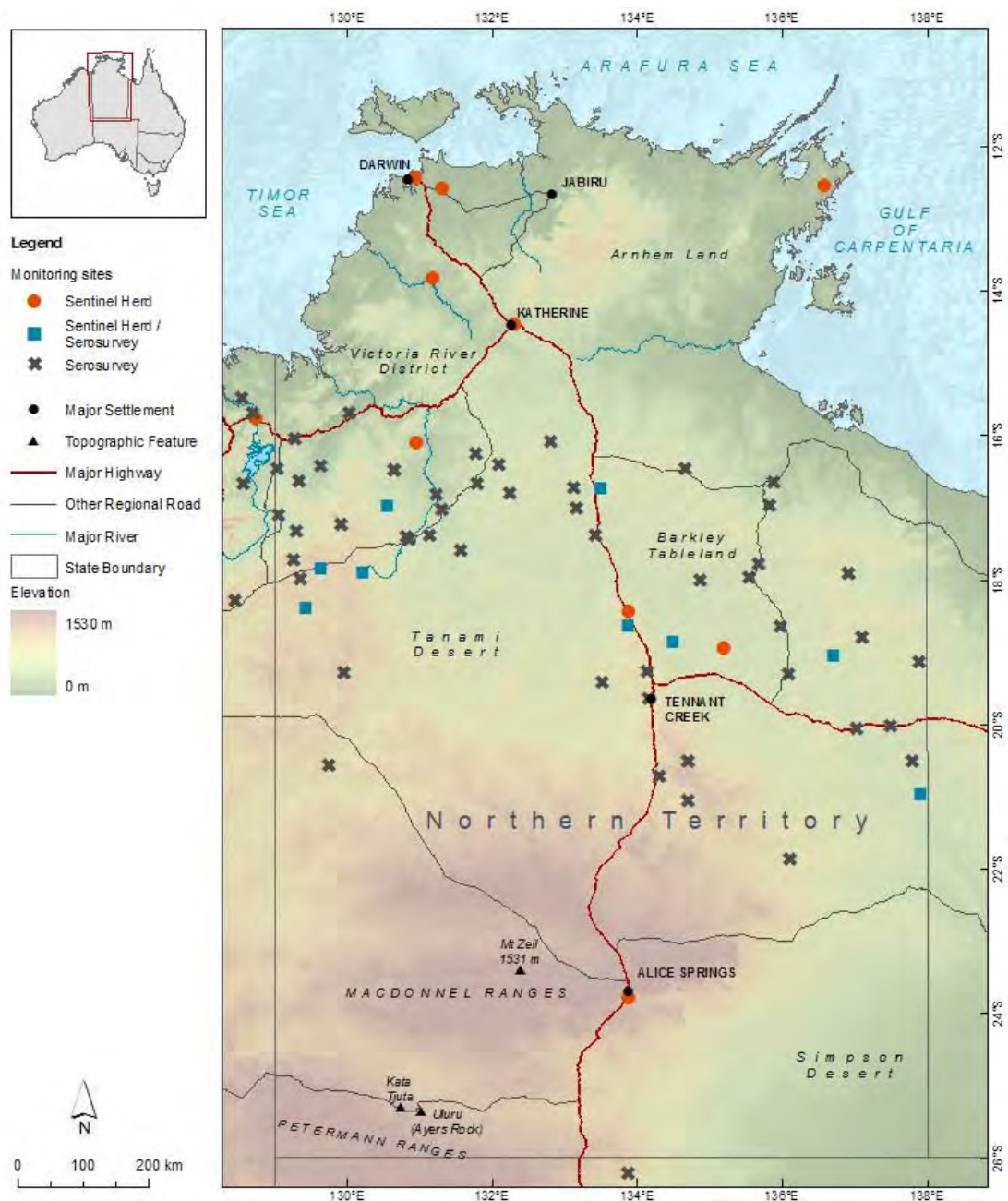


Figure 1: Map of NAMP sites in the NT tested between November 2000 and October 2009

Table 1: Summary of BTV seroconversion data used in this study

Year	No. of tested properties	Prevalence (95% confidence interval)
2000/2001	14	0.36 (0.13-0.37)
2001/2002	18	0.67 (0.41-0.87)
2002/2003	25	0.68 (0.46-0.85)
2003/2004	23	0.74 (0.52-0.90)
2004/2005	18	0.28 (0.10-0.53)
2005/2006	10	0.80 (0.44-0.97)
2006/2007	17	0.70 (0.44-0.90)
2007/2008	37	0.54 (0.37-0.71)
2008/2009	38	0.71 (0.54-0.85)

2.3 Remotely sensed environmental variables

Based on a review of BTV host and vector ecology in Australia, land surface temperature, precipitation, and vegetation indices as well as phenological characteristics of the growing season were identified as most relevant for the presence of the virus. These factors are related to the survival and abundance of the main BTV vector *C. brevitarsis*, which maintains a strong relationship with cattle as the preferred hosts for blood feeding, oviposition and larval development in bovine dung pads. Considering the large area covered in this study, part of which is characterised by a tropical climate, moderate to low spatial resolution satellite data products from MODIS and TRMM delivered at a high temporal resolution were considered most appropriate to monitor environmental conditions associated with vector habitat dynamics.

To derive vegetation indices and land surface temperature, the MODIS MOD11A2 Land Surface Temperature product (LST) as well as the MODIS MCD43A4 Nadir Bi-directional reflectance Adjusted Surface Reflectance data (NBAR), both from Collection 5, were acquired from the Land Processes Distributed Active Archive Center (LP DAAC) using the NASA Warehouse Inventory Search Tool (WIST). The LST product is produced every 8 days at 1 km spatial resolution for day and night time temperatures and achieves an accuracy of better than 1 K (Wan 2008). The NBAR product is produced every 8 days at 500 m spatial resolution from data collected over a 16-day period (Schaaf et al. 2002). Using the MODIS reprojection tool MRT (Dwyer and Schmidt 2006) the single tiles were reprojected from the original sinusoidal projection to a geographical coordinate system, resampled to 0.0023° and mosaiced into a seamless dataset for each 8-day epoch from 2000 to 2009. Low quality pixels, such as those affected by clouds, were rejected using the data quality information supplied. Spatial gaps present in the data were interpolated with a mean filter if the gaps were not wider than five pixels. Temporal gaps were interpolated linearly, if no more than two successive 8-day periods were missing. The NBAR data were further processed to calculate the NDVI as well as the Enhanced Vegetation Index (EVI) (Huete et al. 2002).

Daily rainfall was accumulated at a 0.25° (about 25 km) grid resolution for an intercomparison study of precipitation time series from TRMM and interpolated surfaces from the Australian Bureau of Meteorology rain gauge observations (Renzullo 2008).

2.4 Seasonal bioclimatic variables

To investigate the relationship between the annual serological status of a pastoral property and environmental conditions, a number of seasonal variables were generated for each year. This step also reduced the environmental data layers to a manageable number before analysis, whilst retaining as much information about seasonal characteristics as possible (Robinson, Rogers, and Williams 1997). The aggregation periods that have been selected based on ecological principles and the compositing periods of remotely sensed data are annual (1.11.-31.10.), virus transmission season (1.11.-1.6.) summer (3.12.-25.2.), autumn (26.2.-1.6.) and the previous winter (2.6.-28.8.). Aggregated over these periods, maps of minimum, mean, maximum day and night land surface temperatures, maximum NDVI and EVI, as well as accumulated rainfall were produced. Furthermore, the time integrated NDVI, as an important phenological characteristic of the growing season, was extracted from a smoothed NDVI time series using TIMESAT (Jönsson and Eklundh 2004). Prior to statistical analysis, the 17 environmental variables were averaged over an entire pastoral property to be linked with the seroconversion data using the zonal statistics tool in ArcGIS (ESRI Inc. 2008). Due to the lack of additional data, the whole property was considered of being at equal risk for virus transmission, regardless of any factors that favour or exclude the presence of host and vector habitats. Only pixels with NDVI values below 0.1, representing surface water and bare ground, as well as built-up areas, mines, road reserves, perennial water courses and lakes, and swamps extracted from 1:250,000 topographic base data were excluded from the analysis.

2.5 Statistical model development and validation

Univariate analyses were conducted using the R statistical environment (R Development Core Team 2010) to identify those variables able to discriminate between BTV positive and

negative properties. The Kruskal-Wallis test was used to identify variables significantly associated with property BTV status. Subsequently, variables were assessed pairwise for collinearity to avoid using variables with redundant information in the same model.

A non-parametric Generalised Additive Model (GAM) approach (Hastie and Tibshirani 1990) was selected due to its flexibility and the ability to explore the often non-linear shape of the response curve for each variable efficiently. Responses to changes in environmental variables that cannot be described by linear or higher order functions are frequently found in ecological and epidemiological modelling, where e.g. rising temperature might initially increase vector activity, but after reaching an optimum temperature range, an additional increase in temperature might have an adverse effect (Gubbins et al. 2008). GAMs were fitted using the `gam()` function implemented within the `mgcv` package (Wood 2004) in R. The model building process started with a model that included all predictor variables with a Kruskal Wallis $p < 0.2$. A stepwise exclusion approach was then applied, removing the least significant variable in model until all variables were significant ($p < 0.05$). The p-value provided by the `gam()` function for each smooth term is approximate and based on the test statistic motivated by Nychka's (1988) analysis of the frequentist properties of Bayesian confidence intervals for smoothers (Wood 2004). Where variables had the same or a similar p-value, those variables that are known to be important for the BTV cycle were retained in the model. Hence, the model was developed utilising expert knowledge as well as statistics. This minimises the risk of potentially using relationships which may be statistically significant but have no grounding in the ecological process being modelled. Several models were developed using this hybrid approach and the best model was chosen by maximising the discriminatory ability, given by the area under the receiver operator characteristic (ROC) curve (Pearce and Ferrier 2000).

The effects of spatial autocorrelation have not been considered when building the fixed effects GAM described in this study. A typical approach to identify unaccounted for second-order spatial effects, is to examine the model residuals for evidence of spatial autocorrelation. If there is evidence for spatial autocorrelation, the model can be extended to account for the spatial dependence between sites. The peculiarities of the data that are dealt with in this study made application of this approach difficult. Firstly, due to the large distances between sample sites it is unlikely that spatial autocorrelation was present. Secondly, the samples that have been used for model building describe the BTV status over several years and some stations have multiple representations. Consequently, residuals would have to be analysed for each year separately. In the case of evidence for autocorrelation and considering that the effects will vary between years, distribution models would be needed to be developed on an annual basis. Given that the fixed effects model has originally been built from eight years of data comprising only 162 samples, reducing the sample size to between ten (season 2005/2006) and 25 samples annually (season 2002/2003) would lead to instability in the model. It was therefore concluded that the data available in this study would not be sufficient to support an investigation of spatial autocorrelation. The previously developed stable and ecologically meaningful model is therefore used to generate annual prediction maps.

In contrast to the station average variables used for model development, prediction maps for the seasons 2000/2001 through to 2008/2009 were generated on a pixel basis, using the annual bioclimatic variable raster maps. In order to process data for the whole study area, the nominal spatial resolution was reduced to 1 km to avoid computer memory limitations. It should be noted that areas outside the bounds of pastoral properties are unlikely to support virus survival due to the lack of cattle as the major group of hosts. Nevertheless, area-wide prediction maps have been generated to delineate areas where the environmental conditions might in principle provide suitable host and vector habitats. Although there is no evidence that native animals are involved in BTV transmission, recent laboratory experiments have demonstrated that camels, which are plentiful in central Australia, may act as hosts for BTV (Batten et al. 2011). Knowing about areas with a higher probability for virus transmission also aids in logistic planning of livestock movement and if necessary, precautionary measures can be taken to avoid contact between hosts and potentially infected vectors.

External validation of the model was conducted with data from the season 2008/2009, which were not used for model development. Model validation was performed in two ways. Firstly, we predicted the probability for observing BTV seroconversion in 2008/2009 based on station average variables (which weren't included in model development) and then assessed the model predictions

against the serological test results from that season. The second method took the following approach. Since model output was expressed in terms of BTV probability for each 1 km × 1km pixel that comprised the study area, it was of interest to determine what cut point should be used to classify locations as BTV positive. For the model developed using data for 2000/2001 to 2007/2008, the sensitivity and specificity of model predictions as a function of a series of candidate cut points ranging from 0 to 1 was plotted and the cut point that maximised both sensitivity and specificity identified. Predictions of BTV probability were then made using the 2008/2009 data. All pixels with a predicted BTV probability greater than the identified cut point of 0.6 were then classified as positive and compared with the observed BTV status for each station for 2008/2009.

3 RESULTS

All of the 17 bioclimatic variables initially considered for model development showed significant difference between positive and negative sites with a Kruskal Wallis $p < 0.05$. Although only observable at the least significant digits of the p-value, the strongest difference was found for the vegetation and rainfall variables. Strong correlations were generally found between vegetation and rainfall variables and also between seasonal and summer rain, indicating that most rainfall occurs during the summer months. The strongest correlation was present between maximum NDVI and maximum EVI variables for any aggregation period, which meant that only one of these two variables should be included in the model. In relation to temperature, the highest correlations were found between the summer and autumn day temperature variables.

The GAM with the best ability to discriminate between BTV positive and negative stations, with an area under the ROC curve of 0.8644, is summarised in Table 2. In this model, maximum seasonal NDVI (*maxndviss*), mean day temperature of the warmest 8-day period in autumn (*maxlstdau*), mean day temperature in summer (*meanlstdsu*) and mean night temperature of the previous winter months (*meanlstnpw*) are the predictor variables.

Table 2: Significance of smooth (*maxndviss*, *maxlstdau*, *meanlstnpw*) and parametric (*meanlstdsu*) model terms in the BTV distribution model

Model terms	Df/edf	Chi-square	p-value
s(maxndviss)	3.001	11.001	0.0117*
s(maxlstdau)	3.254	9.551	0.0282*
s(meanlstnpw)	7.015	7.699	0.0241*
meanlstdsu	1	5.087	0.0416*

*: The variable is statistically significant ($p \leq 0.05$)

The GAM plot (Figure 2) shows the smoothers for these four covariates and the 95% confidence interval. The horizontal axis of each plot shows the density of samples for each value of each covariate in the model. The plots show that the effect of one of the four variables on the outcome can be expressed by a linear function (*meanlstdsu*), while the other three variables were best described by a unimodal smoothing function (*maxndviss*, *maxlstdau*, and *meanlstnpw*). Model coefficients and the shape of the smooth functions indicate that the probability of BTV presence is generally higher in areas with high NDVI, although this relationship is not linear. High average summer day temperatures on the other hand decrease the probability of BTV presence, which could be explained by the risk of desiccation, particularly in the southern arid part of the Northern Territory. Maximum autumn day temperatures, or more precisely the average day temperature of the warmest 8-day period in autumn, had a positive effect on BTV presence up to about 40°C, but higher temperatures reduced the likelihood of virus activity. Higher mean day temperatures in winter were associated with BTV detection.

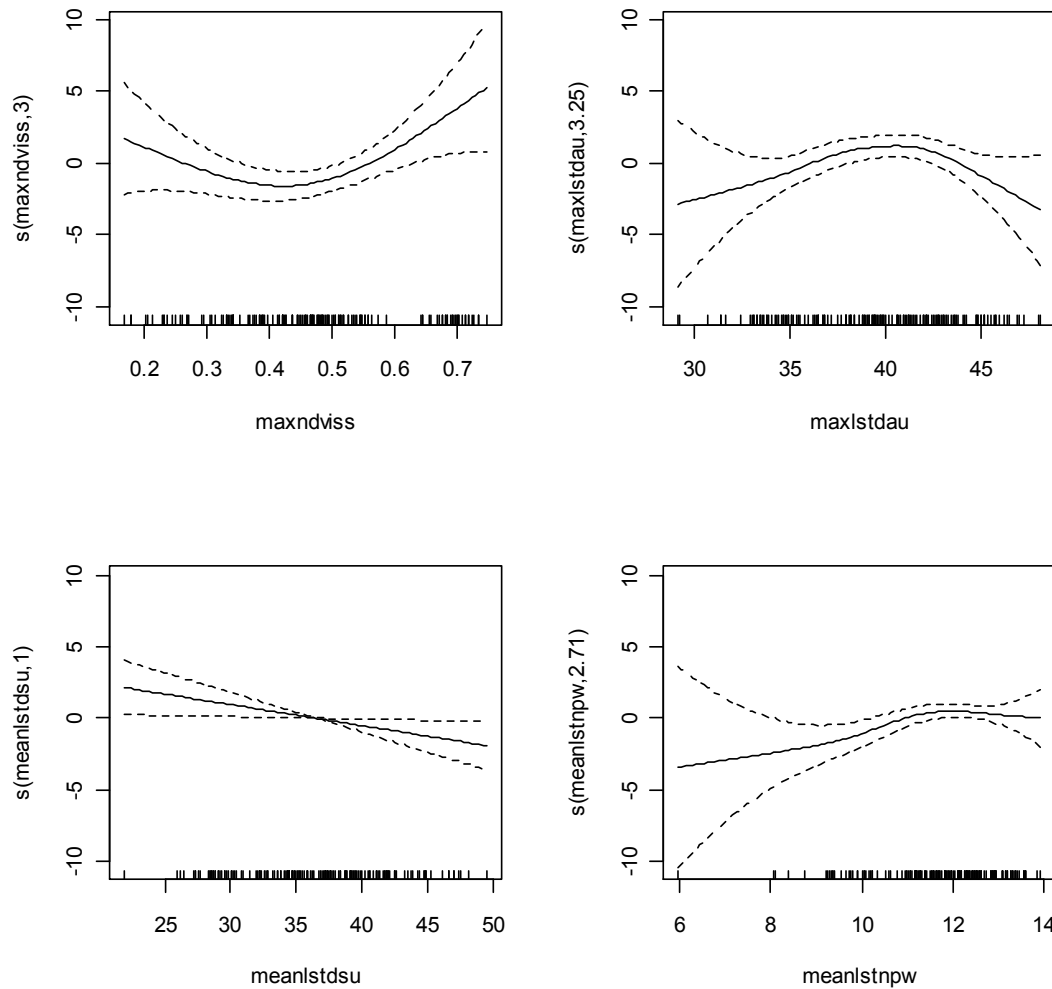


Figure 2: Smoothed fits to the variables *maxndviss*, *maxlst dau*, *meanlst dau* and *meanlst npw* used in the final BTV distribution model for the Northern Territory (solid line). Dashed lines give 95% confidence intervals.

The maps in Figure 3 show, for each year, the estimated probability that cattle tested in an area will be BTV seropositive. Superimposed on each map are the points positioned at the centroid of each station that took part in testing, colour coded according to BTV status. Predictions from the GAM model consistently spatially replicated the distribution of BTV in cattle tested between 2000 and 2009. In particular, the season 2004/2005, for which dry and hot conditions were experienced throughout the study area, with associated low virus activity, demonstrates the predictive capabilities of the model. The general pattern of decreasing probability of BTV presence from North to South is present throughout the years analysed, but the divide between high and low probability varies between the years and largely follows the observed BTV presence. Also, smaller patches of high and low probability are depicted well (e.g. season 2003/2004) and are confirmed by the test results. Comparison of predicted and observed BTV occurrence also reveals that some positive test results are found in areas of low predicted probability. However, often there is a small area within a cattle station with a high probability. Without the knowledge of the true geographical origin of the tested animals, one can assume that they could have come from these areas. It is therefore recommended that the maximum observed probability be used as a guideline, when deciding if a property is to be declared BTV positive or negative.

External validation based on data from 38 pastoral stations tested between November 2008 and October 2009 resulted in an area under the curve (AUC) of 0.8215, which again demonstrated that the model had good discriminatory capabilities based on the classification of AUC values by Hosmer and Lemeshow (2000). A number of measures have been derived from the classified presence/absence maps using both the station average environmental variables as well as the

maximum predicted probability per station. As can be seen in Table 3, using the station average variables for predictions on a station level resulted in the highest overall accuracy (measured by Cohen's Kappa). However, using the maximum predicted probability per station decreased the likelihood of false negative classifications, which is a crucial factor when the objectives of a surveillance program (as in this case) was to detect seroconversion, particularly in areas that have previously been free of virus activity.

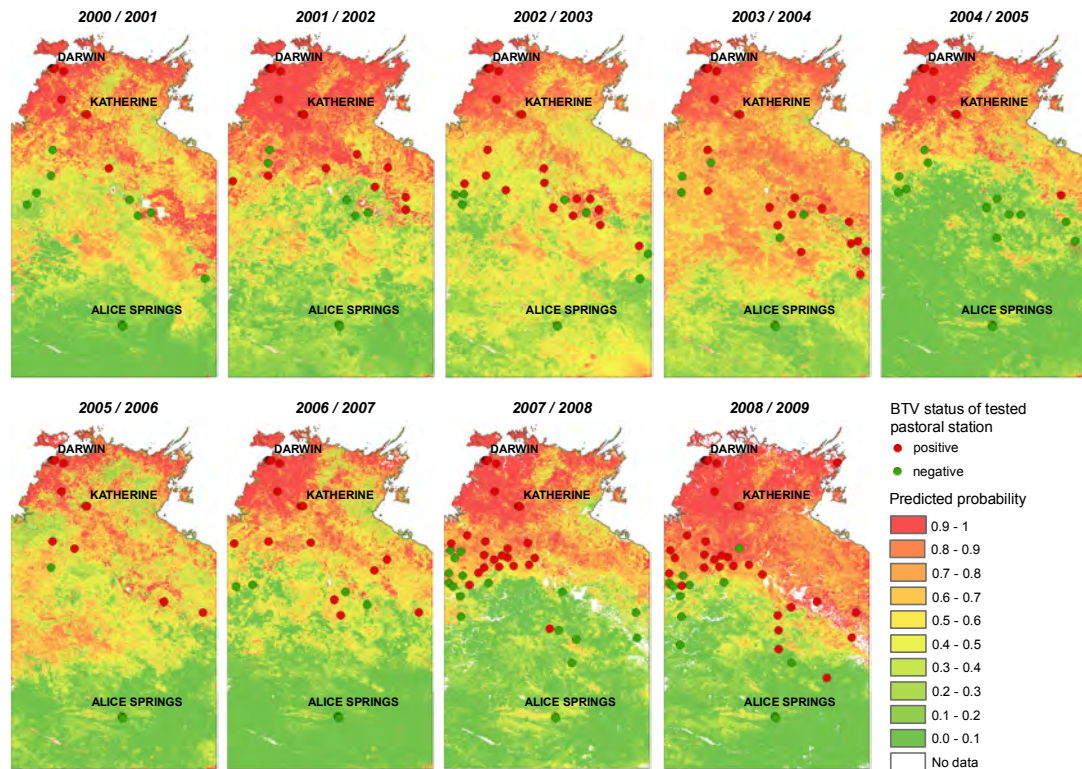


Figure 3: Predicted BTV seropositivity in the NT for the seasons 2000/2001 to 2008/2009 in relation to the surveyed presence/absence data

Table 3: Accuracy measures for the classified predicted presence/absence maps

Measure	Predictions from station average variables	Maximum predicted probability per station
Correct classification rate	78.9%	76.3%
Sensitivity	81.5%	100%
Specificity	72.7%	18.2%
False positive rate	27.3%	81.8%
False negative rate	18.5%	0%
Positive predictive power	88%	75%
Negative predictive power	61.5%	100%
Cohen's Kappa	0.514	0.24

4 CONCLUSION

This study presents a novel application of remotely sensed environmental variables used in a generalised additive model to predict BTV seropositivity in Northern Australia. The selected environmental data and the modelling approach are well suited to reproduce ground observations. However, the model can only be as good as the input data, and as such, the results of this study are seen as complimentary to the data produced by conventional surveillance. What this study has demonstrated well, is that the spatial gaps between the sampled pastoral properties may be filled

to provide a more complete picture of potential virus activity. To a limited extent, the model can also be used to forecast BTV seropositivity. The fact that the model uses data collected during the previous winter as well as the summer and autumn months means that after data acquisition and processing, a prediction can be generated at the end of June each year. Depending on the timing of the annual muster, during which cattle are accessible for blood sampling, most data collected as part of serological surveys are not available in the NAMP database earlier than the predictions. The model therefore has the capability of acting as an early indicator for potential virus activity hotspots to indicate where monitoring should be intensified.

Model building should not be considered a static process. As seen in Europe, where distribution models for BTV vectors have been built using long term climate records (e.g. Tatem et al., 2003), the predictions are soon outdated and therefore less likely to predict incursion of new virus strains and vectors and their vast expansion under a changing climate. This happened recently with BTV 8 (Mellor et al. 2009). A similar scenario is also likely to happen in Australia, where the windborne incursion of vectors infected with a new BTV strain from Southeast Asia pose an ongoing threat to livestock. While surveillance in northern Australia is able to detect these incursions, it is important to constantly monitor the environmental conditions that may create favourable habitats for BTV vectors in regions beyond the current endemic zone. We suggest using the increasing archive of virus and environmental data to gradually recalibrate the model on an annual basis. It is anticipated that the significance of single parameters may change with the availability of a broader data base, e.g. due to the effects of climate change.

This research has demonstrated the limitations of using data from an operational surveillance system that has been designed predominantly to fulfil international biosecurity requirements, such as NAMP. In order to facilitate more in-depth spatio-temporal studies of virus environmental relationships based on NAMP data, the following improvements are recommended:

- Record spatial information more accurately, including the location of vector traps, bleeding sites and, if known, the location(s) at which tested cattle have been held prior to bleeding.
- Maintain a core network of sentinel herds in strategic locations which are tested at regular intervals. This will provide a continuous record of virus activity data and facilitate comparison of inter-annual variations. Bluetongue surveillance in the Northern Territory, based on government research stations, provides an example of such a network.
- Select sites for serological surveys, where possible, based on perceived risk (e.g. informed by previous survey results and modelled probability).

Presently, the study focuses on environmental variables that can be derived in a timely and efficient manner using largely pre-processed and validated datasets. These are not available for all factors known to be associated with BTV. Further research is necessary to investigate other factors that might be associated with virus presence. Of particular interest are the effects of humidity, air temperature, landscape structure, and cattle density. Potential data sources need to be identified and assessed. Potential for future research has been identified in the 2.5 decades of data from the Advanced Very High Resolution Radiometer (AVHRR). These could be analysed to make full use of the long records of epidemiological data from more than 30 years of arbovirus surveillance on Australia. Also the definition of alternative aggregation periods for the bioclimatic variables is worth investigating.

ACKNOWLEDGEMENTS

This research was funded by the Australian Biosecurity Cooperative Research Centre for Emerging Infectious Disease (AB-CRC), Project 3.036R, represented by the project partners Curtin University of Technology, Department of Agriculture and Food Western Australia, Berrimah Veterinary Laboratories, Northern Territory Department of Primary Industries, Fisheries and Mines (DPIFM), University of Western Australia (Arbovirus Surveillance and Research Laboratory) and the Mosquito Borne Disease Control Unit, WA Department of Health. We want to acknowledge all contributors of the National Arbovirus Monitoring Program for access to the surveillance data and scientific advice. We also want to thank Luigi Renzullo from CSIRO Land and Water, Canberra for providing the TRMM 3B42 daily rainfall accumulates for the entire study period.

REFERENCES

- Animal Health Australia. 2001. *The history of bluetongue, Akabane and ephemeral fever viruses and their vectors in Australia*. Deakin, ACT, Australia: Animal Health Australia.
- Animal Health Australia. 2008. *National Arbovirus Monitoring Program - 2007-2008 Report*. Australian Animal Health Council.
- Animal Health Australia. 2009. *National Arbovirus Monitoring Program - 2008-2009 Report*. Australian Animal Health Council.
- Batten, C. A., B. Harif, et al. 2011. Experimental infection of camels with bluetongue virus. *Research in Veterinary Science* 90 (3): 533-535.
- Baylis, Bouayoune, et al. 1998. Use of climatic data and satellite imagery to model the abundance of *Culicoides imicola*, the vector of African horse sickness virus, in Morocco. *Medical and Veterinary Entomology* 12 (3): 255-266.
- Calvete, C., R. Estrada, et al. 2009. Ecological correlates of bluetongue virus in Spain: Predicted spatial occurrence and its relationship with the observed abundance of the potential *Culicoides* spp. vector. *The Veterinary Journal* 182 (2): 235-243.
- Cameron, A. R. 2000. Development of a tempero-spatial disease distribution model for arboviruses and their vectors in Australia. *Proceedings of the IXth Symposium of the International Society for Veterinary Epidemiology and Economics (ISVEE)*, Breckenridge, Colorado, USA, 6-11 August, 2000.
- Cameron, A. R. 2004. Data management and analysis systems for bluetongue virus zoning in Australia. *Veterinaria Italiana* 40 (3): 365-368.
- De La Rocque, S., V. Michel, et al. 2004. Remote sensing and epidemiology: examples of applications for two vector-borne diseases. *Comparative Immunology, Microbiology and Infectious Diseases* 27 (5): 331-341.
- Dwyer, J., and G. Schmidt. 2006. The MODIS Reprojection Tool. In *Earth Science Satellite Remote Sensing*, ed. J. J. Qu, W. Gao, et al., 162-177. Springer Berlin Heidelberg.
- Ebert, E. E., J. E. Janowiak, et al. 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society* 88 (1): 47-64.
- ESRI Inc. 2008. *ArcGIS, version 9.3* (program).
- Gubbins, S., S. Carpenter, et al. 2008. Assessing the risk of bluetongue to UK livestock: Uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number. *Journal of the Royal Society Interface* 5 (20): 363-371.
- Guis, H., A. Tran, et al. 2007. Use of high spatial resolution satellite imagery to characterize landscapes at risk for bluetongue. *Veterinary Research* 38: 669-683.
- Hastie, T. J., and R. J. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.
- Hay, S. I. 2000. An overview of remote sensing and geodesy for epidemiology and public health application. In *Remote Sensing and Geographical Information Systems in Epidemiology, Advances in Epidemiology* 47, ed. S. I. Hay, S. E. Randolph, et al., 1-35. London: Academic Press.
- Hay, S. I., A. Graham, et al., eds. 2006. *Global Mapping of Infectious Diseases: Methods, Examples and Emerging Applications, Advances in Parasitology* 62. London: Academic Press.
- Hosmer, D. W., and S. Lemeshow. 2000. *Applied logistic regression*. 2nd ed, *Wiley series in probability and statistics Texts and references section*. New York: Wiley.
- Huete, A., K. Didan, et al. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83 (1-2): 195-213.

- Jönsson, P., and L. Eklundh. 2004. TIMESAT--a program for analyzing time-series of satellite sensor data. *Computers & Geosciences* 30 (8): 833-845.
- Kirkland, P. D. 2004. Bluetongue viruses, vectors and surveillance in Australia – the current situation and unique features. *Veterinaria Italiana* 40 (3): 47-50.
- Klingseisen, B. 2010. Spatio-temporal modelling of bluetongue virus distribution in Northern Australia based on remotely sensed bioclimatic variables. PhD Thesis, Department of Spatial Sciences, Curtin University of Technology, Perth, Western Australia
- Mellor, P. S., S. Carpenter, et al. 2009. Bluetongue in Europe and the Mediterranean Basin. In *Bluetongue*, 235-264. London: Academic Press.
- Melville, L. F. 2004. Bluetongue surveillance methods in an endemic area: Australia. *Veterinaria Italiana* 40 (3): 184-187.
- Murray, M. D. 1995. Influences of vector biology on transmission of arboviruses and outbreaks of disease: the *Culicoides brevitarsis* model. *Veterinary Microbiology* 46 (1-3): 91-99.
- Murray, M. D., and H. A. Nix. 1987. Southern Limits of Distribution and Abundance of the Biting-Midge *Culicoides-brevitarsis* Kieffer (Diptera, Ceratopogonidae) in Southeastern Australia - an Application of the Growest Model. *Australian Journal of Zoology* 35 (6): 575-585.
- Nychka, D. 1988. Bayesian confidence intervals for smoothing splines. *Journal of the American Statistical Association* 83: 1134-1143.
- Oliver, G. R. 2004. Bluetongue trade issues – an Australian perspective. *Veterinaria Italiana* 40 (4): 688-690.
- Ostfeld, R. S., G. E. Glass, et al. 2005. Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology & Evolution* 20 (6): 328-336.
- Pearce, J., and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133 (3): 225-245.
- Peel, M. C., B. L. Finlayson, et al. 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* 11 (5): 1633-1644.
- Purse, B. V., B. J. J. McCormick, et al. 2007. Incriminating bluetongue virus vectors with climate envelope models. *Journal of Applied Ecology* 44 (6): 1231-1242.
- Purse, B. V., A. J. Tatem, et al. 2004. Modelling the distributions of *Culicoides* bluetongue virus vectors in Sicily in relation to satellite-derived climate variables. *Medical and Veterinary Entomology* 18 (2): 90-101.
- R Development Core Team. 2010. *R: A Language and Environment for Statistical Computing*. Vienna, Austria.
- Renzullo, L. 2008. Considerations for the blending of multiple precipitation data sets for hydrological applications. *Proceedings of the 4th International Precipitation Working Group Workshop*, Beijing, China, 13-17 October 2008.
- Robinson, T., D. Rogers, et al. 1997. Univariate analysis of tsetse habitat in the common fly belt of Southern Africa using climate and remotely sensed vegetation data. *Medical and Veterinary Entomology* 11 (3): 223-234.
- Schaaf, C. B., F. Gao, et al. 2002. First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment* 83 (1-2): 135-148.
- Scharlemann, J. P. W., D. Benz, et al. 2008. Global Data for Ecology and Epidemiology: A Novel Algorithm for Temporal Fourier Processing MODIS Data. *PLoS ONE* 3 (1): e1408.
- Tatem, A. J., M. Baylis, et al. 2003. Prediction of bluetongue vector distribution in Europe and north Africa using satellite imagery. *Veterinary Microbiology* 97 (1-2): 13-29.
- Tatem, A. J., S. J. Goetz, et al. 2004. Terra and Aqua: new data for epidemiology and public health. *International Journal of Applied Earth Observation and Geoinformation* 6 (1): 33-46.

- Wan, Z. 2008. New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products. *Remote Sensing of Environment* 112 (1): 59-74.
- Ward, M. P., and M. C. Thurmond. 1995. Climatic factors associated with risk of seroconversion of cattle to bluetongue viruses in Queensland. *Preventive Veterinary Medicine* 24 (2): 129-136.
- Wilson, B. A., P. S. Brocklehurst, et al. 1990. *Vegetation Survey of the Northern Territory, Australia. Technical Report No. 49*. Darwin: Conservation Commission of the Northern Territory. http://nt.gov.au/nreta/natres/natveg/vegmapping/pdf/ntveg_rpt.pdf (accessed
- Wittmann, E. J., P. S. Mellor, et al. 2001. Using climate data to map the potential distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in Europe. *OIE Revue Scientifique et Technique* 20 (3): 731-740.
- Wood, S. N. 2004. Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models. *Journal of the American Statistical Association* 99 (467): 673-686.

Innovative applications of remotely sensed evapotranspiration and vegetation cover

Kathryn Sheffield*, Mohammad Abuzar

Department of Primary Industries
Parkville, Victoria, Australia

*kathryn.sheffield@dpi.vic.gov.au

Des Whitfield, Andy McAllister, Mark O'Connell, Lexie McClymont

Department of Primary Industries
Tatura, Victoria, Australia

ABSTRACT

Competing demands for a limited water supply has led to a need for more effective regional water management to improve the balance between environmental and commercial water uses within catchments. Improved water management will be achieved with a greater understanding of how much water is used, and how much water is required, by various agricultural enterprises and native vegetation. The work presented in this paper was done to provide this type of required information and to fulfil a need for affordable and objective monitoring and reporting tools for water and vegetation status across landscapes.

This project developed and tested the Department of Primary Industries SEBAL-METRIC approach (DPI SM) to calculate surface energy balance components using satellite imagery. This approach was essentially based on the Mapping Evapotranspiration with Internalized Calibration (METRIC) approach (Allen et al., 2007), which is a variation of the Surface Energy Balance Algorithm for Land (SEBAL) approach (Bastiaanssen et al., 1998; Bastiaanssen et al., 2005). Evapotranspiration (ET) and vegetation cover (Normalized Difference Vegetation Index: NDVI) were calculated from Landsat 5 TM satellite imagery. The results are presented in terms of the Triangular Irrigated Crop Response Framework (TICRF), based on data of Tasumi et al (2005). The TICRF provided a basic structure for the interpretation and analysis of multiple sources of ET-NDVI data. It was also used to identify well-watered crops, which can then form the basis for crop and region specific crop water requirement calculations.

The DPI SM approach has been applied to a diverse range of land uses within major irrigation regions of south-east Australia. ET-NDVI relationships, in the TICRF context, showed that most irrigated crop ET observations were confined to the TICRF. This paper presents selected case studies which illustrate the application of the DPI SM approach in Victoria, Australia. The finding that crops showed a linear relationship between ET and NDVI implies that knowledge of crop type and weather is insufficient for effective irrigation water management. This project has shown that vegetation cover provides a more explicit guide to crop water requirement than knowledge of the crop type alone.

The methods presented are affordable, comprehensive and repeatable in comparison with other approaches of quantifying plant water use. Results from this project have improved understanding of the relative roles of crop type, vegetation cover, regional weather and irrigator behaviour in water management.

KEYWORDS: Regional water use, NDVI, Landsat 5 TM, irrigated agriculture

1 INTRODUCTION

Competing demands for a limited water supply has led to a need for more effective farm and regional water management to improve the balance between environmental and commercial water uses within catchments. Improved water management will be achieved with a greater understanding of how much water is used, and how much water is required, by various agricultural enterprises and native vegetation. However, there is a lack of knowledge regarding water use and water requirement of a wide range of irrigated crops and native vegetation in south-east Australia.

Water availability is a major limiting factor for crop growth and yields of crops grown in Australia. There is a lack of objective technical data that describe optimal water requirement for the range of irrigated crops grown in south-east Australia. Similarly, landscape-scale information relating to water requirements of native vegetation, particularly within riparian systems is lacking. This information is required for an understanding of hydrology and vegetation dynamics within riparian zones, which is important for the development of ecologically appropriate and effective management plans for native systems (Chong and Ladson, 2003; Lawrence and Colloff, 2008). The work presented in this paper was done to provide this type of required information and to fulfil a need for affordable and objective monitoring and reporting tools for water and vegetation status across landscapes.

This study assumes that measures of evapotranspiration (ET) and vegetation cover, derived from satellite imagery, can provide information that is helpful to agricultural and natural resource management areas. Satellite remote sensing of ET is proposed as an innovative and affordable means of acquiring comprehensive water use data, specific to enterprise type and region. The use of satellite remote sensing in this way extends the application of established scientific principles of vegetation water use to a wide range of enterprises, and also extends these principles from farm to regional and national scales using a consistent, comparable method.

Calculations of surface energy balance and ET, using satellite remote sensing data, became more widely used following the publication of the Surface Energy Balance Algorithm for Land (SEBAL) (Bastiaanssen et al., 1998; Bastiaanssen et al., 2005). The model was designed for use in data poor landscapes, and overcame some of the issues associated with previous approaches (Bastiaanssen et al., 1998). The success of the SEBAL approach is strongly associated with the use of hot and cold 'anchor' pixels within the image, which are used to constrain estimates of H. SEBAL is based on the assumption that LE is equal to zero at a selected hot pixel in an image scene, and that H is equal to zero at a selected cold pixel in an image scene.

Allen et al (2007) adapted the SEBAL approach for irrigation applications, taking advantage of data available from automatic weather stations located in agricultural areas. This was published as the Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration model (METRIC). Allen et al (2007) redefined the cold anchor pixel in terms of reference ET, which aligned METRIC with estimates of crop water requirements as calculated by standardised methods published in the Food and Agricultural Organization of the United Nations Irrigation and Drainage Paper No 56 (FAO 56) (Allen et al., 1998). The METRIC model defines the cold pixel as a well-irrigated alfalfa or lucerne crop with near complete vegetation canopy cover, where LE is equal to reference ET (Allen et al., 2005). Reference ET can be calculated as either 'short' grass reference crop ET (ET_0) or 'tall' grass reference crop ET (ET_r) (Allen et al., 1998).

2 METHODS

To provide information relevant to water management, measures of vegetation cover (Normalized Difference Vegetation Index: NDVI) and water use (ET) were derived from satellite remote sensing data. In this study, Landsat 5 TM imagery was used as it provides regularly acquired imagery containing both optical and thermal information. Ground-based weather data from weather stations located near our study areas was sourced from www.weatherzone.com.au.

2.1 Satellite-based estimates of ET and NDVI

NDVI was derived using red and near-infrared (NIR) reflectance. The index has been shown to respond to levels of green biomass, leaf chlorophyll content, and vegetation water stress (Jensen, 2000; Liang, 2004). It is calculated as:

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red}). \quad (1)$$

Satellite-based estimates of ET were calculated by deriving ET from estimations of components of the surface energy balance. The surface energy balance equation is given as:

$$\text{LE} = \text{Rn} - \text{H} - \text{G}. \quad (2)$$

In Equation 2, G is the soil heat flux, H is the sensible heat flux, Rn is the net radiation flux at the surface, and LE is the latent heat flux. LE represents the residual energy available for ET.

This project developed and tested the Department of Primary Industries SEBAL-METRIC approach (DPI SM) to calculate surface energy balance components using Landsat 5 TM imagery (Whitfield et al., 2010a; Whitfield et al., 2010b). This approach was essentially based on the METRIC approach, as METRIC retains the essential features of the SEBAL algorithm and was customised for irrigation applications. Details of the algorithms used are documented in Allen et al (2005) and Allen et al (2007). Requirements for hot and cold anchor pixel selection used in the DPI SM are listed in Table 1. In both cases, the selected pixel was located within a contiguous area, away from an edge or boundary. The cold pixel selected was usually covered by an irrigated summer crop such as lucerne, tomatoes, or carrots. The DPI SM uses the surface roughness relationship published in Teixeira et al (2009) to account for land cover and vegetation height variation across the landscape.

Table 1. Hot and cold anchor pixel selection requirements for DPI SM

Hot anchor pixel	Cold anchor pixel
<ul style="list-style-type: none">▪ Dry, hot and bare land surface▪ Negligible ET ($\text{ET} = 0$)▪ NDVI values 0.05 – 0.15▪ Land surface temperature in the upper 2.5% of image values▪ Located within a contiguous area, not adjacent to an edge or boundary associated with a change in land condition	<ul style="list-style-type: none">▪ Well-watered, full cover crop of known type▪ Vegetation height 0.4-0.6m▪ $\text{ET} = \text{ET}_r$▪ NDVI values > 0.70▪ Land surface temperature in the lower 2.5% of image values▪ Located within a contiguous area, not adjacent to an edge or boundary associated with a change in land condition

2.2 Analysis of ET and NDVI data

ET and NDVI for different land uses within major irrigation regions in south-east Australia were calculated. ET values are presented as fraction of reference ET (ETrF) calculated as ET divided by ET_r (Allen et al., 2005). This calculation gives a standardised measure of ET across differing land uses and image acquisition dates. ETrF values of approximately one indicate near-maximal ET rates. ETrF values of near zero represent little or no ET. ETrF values greater than one indicate ET rates exceeding the expected maximal level, generally due to wet surface soil conditions or the presence of surface water.

The results are also presented in terms of the Triangular Irrigated Crop Response Framework (TICRF), based on ET-NDVI responses of irrigated broadacre crops in Idaho, USA, reported by Tasumi et al (2005) and shown in Figure 1. The TICRF provides a basic structure for the interpretation and analysis of multiple sources of ET-NDVI data. The TICRF can also be used to identify well-watered crops which can then form the basis for crop and region specific crop water

requirement calculations. Tasumi et al (2005) showed that ET-NDVI relationships can be used to partition transpiration and soil evaporation using the following:

$$K_c = K_{cb} + K_e. \quad (3)$$

In Equation 3, K_c is the basal crop coefficient, accounting for both crop and surface evapotranspiration. K_{cb} is the crop coefficient accounting for crop transpiration, while K_e represents evaporation from the soil surface.

The TICRF describes the lower and upper bounds on NDVI in irrigated crops, the theoretical upper limit to irrigated crop ET, and a lower line that describes when crops are short of water (irrigation refill line, also known as a K_{cb} line). The irrigation refill line describes the lower, NDVI-dependent threshold of crop water use and can be used as a trigger for irrigation in farm management practices. An ETrF offset from the irrigation refill line was also calculated as:

$$dK_{cb} = ET/ET_r - K_{cb} \quad (4)$$

Statistics derived from the crop- and date-specific ET-NDVI responses are important to the derivation of objective estimates of the NDVI-dependent crop coefficients, and the NDVI-dependent irrigation refill line, implied by ET-NDVI responses. Mean values of dK_{cb} describe the sample mean ETrF displacement from the irrigation refill line. A 95% confidence interval on dK_{cb} was estimated as $dK_{cb} \pm 1.96$ standard error. NDVI-dependent K_c values are estimated by K_{cb} , supplemented by the mean value of dK_{cb} , subject to the conditions that most observations of dK_{cb} are positive. Whilst these summer crops (with a vegetation height of approximately 0.5m) provided a generic source of cold pixel candidates for the DPI SM approach, it was necessary to adapt the upper limit of the TICRF in order to allow its application to dairy pastures, where maximum rates of ET were described by ET_0 . This analysis allowed irrigation and rainfall inputs to systems to be matched with the ET capability of different land uses.

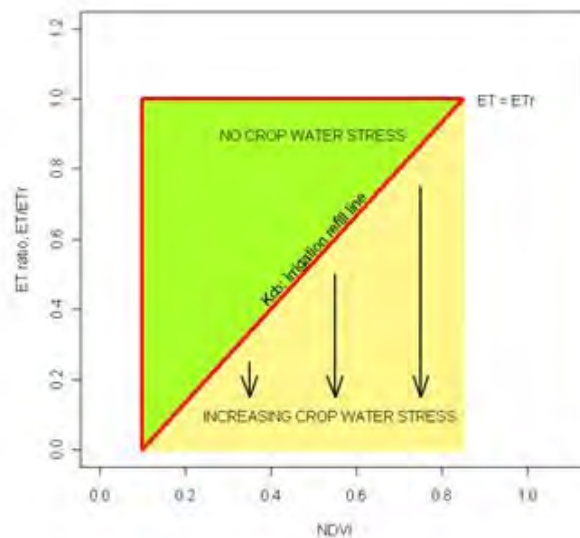


Figure 1. Diagram illustrating the Triangular Irrigated Crop Reference framework (TICRF), based on ET-NDVI responses of irrigated broadacre crops, reported by Tasumi et al (2005). The TICRF is represented the upper green zone enclosed by the red triangle. The diagonal line corresponds with the basal crop coefficient, K_{cb} . The K_{cb} line marks the onset of water stress in crops, and thereby describes the need for irrigation to replenish soil water. The yellow zone represents samples subject to increasing water stress.

3 CASE STUDIES

The DPI SM approach has been applied to a diverse range of land uses within major irrigation regions of south-east Australia. These include:

- Sunraysia Irrigation Region (vines, almonds, citrus)
- Shepparton Irrigation Region (perennial pastures, orchards, summer crops)
- Murrumbidgee Irrigation Area (citrus, vines, rice)
- South Australian Riverland (vines, almonds, citrus)
- Namoi Irrigation Region (cotton)
- Murray River corridor (Barmah Forest, Hattah Lakes)

ET-NDVI relationships, in the TICRF context, showed that most irrigated crop ET observations (>90%) were confined to the TICRF. This suggests that soil water regimes were appropriate in most regions and for most crop types. Maximum ET rates of irrigated crops were consistently described by ET_r . The exception to this was perennial pastures grown on dairy properties in the Shepparton Irrigation Region, where ET_0 provided a more appropriate estimate of reference crop ET.

The following section of this paper presents selected case studies which illustrate the application of the DPI SM approach in Victoria, Australia.

3.1 Perennial pasture on dairy farms in the Shepparton Irrigation Region

Dairy farms in the Shepparton Irrigation Region, Victoria, were identified from local government land use assessments. Within these identified properties, areas of irrigated perennial pasture were mapped using Landsat 5 TM data, following methods outlined in Abuzar et al (2008). Perennial pasture associated with dairy properties was identified as having active vegetation growth during spring, summer and autumn seasons. ET and NDVI results presented here were extracted from four clear-sky Landsat 5 images acquired during the irrigation season (September 2008 – April 2009). Scatterplots of NDVI and ET values extracted from the imagery are shown in Figure 2.

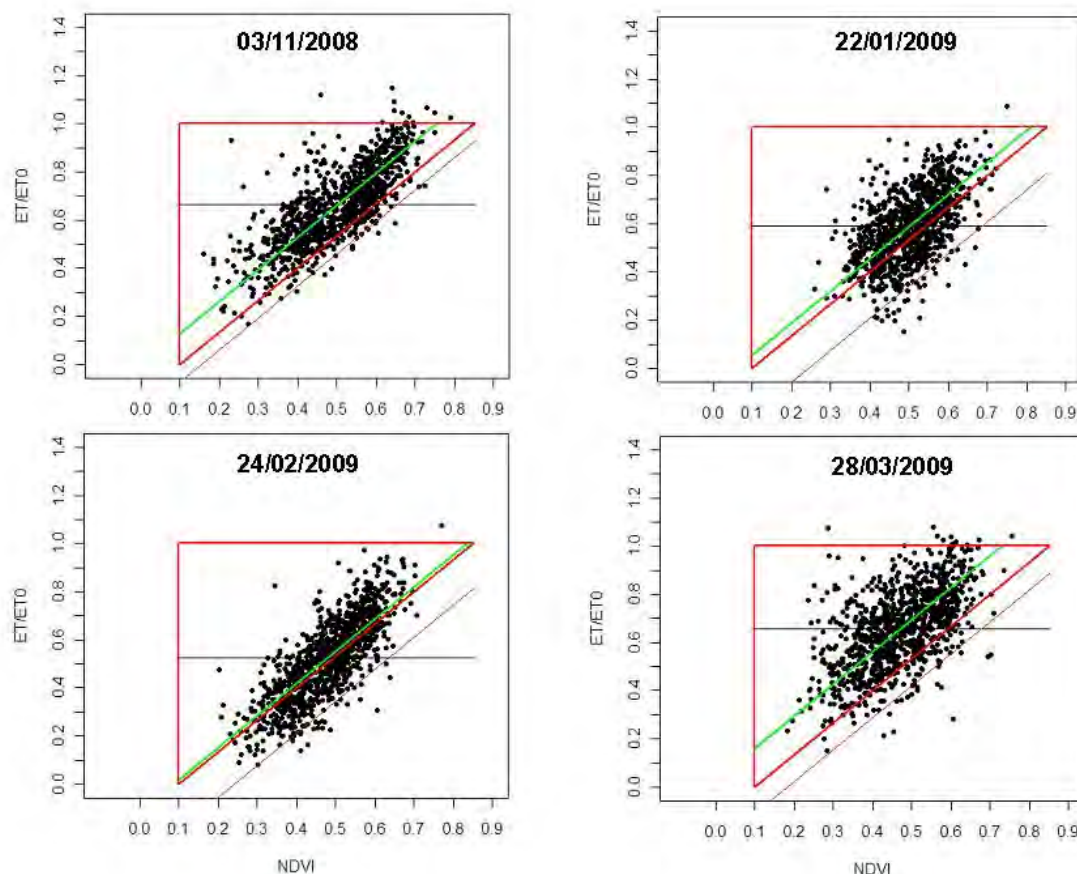


Figure 2. Scatterplots of NDVI and ET values for perennial pastures located on dairy properties in the Shepparton Irrigation Region, derived from Landsat 5 TM images acquired on four different dates. The results are presented in terms of the TICRF, which is depicted by the red triangle on each graph. The TICRF has been revised to ensure a maximum rate of ET corresponds with ET_0 . Green lines show mean ET displacement (dKcb) from the irrigation refill line. Brown lines show the lower 95% confidence interval of dKcb. Horizontal black lines show mean ET/ ET_0 ratio (ETrF)

Figure 2 shows that ET of perennial pastures reached a level equivalent to ET_0 in November 2008, where NDVI values exceeded 0.65. At this time, most ET observations exceeded the Kcb limit associated with the TICRF. However, data obtained from images acquired during summer (January and February) showed lower maximal ET rates and more observations located below the irrigation refill line. These images were acquired at a time when evaporative demand is typically at a maximum in south-east Australia. The mean ET/ ET_0 ratio fell from 0.66 in November, to 0.59 and 0.52 in January and February respectively. The mean ET/ ET_0 ratio increased to 0.66 in March 2009, following rainfall in February and March (4mm and 21mm respectively).

A comparison of ET observations with the Kcb irrigation refill line of the TICRF in November 2008 suggests that most pastures were adequately supplied with water. In summer and autumn, observations were increasingly in the suboptimal range implied by the TICRF. A linear relationship between ET and NDVI is shown, which can provide information for irrigation management in addition to weather and crop type knowledge. This relationship, in addition to analysis with reference to the TICRF can provide information leading to more effective irrigation water management.

3.2 Sunraysia Irrigation Region – perennial horticultural crops

The DPI SM approach was applied to known horticultural crops in the Sunraysia Irrigation Region, located near Mildura in Victoria. Land use data were used to identify locations of vines and citrus orchards. Grape production dominated the horticultural land uses, with 66% of irrigation blocks in the region identified as grapevines. The second most prevalent horticultural land use was citrus orchards (18% of irrigation blocks). Figure 3 shows that the ET-NDVI relationships of irrigation blocks showed a strong compliance with the TICRF, and that there was a wide range of NDVI present across these horticulture blocks. Mean ET/ET_r (denoted by black lines in Figure 3) was 0.54 and 0.57 for grapevine and citrus land uses respectively. Mean ET/ET_r offset, $dKcb$, shown as green lines in Figure 3 are the most appropriate estimates of K_c for both grapevines and citrus. NDVI-dependent K_c estimates of $(K_{cb} + 0.18)$ and $(K_{cb} + 0.15)$ describe mean rates of water use for both grapevines and citrus respectively. There is minimal likelihood of water stress at these values of K_c . The lower 95% confidence limit (shown as a brown line in Figure 3) closely corresponds with the irrigation refill line. Similar to data presented for perennial pastures, ET and NDVI showed a strong linear relationship.

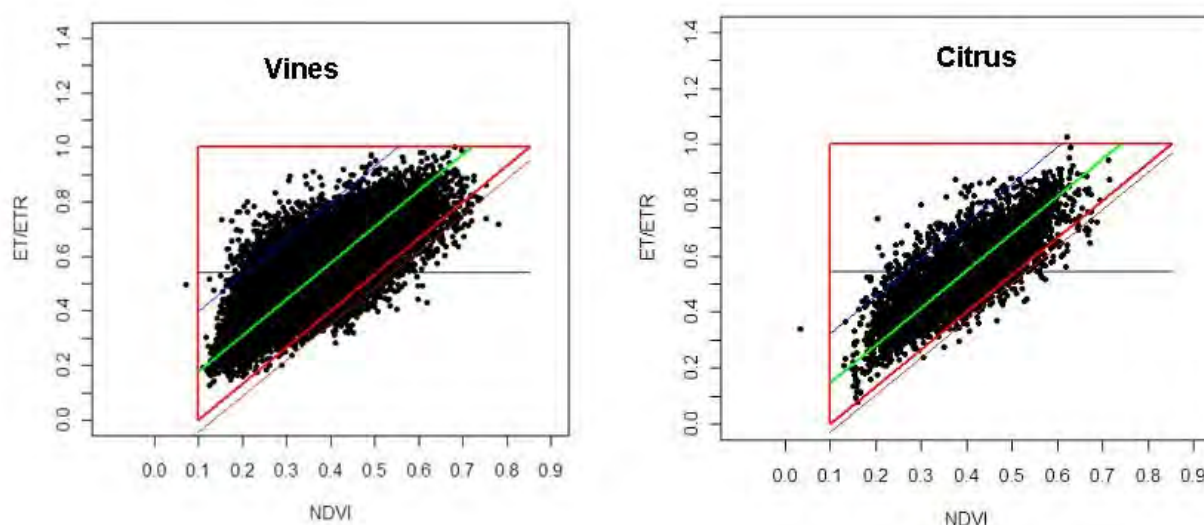


Figure 3. Scatterplots of NDVI and ET values for the two major horticultural enterprises located in the Sunraysia Irrigation Region, derived from Landsat 5 TM images acquired on 5/1/2009. The results are presented in terms of the TICRF, which is depicted by the red triangle on each graph. Green lines show mean ET displacement (dK_{cb}) from the irrigation refill line. Blue lines shown the upper 95% confidence interval on dK_{cb} . Brown lines show the lower 95% confidence interval of dK_{cb} . Horizontal black lines show mean ET/ET_r ratio (ET_rF)

Analysis of images acquired across multiple years provides additional information about the robust nature of this type of analysis. This was particularly important as during the 2008-2009 irrigation period south-east Australia was experiencing extended drought conditions, with an approximate average annual rainfall of 200mm and 230mm recorded at the Mildura weather station during 2008 and 2009 respectively (data sourced from www.weatherzone.com.au). However, during the 2009-2010 irrigation period, these drought conditions no longer prevailed and approximately 490mm of rainfall was recorded in 2010. Figure 4 shows ET-NDVI values for grapevines, and their relationship to the TICRF during December 2008, January 2009 and December 2010. Median NDVI values for 2010 were slightly higher than median values observed in the 2008-2009 irrigation period (0.5 and 0.4 respectively), suggesting a slightly larger crop, in terms of vegetation cover, was grown in December 2010. Figure 4 illustrates that while data collected in January 2009 was largely consistent with the NDVI-dependent ET limits described by the TICRF, the ET rates recorded during December 2008 and December 2010 frequently fell below the irrigation refill line depicted by the TICRF. This suggests that there was insufficient water to maintain maximum rates of ET in December 2008 and 2010.

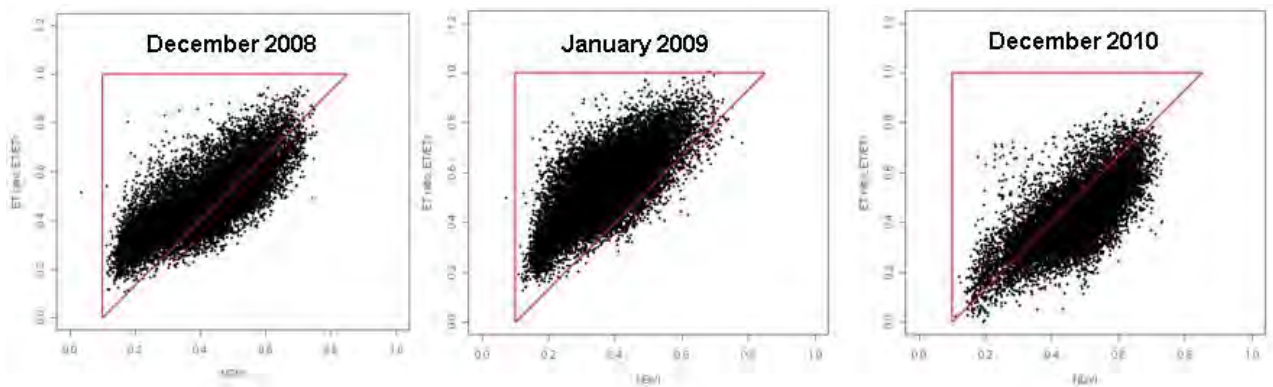


Figure 4. Scatterplots of NDVI and ET values for grapevine blocks located in the Sunraysia Irrigation Region, derived from Landsat 5 TM images acquired during December 2008, January 2009 and December 2010. The results are presented in terms of the TICRF, which is depicted by the red triangle on each graph

3.3 Barmah Forest – riparian vegetation

The potential application of ET-NDVI relationships to the evaluation of vegetation condition and water use in ecologically significant riparian vegetation was also assessed. Figure 5 shows the ET-NDVI relationship in the Barmah Forest, located along the Murray River.

Most observations were located within the bounds of the TICRF. Based on this data, most riparian observations can then be assumed to comply with the basic principles of water use and ET demonstrated by agricultural crops. This includes the expected increase in ET with increasing vegetation cover (NDVI), and the increased rates of ET associated with enhanced soil evaporation when the soil is wet. Large rates of ET were seen where low NDVI values were observed ($NDVI < 0.1$). These observations were consistent with ET rates expected from areas of wet, sparsely vegetated soil or areas of free surface water. These scenarios are to be expected in a riparian area such as the Barmah Forest, where inundation is a feature of these native vegetation systems. In some cases, ET exceeded the upper limit of the TICRF ($ET = ET_r$) for values of NDVI between 0.1 and 0.85. This is indicative of the presence of standing water on the soil surface located at the base of vegetation cover such as grassy woodland forests and wetland vegetation. Observations of low ET and low NDVI values occurred throughout areas of the Barmah Forest. These observations provide a quantitative measure of the extent of aridity and degradation of native vegetation.

The riparian data clearly show that ET-NDVI relationships provide a repeatable, affordable source of qualitative and quantitative information to describe status and activity of riparian vegetation. However, it should be noted that observations were sourced from a single image and that this approach has not been extensively tested over riparian vegetation systems to date. However, consistencies in these observations with the TICRF suggest that the DPI SM approach may provide an objective reporting mechanism and approximate estimates of the evaporative water requirement of riparian vegetation.

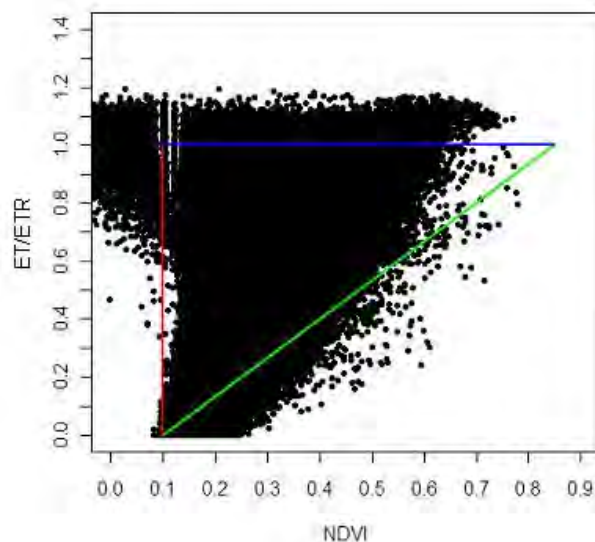


Figure 5. Scatterplot of NDVI and ET values for pixels located within the Barmah Forest, derived from Landsat 5 TM images acquired on 4/11/2008. The results are presented in terms of the TICRF, which highlights the lower limit of vegetation cover (red line), the maximum ET rate (blue line) and the irrigation refill line (green line)

4 CONCLUSION

Satellite remote sensing of ET has been demonstrated to be an innovative and affordable means of acquiring comprehensive water use data, specific to enterprise type and region. ET-NDVI relationships were used in this project to establish a consistent, objective biophysical context for results. ET-NDVI relationships can describe maximum rates of crop ET, and their dependence on NDVI. Previous studies have shown that crop water use is dependent on an increase in crop leaf area (measured as NDVI), in addition to influencing factors such as rainfall and irrigation (Tasumi et al., 2005). The use of satellite remote sensing in this way extends the application of established scientific principles of vegetation water use to a wide range of enterprises, and also extends these principles from farm to regional and national scales using a consistent, comparable method.

Traditionally, estimates of crop evapotranspiration are based on generic crop coefficients and weather data. The findings that linear relationships existed between ET and NDVI, and that NDVI varied greatly between fields implies that knowledge of crop type and weather alone is insufficient for effective irrigation water management. This project has shown that vegetation cover (NDVI) provides a more explicit guide to calculating crop water requirement than knowledge of the crop type alone.

This project has established methods and approaches that allow for improved on-farm water management in Australia, and improved objective water planning at farm to catchment scales. The methods presented are affordable, comprehensive and repeatable in comparison with the traditional approaches. Results from this project have improved understanding of the relative roles of crop type, vegetation cover, regional weather and irrigator behaviour in water management.

ACKNOWLEDGEMENTS

This work was supported by the Department of Primary Industries, the Department of Sustainability and Environment, the National Water Commission and the CRC Irrigation Futures. Cress Savage provided a substantial amount of IDL code to process the imagery used in this work.

REFERENCES

- Abuzar, M., McAllister, A., Whitfield, D., Morse-McNabb, E. & Savige, C. (2008) Remote Sensing Tools and Approaches to Integrated Irrigation Water Management at Farm and Regional Scales. *14th Australasian Remote Sensing and Photogrammetry Conference*. Darwin, Australia.
- Allen, R., Tasumi, M. & Trezza, R. (2005) METRIC - Mapping Evapotranspiration at High Resolution. Applications Manual for Landsat Satellite Imagery. Version 2.0. University of Idaho.
- Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. (1998) Crop evapotranspiration - Guidelines for computing crop water requirements. Rome, FAO (Food and Agricultural Organization of the United Nations).
- Allen, R. G., Tasumi, M. & Trezza, R. (2007) Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC) - Model. *Journal of Irrigation and Drainage Engineering*, vol. 133, pp. 380-394.
- Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A. & Holtslag, A. A. M. (1998) A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *Journal of Hydrology*, vol. 212-213, pp. 198-212.
- Bastiaanssen, W. G. M., Noordman, E. J. M., Pelgrum, H., Davids, G., Thoreson, B. P. & Allen, R. G. (2005) SEBAL Model with Remotely Sensed Data to Improve Water-Resources Management under Actual Field Conditions. *Journal of Irrigation and Drainage Engineering*, vol. 131, pp. 85-93.
- Chong, J. & Ladson, A. R. (2003) Analysis and Management of Unseasonal Flooding in the Barmah-Millewa Forest, Australia. *River Research and Applications*, vol. 19, pp. 161-180.
- Jensen, J. R. (2000) *Remote Sensing of the Environment: An Earth Resource Perspective*, Upper Saddle River, New Jersey, USA, Prentice Hall.
- Lawrence, L. & Colloff, M. (2008) Rescuing the Barmah Forest. *ECOS*, vol. 141, pp. 22-24.
- Liang, S. (2004) *Quantitative Remote Sensing of Land Surfaces*, New Jersey, USA, John Wiley & Sons, Inc.
- Tasumi, M., Allen, R. G., Trezza, R. & Wright, J. L. (2005) Satellite-Based Energy Balance to Assess Within-Population Variance of Crop Coefficient Curves. *Journal of Irrigation and Drainage Engineering*, vol. 131, pp. 94-109.
- Teixeira, A. H. d. C., Bastiaanssen, W. G. M., Ahmad, M. D. & Bos, M. G. (2009) Reviewing SEBAL input parameters for assessing evapotranspiration and water productivity for the Low-Middle Sao Francisco River basin, Brazil. Part A: Calibration and validation. *Agricultural and Forest Meteorology*, vol. 149, pp. 462-476.
- Whitfield, D., McAllister, A., Abuzar, M., Sheffield, K., O'Connell, M. & McClymont, L. (2010a) Applications of DPI SEBAL-METRIC to horticultural crops in the Riverland region of South Australia, and riparian vegetation in the Murray River corridor in Victoria. Taturam Victoria, Department of Primary Industries.
- Whitfield, D., O'Connell, M., McAllister, A., McClymont, L., Abuzar, M. & Sheffield, K. (2010b) SEBAL-METRIC Estimates of Crop Water Requirement in Horticultural Crops Grown in SE Australia. *28th International Horticultural Congress*. Lisbon, Portugal.

Disaster Change Detection Using Airborne LiDAR

John Trinder

School of Surveying and Spatial Information Systems, The University of New South Wales
UNSW SYDNEY NSW 2052, Australia
j.trinder@unsw.edu.au

Mahmoud Salah

Dept. of Surveying, Faculty of Engineering Shoubra, Benha University
108 Shoubra Street, Cairo, Egypt
engmod2000@yahoo.com

ABSTRACT

Potential applications of airborne LiDAR for disaster monitoring include flood prediction and assessment, monitoring of the growth of volcanoes and assistance in the prediction of eruption, assessment of crustal elevation changes due to earthquakes, and monitoring of structural damage after earthquakes. Change detection in buildings is an important task in the context of disaster monitoring, especially after earthquakes. The paper will describe the capability of airborne LiDAR for rapid change detection in elevations, and methods of assessment of damage in made-made structures. The approach is to combine change detection techniques such as image differencing, principal components analysis (PCA), minimum noise fraction (MNF) and post-classification comparison based on support vector machines (SVM), each of which will perform differently, based on simple majority vote. In order to detect and evaluate changes in buildings, LiDAR-derived DEMs from two epochs were used, showing changes in urban buildings due to construction and demolition. To meet the objectives, the detected changes were compared against reference data that was generated manually.

The comparison is based on three criteria: overall accuracy; commission and omission errors; and completeness and correctness. The results showed that the average detection accuracies were: 84.7%, 88.3%, 90.2% and 91.6% for post-classification, image differencing, PCA and MNF respectively. On the other hand, the commission and omission errors, and completeness and the correctness of the results improved when the techniques were combined, compared to the best single change detection method. The proposed combination of techniques gives a high accuracy of 97.2% for detection of changes in buildings, which demonstrates the capabilities of LiDAR data to detect changes, thus providing a valuable tool for efficient disaster monitoring and effective management and conservation.

KEYWORDS: LiDAR, Change detection, Building extraction, Feature extraction.

1 INTRODUCTION AND RELATED WORKS

An up-to-date building database is a crucial requirement for reliable disaster damage assessment. Change detection employing LiDAR (Light Detection and Ranging) data is a useful tool for damage detection, particularly for collapsed multi-floored buildings (Tuong et al., 2004). LiDARs are active acquisition systems equipped with a laser scanner, a Global Positioning System (GPS) receiver and an Inertial Navigation System (INS). They emit infrared laser pulses at high frequency and record the time of flight of the return pulses. By combining the LiDAR distance with GPS and INS data the X, Y and Z coordinates of ground points can be determined. The intensity of

the returns can also be recorded. There are basically two types of LiDAR systems: discrete and waveform. A short pulse (~ns) is emitted from the laser and in discrete systems one or more discrete distances and intensities are recorded. Waveform systems record the full waveform of the return signal.

Methods of change detection can mainly be divided into two categories:

- The determination of the difference of classifications of a surface obtained at two periods;
- The direct determination of change between two data sets.

Detecting changes by supervised classification is unreliable when the appearances of non-buildings and buildings are similar. Furthermore, using spectral information to detect change does not consider the situation when the differences occur in shape instead of colour (Huang and Chen, 2007). A number of research results, such as Knudsen and Olsen (2003), Matikainen et al. (2004), Walter (2004a, b) and Nielsen and Canty (2011) belong to the first category above. The second category [Murakami et al., 1999; Jung, 2004] is unable to determine the land category because no classification is used. It is also observed that trees often cause mistakes in the output of research.

Even though aerial photography has been conventionally employed for change detection [Niederöst, 2001; Knudsen and Olsen, 2003; Walter, 2004a; Walter, 2004b], aerial photography is subject to several unavoidable problems such as: shadows in the scenes acquired over dense urban areas with many skyscrapers; the spectral information of certain features in aerial photography is diverse and ill-defined (Knudsen and Olsen, 2003); perspective projection causes relief displacement of buildings, which requires height information to correct. Therefore, the employment of LiDAR data rather than spectral information derived from aerial photographs offers important advantages (Tuong, et al., 2004). It allows obtaining 3D point clouds of the surface with high density as well as high accuracy. Moreover, the method is capable of collecting data over large areas in a short time (Baltsavias, 1999).

Instead of the multi-spectral imagery that was often used in the past, many change detection methods using LiDAR data have been proposed. Murakami et al., (1999) carried out change detection of buildings using LiDAR data in Japan. That study was a simple comparison between two datasets. Tuong et al., (2004) presented an automatic method for LiDAR-based change detection of buildings in dense urban areas. Walter (2004b) used LiDAR data in an object-based classification to determine the land-use category after the observation of land phenomena. Matikainen et al. (2004) divided a LiDAR point cloud into homogeneous areas, and then extracted information to discover the building areas for change detection. Girardeau-Montau et al. (2005) directly used point-to-point position relations for change detection. Huang and Chen (2007) included LiDAR data and aerial images to detect the changes of building models. Brzank et al., (2009) presented a new method to detect and evaluate morphologic changes of the Wadden Sea based on the extraction of structure lines of tidal channels from LiDAR data. Chien and Lin (2010) developed a new method to find changes within 3D building models in the region of interest with the aid of LiDAR data. Their modelling scheme comprises three steps, namely, data pre-processing, change detection in building areas, and validation. Research findings clearly indicate that the double-threshold strategy improves the overall accuracy from 93.1% to 95.9%.

It is worth mentioning that, as change detection is an important step in data updating, some methods used spectral-based methods such as the iterative principal components analysis (IPCA) to determine temporal distance in feature space and combine it with a Bayesian decision rule to determine the presence of change (Spitzer *et al.*, 2001). Clifton (2003) describes training neural networks to learn expected changes between images and to then identify pixel changes which do not match what is “expected”. Hashimoto et al. (2011) proposed a knowledge-based change detection approach, which can obtain change information that includes not only land cover changes, but also contextual changes, such as types of damage caused by natural hazards. This approach mainly consists of two processes: information extraction and change inference using a Bayesian network. Information extraction employs object-based image analysis for extracting spatial information. Change inference uses extracted information and the Bayesian network constructed from knowledge of the change detection process. To demonstrate this approach, change detection of mudslide damage caused by heavy rain in Yamaguchi Pref., Japan was conducted. Some other methods used multi-temporal high-resolution imagery to detect changes in spectral difference or used supervised classification to determine building positions for comparisons of two epochs for change detection (Knudsen and Olsen, 2003; Kumar, 2011).

This paper describes a proposed workflow for LiDAR-based change detection. The paper is organised as follows. Section 2 describes the study areas and data sources. Section 3 describes the experiments while Section 4 presents and evaluates the results. We summarise our results in Section 5.

2 STUDY AREAS AND DATA SOURCES

2.1 LiDAR data

No data was available before and after an earthquake, but in order to demonstrate the capability of the proposed change detection method, two LiDAR data sets were available, acquired on different occasions over Tokyo, which is an earthquake-prone area in Japan. Survey flights were carried out by Asia Air Survey Co. Ltd in June 1999 and February 2004. The sensor and scene characteristics of the used data are summarized in Table 1. Data was provided in 1m² grid formats over a dense urban area which includes residential buildings, large buildings, a network of main and local roads, open and green areas as well as trees. The pulse intensity is recorded in more recent LiDAR systems but it has not been a concern in this study. Figure 1 shows the acquired LiDAR data in grid format. It is worth mentioning that the test areas have been used in previous studies (Tuong et al., 2004) for change detection, which used a simple method to form the difference image based on a histogram thresholding and different reference data. The results showed 55% error in detecting new demolitions.

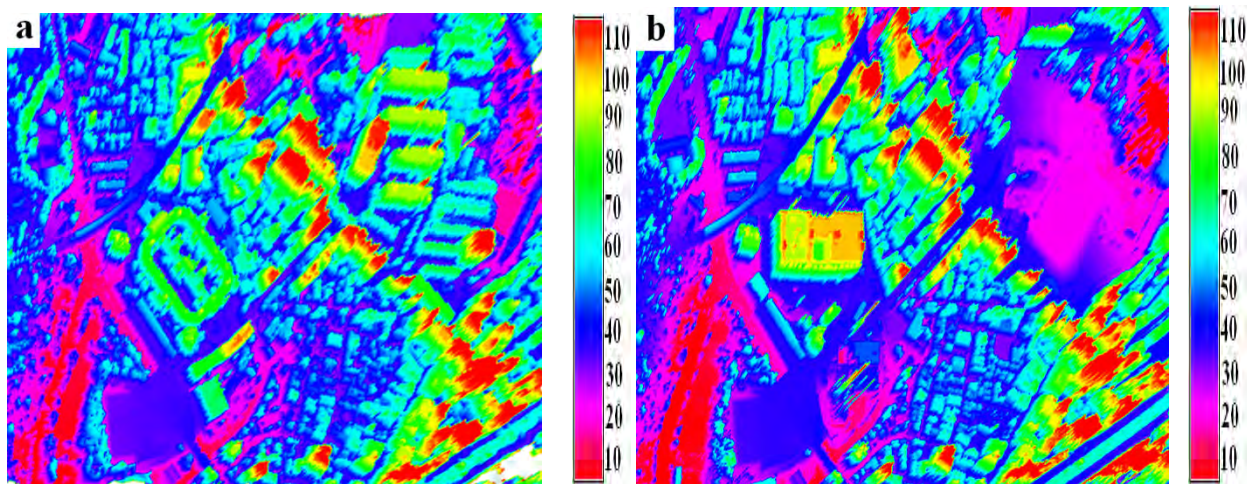


Figure 1: Acquired LiDAR data in grid format a) June 1999, and b) February 2004.

Table 1: Characteristics of LiDAR datasets.

	First Dataset	Second Dataset
Spacing across track (m)	0.15	0.85
Spacing along track (m)	1.5	1.48
Vertical accuracy (m)	0.15	0.10
Horizontal accuracy (m)	0.5	0.5
Density (Points/m ²)	4	2.5
Sampling intensity (mHz)	125	150
Wavelength (μm)	1.56	1.064
Average altitude (m)	800	1450
Laser swath width (m)	750	777.5
Acquisition date	June 1999	February 2004

2.2 Reference data

In order to accurately evaluate the performance of the proposed change detection method, changes were visually interpreted and digitized independently of their size. The reference data is a four-class thematic image, typically divided into the four categories of: '*background*'; '*decreased*', '*increased*'; and '*unchanged*' as shown in figure 2.

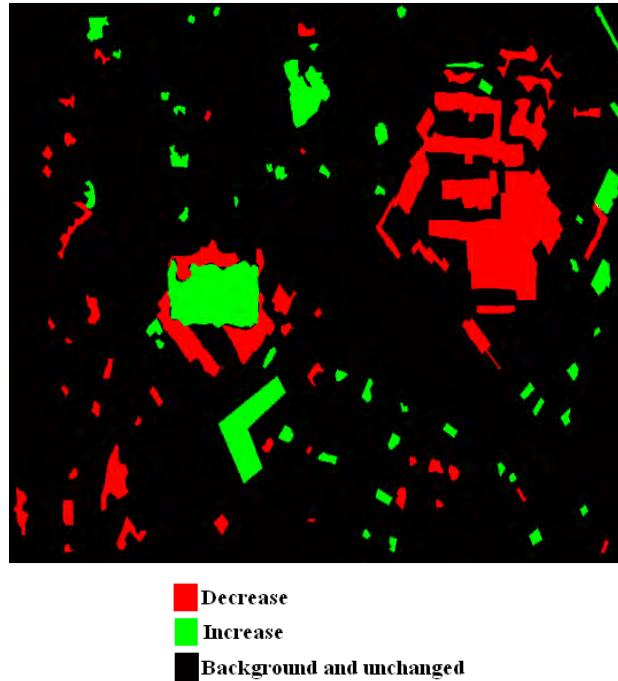


Figure 2: Manually digitized reference changes.

3 METHODOLOGY

3.1 Pre-processing

First, both DEMs were registered to each other based on a projective transformation. The registration process resulted in small Root Mean Square Errors (RMSE) that did not exceed 0.15m in both X and Y directions. Following the transformation, the images were resampled to 1m pixel size. A grid format is preferred to the raw point cloud format to speed up the processing, particularly when there is a direct comparison of the two datasets. In order to obtain a high image quality and to reduce the processing time, a bilinear interpolation was applied for the resampling process. The bilinear interpolation can result in a better quality image than nearest neighbourhood resampling and requires less processing than cubic convolution.

3.2 Main-processing

Four different change analyses were performed to evaluate the efficacy of LiDAR data for detecting changes occurring at two different temporal scales. The four methods include: image differencing; principal components analysis (PCA); minimum noise fraction (MNF); and Post-Classification based on support vector machine (SVM). After these steps, a simple majority vote has been applied to generate the change detection image. All the methods proposed in this research were implemented through programs generated by the authors in a Matlab environment. An interface was developed to enable the user to: detect changes through the aforementioned four methods; combine votes derived from all methods; generate a change detection image; and evaluate the change detection results. The workflow for this investigation is shown in figure 3.

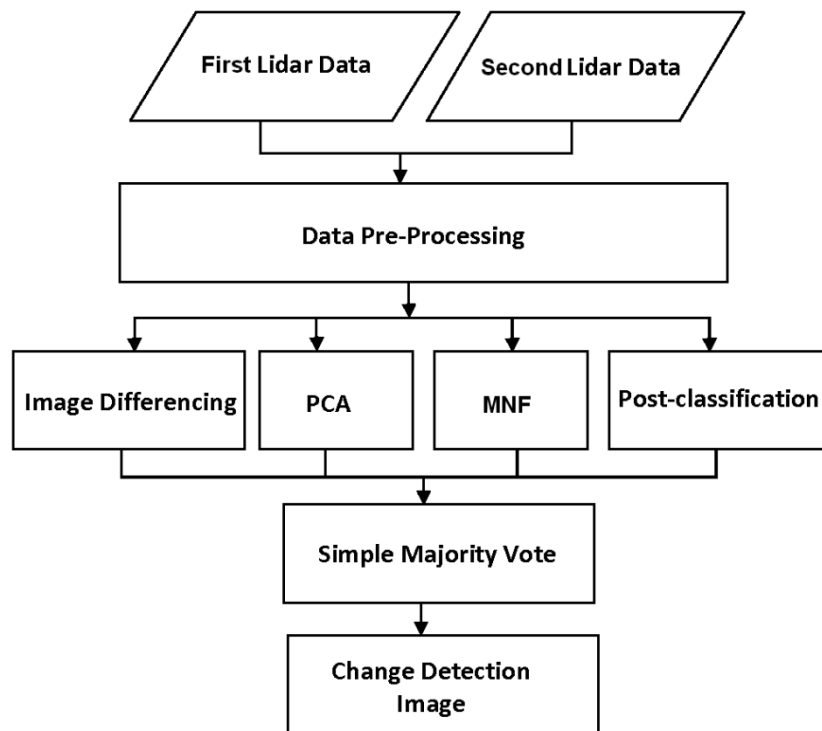


Figure 3: The workflow for the proposed change detection method.

In the first change detection technique, differences between the two DEM images that exceed a user-specified threshold of 10 pixels in area and 0.30m in height, double of the LiDAR system accuracy, were computed and highlighted. In the image differencing method, the second image is subtracted from the first image to provide the difference and highlight changes. The second image is more recent and the differences reflect changes over time. After application of image differencing, increases in height values that are more than the predefined thresholds, are highlighted as *increases*, while decreases in height values that are more than the predefined thresholds, are highlighted as *decreases*. The result is a grey scale image composed of a single band of continuous data that reflects the changes. The change image is a four-class thematic image, typically divided into the four categories of: *background*, *decreased*, *increased*, and *unchanged*. Although the calculation is simple, the interpretation requires knowledge about the area, because every difference relates to a certain location but not necessarily to the same object.

In the second change detection technique, Principal Components Analysis (PCA) has been applied to detect changes. PCA is commonly applied for orthogonal data transformations. PCA maximizes the spectral variability detected by decreasing the redundancy of information contained in multiple spectral bands (Armenakis et al., 2003). PCA components are based on statistical relationships that are difficult to interpret, and are variable between different landscapes and different dates for a single landscape (Collins and Woodcock 1994). PCA is a linear transformation of the data along perpendicular axes of maximum variance between data sets (Legendre and Legendre 1998). The first eigenvector sorts pixels along an axis of highest correlation between data sets. Pixels on this axis have not significantly changed between the two images. The second eigenvector is perpendicular to the first, and therefore sorts pixels that represent differences between data sets.

In the third change detection technique, the minimum noise fraction Transform (MNF) as modified from Green et al. (1988) has been performed to detect changes. MNF is a linear transformation that consists of the following separate principal components analysis rotations: (i) The first rotation uses the principal components of the noise covariance matrix to decorrelate and rescale the noise in the data (a process known as noise whitening), resulting in transformed data in which the noise has unit variance and no band-to-band correlations; (ii) The second rotation uses the principal components derived from the original image data after they have been noise-whitened by the first rotation and rescaled by the noise standard deviation. The inherent dimensionality of the data is determined by examining the final eigenvalues and the associated images. For the best

results, and to save disk space, only those bands with high eigenvalues have been output. Images with eigenvalues close to 1 are mostly noise.

In the fourth change detection technique, post-classification comparison was performed in order to detect changes. The two DEM images were classified using a Support Vector Machine classifier (SVM), then the classification results were compared and the differences were extracted. The objective is to classify the input data into four primary classes of interest, namely buildings, trees, roads, and ground. SVMs are based on the principles of statistical learning theory (Vapnik, 1979) and delineate two classes by fitting an optimal separating hyperplane (OSH) to those training samples that describe the edges of the class distribution. As a consequence they generalize well and often outperform other algorithms in terms of classification accuracies. Furthermore, the misclassification errors are minimized by maximizing the margin between the data points and the decision boundary. Since the One-Against-One (1A1) technique usually results in a larger number of binary SVMs and then in subsequently intensive computations, the One-Against-All (1AA) technique was used to solve for the binary classification problem that exists with the SVMs and to handle the multi-class problems. The Gaussian radial basis function (RBF) kernel has been used, since it has proved to be effective with reasonable processing times in remote sensing applications.

Then a simple majority vote, which can be more effective than more complex voting strategies (Waske, 2007), is used to generate the final result. If change detection algorithm c_i assigns a given pixel to class label ω_j , then we say that a vote is given to ω_j . After counting the votes given to each class label by all detection algorithms, the class label that receives the highest number of votes is taken as the final output. It is worth mentioning that all votes are of equal weight and independent of height differences. When the four detection methods give completely different decision for a given pixel, which does not convey any information, the decision from the method with highest overall detection accuracy is considered.

As a last step, the smaller detected regions were merged into larger neighbouring homogeneous ones or deleted according to an arbitrary 1 m distance and 30 m² area thresholds respectively. The area threshold represents the expected minimum change area, while the distance threshold was set to 1 m to fill in any gaps within the detected region. Regions were retained if they were larger than the given area threshold and/or were adjacent to a larger homogeneous region by a distance less than 1 m. Finally, region borders were cleaned by removing structures that were smaller than 5 pixels and that were connected to the region border. There was a compromise between cleaning thresholds less than 5 pixels, which may leave the original buildings uncleaned, and thresholds greater than 5 pixels which may remove parts of the detected region. The result was an image that represents the detected changes without noisy features and also without holes.

3.3 Evaluation of the change detection results

In order to evaluate the performance of the adopted method for change detection from LiDAR data, the results have been checked based on three different methods which include: (i) The overall detection accuracy; (ii) The produced omission and commission errors; and (iii) The completeness and correctness of the results.

The overall accuracy for the detection process was assessed using the reference data and based on equation 1:

$$ODA = \frac{NCP}{NRP} \quad (1)$$

Where ODA is the overall detection accuracy; NCP is the total number of correctly detected pixels and NRP is the total number of reference pixels.

Since the overall detection accuracy is a global measure for the performance of the combination process, two additional measures were used to evaluate the performance of the proposed combination method, namely: commission and omission errors (Congalton, 1991). Unlike overall detection accuracy, commission and omission errors clearly show how the performance of

the proposed methods improves the results or cause a deterioration of results for each individual class compared to the reference data.

$$CE_I = \frac{A_1 + A_2 + A_3}{R_1} \quad (2)$$

$$OE_I = \frac{B_1 + B_2 + B_3}{R_1} \quad (3)$$

CE_I and OE_I are commission and omission errors of class *increased*; A_1 , A_2 and A_3 are the numbers of incorrectly identified pixels of class *increased* associated with classes *decreased*, *background* and *unchanged*; R_1 is the total number of pixels of the class *increased* as observed in the reference data; B_1 , B_2 and B_3 are the numbers of unrecognized pixels that should have identified as belonging to the class *increased*. The same is applicable for classes *decreased*.

On the other hand, in order to evaluate the performance of the change detection process, the *completeness* and the *correctness* (Heipke et al., 1997) of the detected changes were investigated based on a per-pixel as follow:

$$Completeness = \frac{TP}{TP + FN} \quad (4)$$

$$Correctness = \frac{TP}{TP + FP} \quad (5)$$

TP denotes to the number of true positives which is the number of entities that were automatically detected and were available in the reference data. FN relates to the number of false negatives which is the number of entities that were available in the reference data but not automatically detected. FP stands for the number of false positives which is the number of entities that were automatically detected but do not correspond to any entities in the reference data (Rottensteiner et al., 2007).

4 RESULTS AND ANALYSIS

Figure 4 is a typical example showing the results in a sub-area of the whole test area. For the detected changes, the green colour indicates an *increase*, the red colour labels a *decrease* while black colour refers to both *background* and *unchanged*. It can clearly be seen that the important changes occur in buildings. In the middle of the area the large building has been replaced with a new one with a different shape. At the lower right part of the area, some new buildings have been constructed.

Another aspect of interest is where the misclassified pixels were recovered by the combination process. Most corrections occur at the edges of buildings and trees, which demonstrates the effect of between-class variance on the edge pixels which caused many of these pixels to be placed in an incorrect category. It can clearly be seen that the detected changes for PCA are eroded as compared to the reference data. This trend can also be observed for post-classification results. On the other hand, the detected changes for image differencing are larger. However, the erosion effect has been reduced after applying the MNF and the simple majority vote combination.

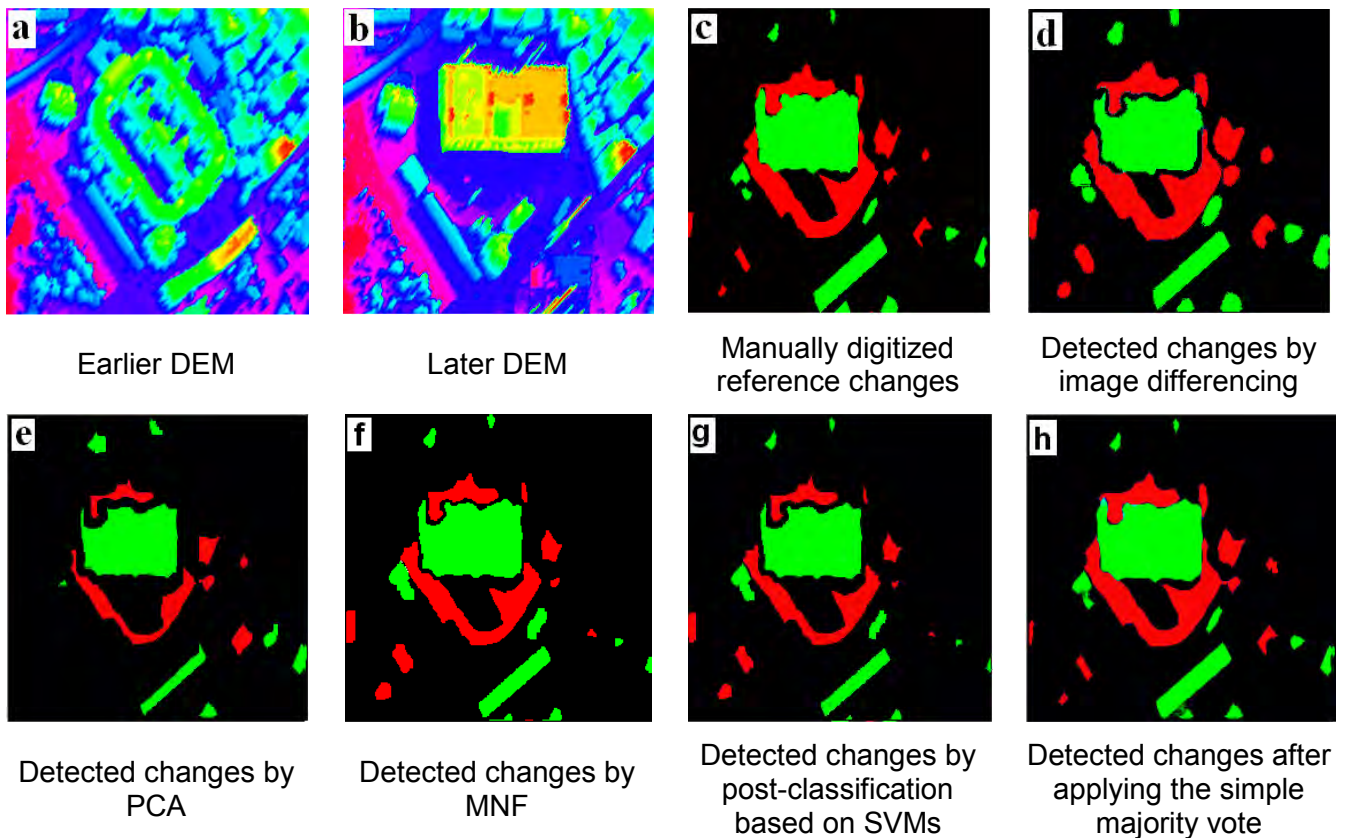


Figure 4: Typical example of the change detection results. Areas in red are classified as *decreased*; areas in green are classified as *increased*; while areas in black refer to both *background* and *unchanged*.

The overall detection accuracies of individual techniques, based on the reference data, are given in Table 2. MNF performed the best with 91.6% detection accuracy, followed by PCA, image differencing and post-classification with detection accuracies of 90.2%, 88.3% and 84.7% respectively.

The improvement in detection accuracies achieved by the combination method compared with the best individual detection technique, MNF, was determined as shown in Table 2. It is clear that the performances of simple majority vote are better than those of single detection methods. The improvement in detection accuracy of 5.6% is obtained from simple majority vote algorithm.

Table 2: Performance evaluation of single detection techniques.

Method	Detection accuracy (%)
image differencing	88.3
PCA	90.2
MNF	91.6
post-classification	84.7
simple majority vote	97.2

Table 3 shows the commission and omission errors based on the proposed method, compared to the commission and omission errors of the best individual methods. It can be seen that a considerable amount of the misclassified pixels have been recovered by the combination process. For the best change detection method, MNF, the omission errors vary from 8.08% to 10.65%, while the commission errors range from 6.17% to 9.28% respectively. For the combined change detection method based on the simple majority vote algorithm, the omission errors range

from 2.87% to 7.03%, while the commission errors range from 3.32% to 6.11% respectively. The most important point to note is that both types of errors using simple majority vote are comparable. This indicates the capabilities of the proposed method to detect changes from LiDAR data. Figure 5, which is a typical example of the error distribution map, showing that omission and commission errors mostly occurred at the building outlines.

It is worth mentioning that some differences along the edges of buildings are not real changes. One possible reason for these differences is the random reflectance of laser pulses. Another possible reason could be the differences between parameters of the two surveying flights. The flight in 1999 was collected in summer while the flight in 2004 was collected in winter; many of detected ‘demolitions’ were trees. It can be concluded that trees are the main cause of commission errors. There seems to be no mechanism to check the omission errors, which supports previously reported findings (Tuong et al., 2004).

Table 3: Change detection errors of the proposed method compared with those of the individual methods. Com. and Om. Stand for commission and omission errors respectively.

	post-classification		image differencing		PCA		MNF		simple majority vote	
	Om. (%)	Com. (%)	Om. (%)	Com. (%)	Om. (%)	Com. (%)	Om. (%)	Com. (%)	Om. (%)	Com. (%)
<i>Increased</i>	19.02	11.93	11.68	9.38	10.30	09.31	8.08	09.28	02.87	03.32
<i>Decreased</i>	21.50	11.98	16.89	11.46	13.09	07.32	10.65	06.17	07.03	06.11

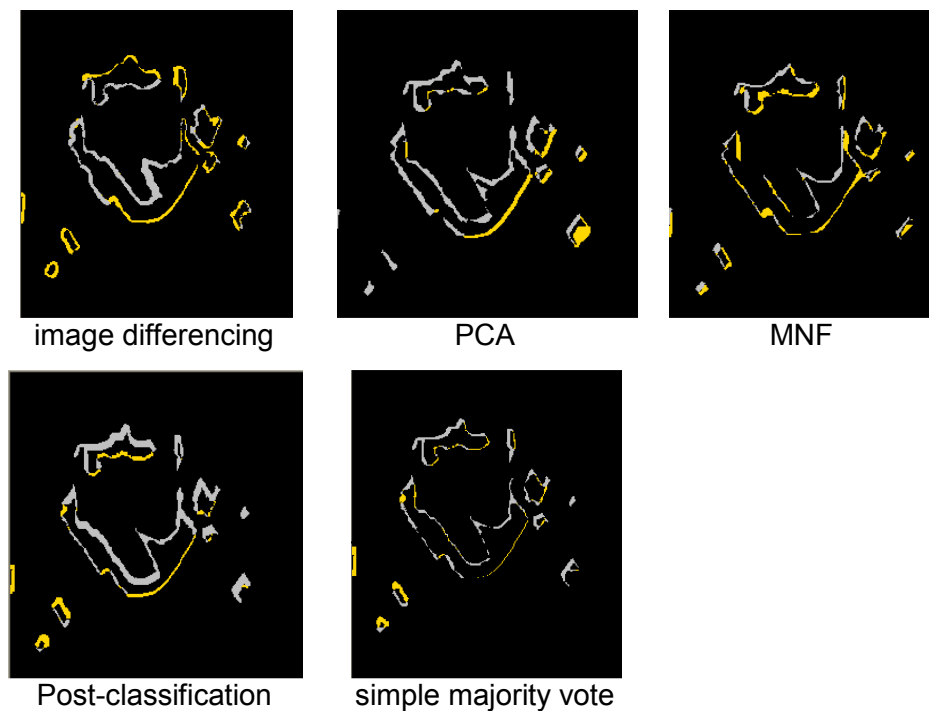


Figure 5: Evaluation of the results of change detection. Black: correct changes in pixels; grey: omission errors; yellow: commission errors.

For the per-region based evaluation, a region was counted as a true positive if at least 90% of its area from the automatically detected results was overlapped by the corresponding area in the reference data. Figure 5 shows the completeness and correctness against the region size. The completeness and correctness of regions around 30m² were around 79% and 75% respectively and these statistics improve as the region size increased. Completeness and correctness were over 91% for all regions larger than 50m². The difference between completeness and correctness is a matter of 0.1–1% except for regions smaller than 50m² where the difference is up to 4%. This further confirms the lower reliability of detecting regions smaller than 50m². It can therefore be

concluded that these tests strongly represent achievable accuracies for detection of changes by the proposed method using LiDAR data.

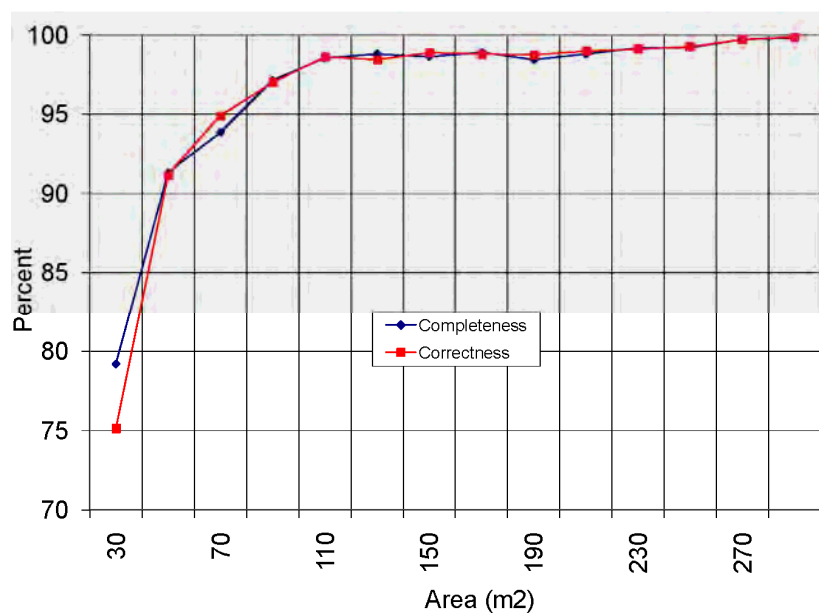


Figure 5: Completeness and correctness derived for the detection process plotted against size of detected areas.

5 CONCLUSION

In this paper, we have applied a powerful method to combine change detection techniques with different performance based on the simple majority vote. To test the algorithm, four change detection methods were based on LiDAR data of different two epochs. The results showed an improvement in terms of detection accuracy as well as omission and commission errors. Detection accuracies of individual algorithms were 84.7%, 88.3%, 90.2% and 91.6% for post-classification, image differencing, PCA and MNF respectively, whereas the proposed fusion algorithm gave an accuracy of 97.2% which is an improvement of around 5.6%. On the other hand, the proposed method showed a high level of automation in change detection process. These results demonstrate the overall advantages of the proposed algorithm for change detection that could be applicable for detecting changes in buildings damaged in a disaster such as an earthquake. If two LiDAR flights could be carried out before and after an earthquake, the change detection results can reveal the collapsed buildings. Although this paper used only heights of buildings for change detection, it is well-prepared and opened to integrate elevation and intensity distribution in future studies. Spectral information from aerial imagery can also be applied along with LiDAR data in order to improve the performance of the proposed method, and to refine the results.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Asia Air Survey Co. Ltd for the provision of the LiDAR datasets.

REFERENCES

Armenakis C., F. Leduc, I. Cyr, F. Savopol, F. Cavayas (2003). A comparative analysis of scanned maps and imagery for mapping applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57:304–14.

- Baltsavias E. (1999). Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3), pp. 199-214.
- Brzank A., Heipke C., Goepfert J. (2009). Morphologic Change Detection in the Wadden Sea from LiDAR Data. *ISPRS Congress*, Volume XXXVII, Part B8, Commission VIII, ISSN 1682-1750, 3 to 11 July 2008, Beijing, China, p.647 ff.
- Chen L. C. (2010). Detection of building changes from aerial images and light detection and ranging (LIDAR) data. *Journal of Applied Remote Sensing*, Vol. 4, 041870 (19 November 2010).
- Clifton C. (2003). Change Detection in Overhead Imagery using Neural Networks. *International Journal of Applied Intelligence*, 18(2):215-234.
- Collins J. B., C. E. Woodcock (1994). Change detection using the Gramm-Schmidt transformation applied to mapping forest mortality. *Remote Sensing of Environment*, 50:267–79.
- Congalton R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1):35–46.
- Girardeau-Montaut D., Roux M., Marc R., Thibault G. (2005). Change Detection on Points Cloud Data Acquired with a Ground Laser Scanner, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI (Pt. 3/W19):30-35.
- Green A. A., Berman M., Switzer P., Craig M. D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal: *IEEE Transactions on Geoscience and Remote Sensing*, v. 26, no. 1, p. 65-74.
- Hashimoto S., Tadono T., Onosatoa M., Horib M. and Moriyamab T. (2011). An Approach for Automatic Change Inference in High Resolution Satellite Images. *34th International Symposium on Remote Sensing of Environment*, The GEOSS Era: Towards Operational Environmental Monitoring, Sydney, Australia, 10-15 April 2011.
- Heipke C., Mayer H., Wiedemann C., Jamet O. (1997). Evaluation of automatic road extraction. *International Archives of Photogrammetry and Remote Sensing*, XXXII-2-3W3, Haifa, Israel, pp. 47–56.
- Huang C., Chen L. (2007). Detection of Building Changes from LiDAR Data and Aerial Imagery. In: *ACRS 2007: proceedings of the 28th Asian conference on remote sensing*, 12-16 November 2007, Kuala Lumpur, Malaysia.
- Jung F. (2004). Detecting Building Changes from Multitemporal Aerial Stereopairs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58:187-201.
- Kumar A. (2011). Temporal Changes in Mangrove Cover Between 1972 and 2001 Along the South Coast of the Arabian Gulf. *34th International Symposium on Remote Sensing of Environment*, The GEOSS Era: Towards Operational Environmental Monitoring, Sydney, Australia, 10-15 April 2011.
- Knudsen T., Olsen B. P. (2003). Automated Change Detection for Updates of Digital Map Databases, *Photogrammetric Engineering and Remote Sensing*, vol. 69 (11), pp. 1289-1297, 2003.
- Legendre P., L. Legendre (1998). *Numerical Ecology*. Elsevier Science, Amsterdam. 853 p.
- Matikainen L., Hyyppä J., Hyyppä H. (2004). Automatic Detection of Changes from Laser Scanner and Aerial Image Data for Updating Building Maps. *International Archives of Photogrammetry*,

Murakami H., Nakagawa K., Hasegawa H. Shibata T. (1999). Change Detection of Buildings Using an Airborne Laser Scanner, *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:148-152.

Niederöst M. (2001). Automated update of building information in maps using medium-scale imagery (1:15,000), Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), Baltsavias, E., Gruen, A., Van Gool, L. (Eds.), Balkema, Lisse, pp. 161-170.

Nielsen A. A., Canty M. J. (2011). A Method for Unsupervised Change Detection and Automatic Radiometric Normalization in Multispectral Data. *34th International Symposium on Remote Sensing of Environment*, The GEOSS Era: Towards Operational Environmental Monitoring, Sydney, Australia, 10-15 April 2011.

Rottensteiner F., Trinder J., Clode S. and Kubik K. (2007). Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance Evaluation and Sensitivity Analysis, *ISPRS Journal of Photogrammetry & Remote Sensing*, 62, pp. 135–149.

Spitzer H., R. Franck, M. Kollwe, N. Rega, A. Rothkirch, R. Wiemker (2001). Change Detection with 1 m Resolution Satellite and Aerial Images. In: *Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium*. Vol. 5, pp. 2256-2258.

Clifton C. (2003). Change Detection in Overhead Imagery using Neural Networks. *International Journal of Applied Intelligence*, 18(2):215-234.

Tuong T.V., Matsuoka M., Yamazaki F. (2004). LIDAR-based Change Detection of Buildings in Dense Urban Areas, *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, Vol. 5, pp.3413-3416.

Vapnik V. (1979). *Estimation of Dependences Based on Empirical Data [in Russian]*. Nauka, Moscow, 1979. (English translation: Springer Verlag, New York, 1982).

Walter V. (2004a). Object-based classification of remote sensing data for change detection, *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 58, pp. 225– 238, 2004.

Walter V. (2004b). Object-based Evaluation of LiDAR and Multispectral Data for Automatic Change Detection in GIS Databases, *IAPRS*, 35(B2):723-728.

Waske B. (2007). Classifying multisensor remote sensing data: concepts, algorithms and applications. *PhD thesis, Bonn University, Germany*.

Evaluation of Photo Imaging Methods for Vegetation Condition Assessment

Armando Apan, Govinda Baral, Ernest Dunwoody

Faculty of Engineering and Surveying & Australian Centre for Sustainable Catchments
University of Southern Queensland, Toowoomba 4350 QLD, Australia
apana@usq.edu.au, GovindaPrasal.Baral@usq.edu.au, Ernest.Dunwoody@usq.edu.au

Lucy Richardson

Condamine Alliance, 310 Anzac Avenue, Toowoomba 4350 QLD, Australia
lucy.richardson@condaminealliance.com.au

Kevin McDougall

¹Faculty of Engineering and Surveying & Australian Centre for Sustainable Catchments
University of Southern Queensland, Toowoomba 4350 QLD, Australia
mcdougak@usq.edu.au

ABSTRACT

The aim of this project was to assess the feasibility of a photo imaging approach in the assessment of vegetation condition attributes in comparison with the Queensland Department of Environment and Resource Management (QDERM) Vegetation Condition Assessment approach. The project applied both the QDERM *BioCondition* monitoring method and an alternative photo imaging analysis approach to the assessment of vegetation condition attributes on different regional ecosystems. The study focused on the comparability of results and the cost-effectiveness of the photo imaging approach in comparison to the standard *BioCondition* assessment method.

Six regional ecosystems (RE) sites, of varying structural complexity, were selected for this study. They contained a diverse range of different vegetation attributes. Each site was assessed using both approaches. A low-cost Canon *PowerShot SX10 IS* camera was used to collect the photographs. It is equipped with a 20x Optical Zoom lens with a focal length of 5.0-100mm that allows shooting a scene from wide-angle to telephoto. For the ground cover data collection, two vertical-down photographs were taken of each quadrat. To capture canopy cover, vertical-up photographs were taken at 5m intervals along the transect midline. A two pole photographic method was developed to estimate tree canopy height.

This study found that the use of photo imaging methods to measure most attributes of vegetation for the *BioCondition* approach is technically possible. However, their application for operational use in ecosystems with closed vegetation canopies is not feasible. The estimation of the vegetation condition variables is constrained, in various degrees, by several factors. The estimation of shrub species richness in grassland and open canopy forests are the only attributes that have potential for operational use. Canopy cover estimates from vertical-up photographs produced comparable *Tree Cover Rank* results compared to the manually based crown cover estimate method. The photographic technique also has good potential for estimating major classes of ground cover in quadrats. Canopy height can be estimated more easily by using a laser range finder than a photo imaging method. In the future, when the cost of data acquisition becomes less expensive, the suitability of a LiDAR system could be considered to quantify the desired vegetation attributes.

KEYWORDS: photo imaging, *BioCondition*, vegetation assessment, low-cost camera

1 INTRODUCTION

The loss, decline and fragmentation of habitat through excessive clearing of native vegetation poses a significant threat to native flora and fauna (e.g. Ford et al., 2009; Maron and Fitzsimons, 2007). This has been occurring in many parts of the world (e.g. in Southeast Asia, see review of Koh and Sodhi, 2010). In Queensland, Australia, the impacts of land clearing on native biota include reduced abundance, localised extinctions and declining viability of populations (Cogger et al., 2003). To abate this problem, conservationists and resource managers need to intensify habitat protection and recovery programs, in tandem with other conservation measures. In pursuing these programs, habitat areas need to be identified, mapped, and their condition assessed.

Habitat condition can be assessed at a range of spatial scales, i.e. from site to regional scales (Gibbons, et al., 2006). These include the following three approaches: on-ground site assessment, spatial modelling, and remote sensing. At a site scale, Gibbons and Freudenberger (2006) reviewed the different tools and techniques for rapid, on-ground assessments of vegetation condition and suggested a framework on developing new approaches. While on-ground assessment of vegetation condition has several uses, this approach is very time consuming and resource-intensive. It is therefore logical to develop innovative methods that can be used to reduce the volume of work and time required without compromising the completeness and accuracy of key information that site-based methods can provide.

In Queensland, the former Environmental Protection Agency (EPA) (now part of the Department of Environment and Resource Management) has conducted several studies to develop methods to survey, classify and map different vegetation communities. The *BioCondition* assessment technique was one of those vegetation condition methods developed for Queensland (Eyre et al., 2006). It provides a measure of how well a terrestrial ecosystem is functioning for the maintenance of biodiversity values. This site-based method considers the structural, compositional and functional aspects of a vegetation community. As it depends on field-based site assessment of vegetation attributes, the amount of time and volume of work can be prohibitive when multiple sites need to be assessed.

The aim of this project was to assess the practicality, suitability, comparability and cost-efficiency of a photo (photographic) imaging approach in the assessment of habitat condition parameters in comparison with the DERM *BioCondition* (field survey) approach. This project applied both the *BioCondition* monitoring method and an alternative photo imaging analysis approach to the assessment of selected habitat condition parameters for a variety of regional ecosystems. The study focused on the comparability of results and the cost-effectiveness of the photo imaging approach in comparison to the standard *BioCondition* method. Six regional ecosystems (RE) sites (i.e. 11.3.21, 11.8.3, 11.8.15, 11.9.4a, 11.9.6, and 12.5.13) were included in this study.

2 USING TERRESTRIAL PHOTOGRAPHY IN BIOCONDITION ASSESSMENT

2.1 The *BioCondition* Approach

Information about the extent and condition of vegetation is necessary for integrated catchment management. Consequently, forest and vegetation assessment and monitoring programmes become integral to resource management efforts by government agencies and for environmental assessment purposes around the world (e.g. FAO, 2007). In Queensland, two major vegetation mapping programs exist: a) the Statewide Landcover and Trees Study (SLATS) and b) the Regional Ecosystems mapping program, both conducted by the Queensland Department of Environment and Resource Management. More recently, the former EPA developed the *BioCondition* framework for vegetation condition assessment.

The aim of the *BioCondition* assessment toolkit is to provide a framework that provides a measure of how well a terrestrial ecosystem is functioning for the maintenance of biodiversity values compared to its undisturbed condition (Eyre et al., 2008; 2006). It is a site-based, quantitative and repeatable assessment procedure that provides a numeric score to each prescribed vegetation attribute. It considers the structural, compositional and functional aspects of a vegetation community.

The BioCondition methods are basically field-based, i.e. data should be collected at the field level. While on-ground assessments can be accurate at fine scales, they can be impractical for assessment across broad scales due to high expense of manpower, resources and time. Thus, it will be beneficial if other techniques can be developed to alleviate some of the key issues with the existing field-based approach. One of the potential techniques that can be used to assess BioCondition and reduce labour is terrestrial photography. Photographs can capture features of interest which can be analysed off-site, thereby providing opportunities to reduce time and effort during field work.

2.2 The Potential of Terrestrial Photography

Several studies have been done using ground based photographs taken using an ordinary digital/manual camera. For instance, the Department of Natural Resources and Water used photographic methods to keep visual records of land features to monitor short and long-term physical change at each location (NRW, 2006b). They have named it as “photopoints”. Photopoints are permanent or semi permanent sites set up from where a series of photographs can be taken over time. For this purpose, the following ways of taking photographs were suggested:

- *Spot Photograph*: (near) vertical down photographs of specific locations from 1.6 m above the ground. This is for recording ground cover and species, and organic litter and bare soil.
- *Trayback Photograph*: taken standing on a vehicle tray back providing an elevation of approximately 3 metres. This is used primarily for assessing ground cover and condition.
- *Landscape Photograph*: they are used for showing shrub or tree layers, or the extent of events on the landscape such as floods or fire.

Images from photopoints can provide a valuable supporting record when monitoring the following (NRW, 2006b):

- Pasture condition, pasture species and yearly pasture use
- Ground cover, organic litter, shrub cover, recruitment of woody plants, tree canopy cover and health, and vegetation density
- Native vegetation area and wetland area
- Native plant richness, large trees, fallen woody material and in-stream habitat
- Impacts on native vegetation, impacts on wetlands
- Farm water flow, gully erosion, hill slope erosion and wind erosion
- Saline land and deep-rooted perennials
- Weed cover and weed species
- Effects of fire, drought, flood, dieback and feral animals
- Wind erosion

In a different application, Gilbert et al. (2009) used a digital camera to monitor *Calluna vulgaris* after a fire. The two trials undertaken, artificial and field-based, demonstrated the value of using digital photography as a tool in measuring vegetation cover. Comparing results from the digital and point quadrat methods indicated that they were not significantly different ($P > 0.05$), permitting confident use of the digital technique. Enhanced speed of data collection was most useful in areas of poor climatic conditions or poor accessibility, such as on upland moors. Less time in the field reduces the effect of sampling fatigue on the results.

Time spent on computer analysis of images can be conveniently interrupted within a comfortable environment.

A study on grassland biomass estimation using ground based digital photographs was conducted in the U.S. (Vanamburg et al., 2005). The results showed that conventional (RGB) digital camera imagery was not useful on a shortgrass prairie for the estimation of aboveground biomass. The complex spectral characteristics of shortgrass prairie systems, especially as vegetation begins to senesce, limited the usefulness of this type of sensor for spectrally distinguishing among substrate components like soil, litter and brown vegetation. Detailed analysis of these data sets showed that shortgrass prairie ecosystems are spectrally very complex. Yet, as vegetation began to senesce throughout the season, these ecosystems exhibit a large amount of spectral overlap among substrate components. This caused considerable error between classes such as brown vegetation and soil, which increased classification error.

In another study, a technique for near ground remote sensing of herbaceous vegetation in tropical woodlands was developed (Northup et al., 1999). The procedures they applied were found efficient in the open eucalypt woodlands of northern Queensland. The technique was relatively cost-effective, and thought to be equally capable in ecosystems with open woody canopies, or in grasslands.

3 METHODS

3.1 Study Area

The study area covered part of the Condamine Catchment (Figure 1). The catchment is located west of the Great Dividing Range in southern Queensland, covering an area of 24,434 km². The area has a highly variable subtropical climate, with an average annual rainfall of 682-955mm, and average temperatures ranging from 3°C to 30°C (NRW, 2006a). The vegetation in the basalt hills is dominated by mountain coolibah, narrow-leaved ironbark and silver leaf ironbark. In soils associated with sandstone areas, patches of brigalow/belah, poplar box, ironbark, bullock and cypress pine are common. The extensive use of the area for agriculture and pasture has resulted in the loss of much of the original vegetation.

This study investigated six regional ecosystem (RE) types of varying structural complexity (Table 1). The vegetation communities include grassland, woodland, open forest, vine thicket and vine forest. These were selected to represent the range of assessable vegetation attributes in which the photo imaging method was to be tested. The grasslands (e.g. RE 11.3.21) were easier to assess than the vine thickets and vine forests (e.g. RE 11.8.3 and 12.5.13).

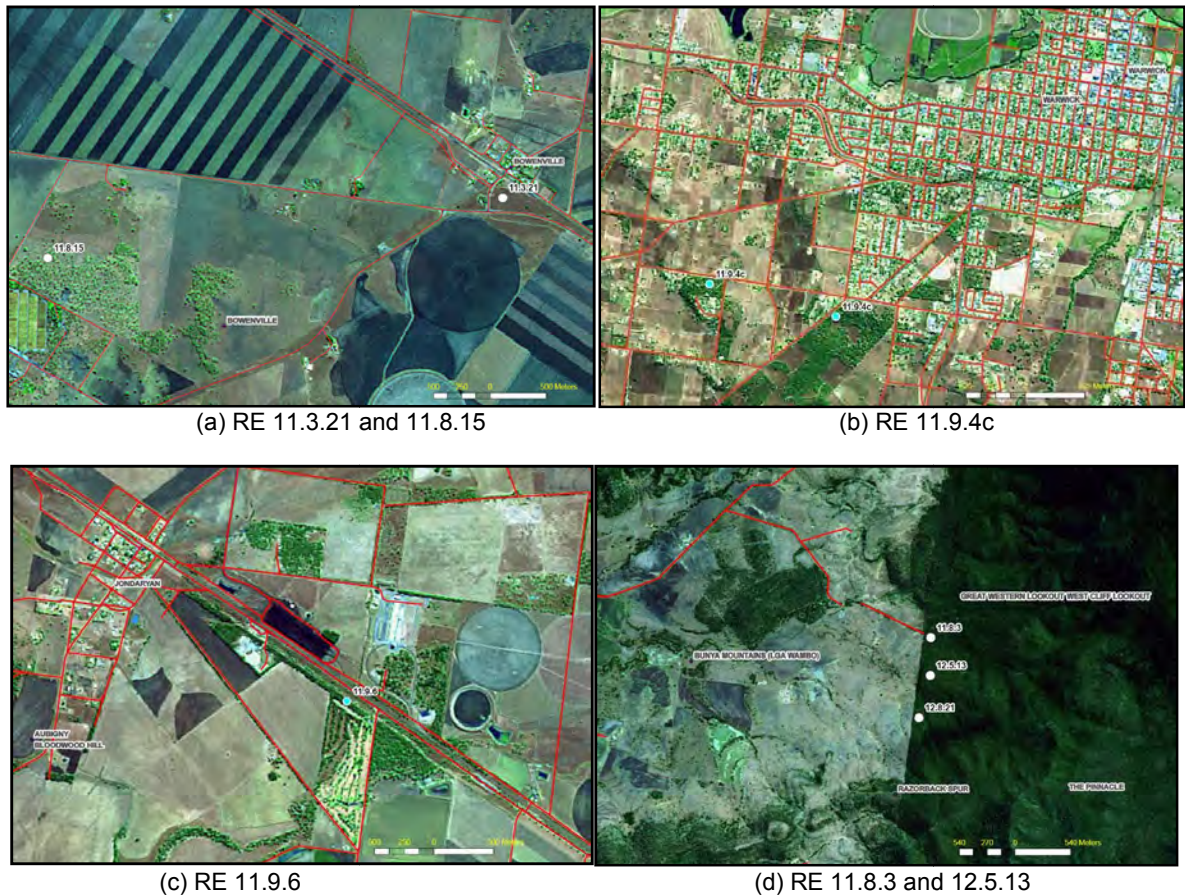


Figure 1: Location and image of the study area.

3.2 Field Data Acquisition

Sample data for assessment of the *BioCondition* of each Regional Ecosystem (RE) site (see Figure 1) were collected by similar methods to those outlined in the Queensland *BioCondition* Assessment Manual (BAM) (Eyre, et al., 2008). The methods were modified by the addition of increased photographic data collection as outlined in section 3.2.3 below.

3.2.1. Transect Data Collection

The position of each 100 x 50 m transect for each RE was recorded using a Garmin Explorer GPS unit. Full Transect data were collected over the entire 100 x 50 m area of each transect. Sub-transect data were collected from a 50 x 10 m area in the centre of each transect. Large Tree measurements were based on up to 10 large trees selected at random (less if 10 trees were not present). Tree and Shrub canopy cover was assessed along the 100m centreline of all RE transects except for REs 11.8.3 and 12.5.3. Canopy cover was only collected along 50 m of centreline for these two REs. The thickness of the vegetation made it impractical to collect data along a 100 m centreline. Canopy cover (both tree and shrub) was recorded by vertical projection over the centreline. As such, it equates to crown cover (Walker and Hopkins (1990) as cited by Eyre et al., 2008). Canopy health scores were also collected according to the BAM procedures (p. 18). The height of the trees was measured using a “TruPulse™” 200 laser range finder.

Table 1 Regional ecosystems sampled in this study

Regional Ecosystems	Short Description	Structure Category	Location of Sample Sites
11.3.21	<i>Dichanthium sericeum</i> and/or <i>Astrebla</i> spp. grassland on alluvial plains. Cracking clay soils	grassland	Bowenville
11.8.15	<i>Eucalyptus brownii</i> or <i>Eucalyptus populnea</i> woodland on Cainozoic igneous rocks	sparse	Wainui
11.9.4a	Semi-evergreen vine thicket or <i>Acacia harpophylla</i> with a semi-evergreen vine thicket understorey on fine grained sedimentary rocks	dense	Warwick
11.9.6	<i>Acacia melvillei</i> +/- <i>A. harpophylla</i> open forest on fine-grained sedimentary rocks	mid-dense	Jondaryan
11.8.3	Semi-evergreen vine thicket on Cainozoic igneous rocks	dense	Bunya Mountain
12.5.13	Microphyll to notophyll vine forest +/- <i>Araucaria cunninghamii</i>	dense	Bunya Mountain

3.2.2. Ground Cover Data Collection

Ground cover data were collected from five 1mx1m quadrats located at 10 m intervals along the transect centreline beginning at 25 m from the origin. At each quadrat (Figure 2), the percentage of vertical cover for each of 10 categories (Eyre et al. 2008) was estimated by averaging the estimates of two observers. Photographs of cover over each quadrat were taken in a standard manner (see Section 3.2.3 below). The manually collected data from 5 quadrats was averaged for each transect.



Figure 2: A 1m x 1m quadrat used in the field

3.2.3. Photographic Data Collection

From the literature review, brainstorming, and pre-testing of the methods, the team developed the following photographic data collection techniques:

- a) A 10 megapixel Canon *PowerShot SX10 IS* was used for this study (Figure 3). This camera features 10-megapixel resolution and new DIGIC 4 Image Processor for high-performance face and motion detection. It is equipped with a 20x Optical Zoom lens with a focal length of 5.0-100mm (35mm film equivalent: 28-560mm) that allows shooting a scene from wide-angle to telephoto. It cost approximately \$460 in 2009.



Figure 3: A 10 megapixel Canon *PowerShot SX10 IS*.

- b) For the ground cover data collection, two vertical photographs (with one as back up) were taken of each quadrat. A special set up was designed to make the photograph vertical (Figure 4).



Figure 4: Set up of the camera.

- c) For the canopy cover, a vertical upward photograph was taken at every 5m interval along the 100m transect to estimate the canopy cover from photographs.

- d) For the tree canopy height, two 2m graduated ranging poles were erected at the 25m and 50m points on the transect. A horizontal photograph of the site was taken from 10m away from the starting point of the transect. A similar record was made at the end of transect. These photographs were used to calculate the height of trees.
- e) Four photographs were taken in four 90 degree directions from midtransect to provide a reference record of the area.

3.3 Data Processing and Analysis

Manually collected data were analysed according to the Queensland *BioCondition* Assessment Manual (BAM) procedures (Eyre, et al., 2008). Different procedures were explored and tested for analysing the photographic data to evaluate the utility of inexpensive and readily available software. This study used *Pixcavator*, *Adobe Photoshop*, and *Adobe Acrobat Professional*.

3.3.1. Transect Data Analysis

a) Manual Analysis

The full transect and sub transect values were recorded directly from the field. Large Tree data (for REs with large trees) were recorded for up to 10 specimens selected at random within the transect. The circumference at breast height (CBH at 1.3m) was recorded and used to calculate the diameter at breast height (DBH). This was averaged for each sample.

The canopy vertical intersect distances along the midline were converted to crown cover percentages. This was expressed as a percentage of the reference crown cover percentage for each RE (rel %) and used to identify the applicable row in Table 4 of the *BioCondition* Assessment Manual. The percent of crown health scores ≥ 3 was calculated and used to identify the applicable column in Table 4. The intersection of the Crown Cover percentage and the Crown Health percentage identifies the overall tree cover health of the ecosystem being investigated.

b) Photographic Data Analysis

To interpret the images taken in the field, the software *Pixcavator* was used. This software allowed us to choose a channel and classify the image using an object-based approach. To determine the percentage cover of the canopy, the blue channel was selected to process the photograph. The blue channel was found suitable for canopy cover estimation as it has a low response to vegetation components and hence gives more dark pixels during classification. After the application of the edge enhancement ("dilation") filter, the image was classified. Generally the software divides the whole image into dark and light classes, i.e. a group of patches of bright or dark pixels in the image based on given threshold values). From this, the percentage of canopy cover can be estimated.

The classification output resulted in three general possibilities (Figure 5):

- a. All canopy cover was classified as dark class and blue sky (not the open sky) seen through the canopy as light class. In this case,
Canopy cover = dark class – light class
- b. In the case where most of the photograph was covered by canopy, then only the holes of blue sky were classified as light class, in this case,
Canopy cover = 100 – light class
- c. Only the canopy cover was classified as dark class, in this case,
Canopy cover = dark class

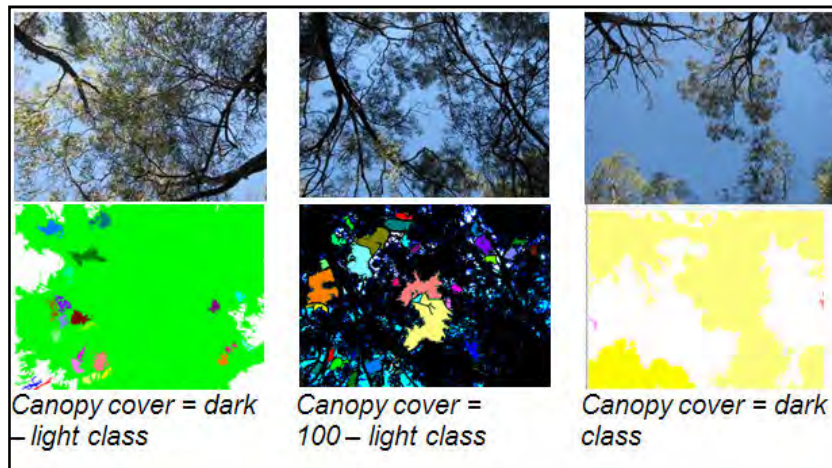


Figure 5: Three potential classification results for canopy cover calculation in *Pixcavator*

2. The average of the canopy cover values was calculated and compared to the result from the manual method.

3.3.2. Ground Cover Data Analysis

Ground cover was estimated in three different ways: manual field estimation, photographic estimation and gridded photographic estimation. Other methods such as pixel classification with Adobe Photoshop software and object classification with Pixcavator software were explored but yielded impractical results. Pixcavator was not found to be a good software to classify different types of ground cover features due to the limited capacity of the software to create multiple classes. Human knowledge of vegetation and cover types (by their texture, pattern and shape) and the colour of features were found more important to analyse the photograph.

a) Manual Field Estimation

Two observers inspected each of the 5 quadrats in each transect and agreed on the percent of each cover type. The figures for each transect were averaged and then grouped into the three *BioCondition* ground cover categories: perennial grasses, perennial non grasses and annual species as well as organic litter cover. The values were compared with reference ecosystem values for each RE.

b) Photographic Estimation

A standard photograph of quadrat 3 from each transect was displayed on a nominal 24" LCD screen using Adobe Acrobat 6.0 software. An informed observer visually estimated the percent of each type of ground cover.

c) Grid Photographic Estimation

A standard nominal grid (1 cm cell size) was displayed over the same photograph using Adobe Acrobat 6.0 (Figure 6). Seven cover class stamps were created and each grid cell was stamped according to the type of majority cover in the cell. The percent of each ground type cover was calculated by summing each class stamp and expressing it as a percentage.



Figure 6: Grid photographic estimation method of ground cover

3.4 Assessment of the Methods Used

As indicated before, the aim of this project was to assess the practicality, suitability, comparability and cost-efficiency of photo imaging approach in the assessment of habitat condition parameters with reference to the *BioCondition* approach. In this regard, we assume the following:

- The photo imaging method should be as simple as possible, i.e. the sensor is a low-cost (<\$500), “off-the-shelf” camera system. This cost limitation precludes the use of a multi-spectral band (e.g. near infrared channel) camera.
- The software system used in digital analysis of the photographs should be inexpensive (<\$500) and relatively easy-to-use, i.e. without the need to perform complex image processing tasks.
- The analyst user of the system does not need to be an expert in image processing, but only to have a basic knowledge of graphics and digital data handling.

In the assessment of the potential photo imaging methods, this study considered two main criteria:

- technical possibility – this refers to the prospect that a parameter (e.g. canopy height) can be measured using the photo imaging method regardless of the resources (time, labour, supporting equipment, etc.)
- operational feasibility – this refers to the prospect that a parameter (e.g. canopy height) can be measured using the photo imaging method within the confine of set resources or costs. A method was considered not operationally feasible if the cost (labour, software, hardware, user training, time, etc.) of implementing it was equal to or greater than the standard *BioCondition* method.

4 RESULTS

4.1 Estimating Transect and Sub-transect Attributes

Table 2 shows the manually collected field attribute measurements for each of the six RE sample sites and the investigating team's evaluation of the technical and operational feasibility of collecting similar information from extensive site photographs. These results show that in grassland and open woodland (savannah) ecosystems (REs 11.3.21, 11.8.15 and 11.9.4c), it is technically possible to extract each attribute with a medium level of difficulty. However it is not operationally feasible because it would take too much time and software skill to do it. Estimation of the canopy height of trees was an exception to this finding in that it was very possible (high rating) to calculate these accurately from photographs. However, because of the time and skill required, it was judged as operationally unfeasible.

Table 2 Estimates of the Feasibility of Measuring Transect and Sub-Transect Attributes from Photographs compared to Field Measurements.

RE category	Measurement Type	Full Transect					Sub Transect						
		Lge Tree spp #	Recruit %	Lge tree c hol	DBH (cm)	Canopy height (m)	Shrub spp #	Grass spp #	Herb + Forb spp	Other spp #	Weed cvr %	Fallen logs #	Fall logs decay %
RE 11.3.21 (Bowenville)	Field Measurement	No trees in this regional ecosystem transect sample*					0	2	5	1	0	No trees in this regional ecosystem sample*	
	Reference Value						0	7	11	4	0		
	Photo Meas.(Tech Possible)						Med	Med	Med	Med	Med		
	Photo Meas. (Oper Feasible)						Yes	No	No	No	No		
RE 11.8.15 (Wainui)	Field Measurement	1	100	13	62.39	10.32	1	2	5	1	1	2	13
	Reference Value	1	-	-	54	22	5	6	13	2	0	40	-
	Photo Meas.(Tech Possible)	Med	Med	Med	Med	High	Med	Med	Med	Med	Med	Med	Med
	Photo Meas. (Oper Feasible)	No	No	No	No	No	Yes	No	No	No	No	No	No
RE 11.9.4c (Warwick)	Field Measurement	1	5	2	38.52	7.94	2	5	2	0	5	8	0
	Reference Value	4	-	-	38	16	8	8	5	5	0	6	-
	Photo Meas.(Tech Possible)	Med	Med	Med	Med	High	Med	Med	Med	Med	Med	Med	Med
	Photo Meas. (Oper Feasible)	No	No	No	No	No	No	No	No	No	No	No	No
RE 11.9.6 (Jondaryan)	Field Measurement	5	100	12	34.03	6.96	1	2	4	0	1	>10	>10
	Reference Value	3	-	-	17.3	15	4	8	6	4	0	3	-
	Photo Meas.(Tech Possible)	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
	Photo Meas. (Oper Feasible)	No	No	No	No	No	No	No	No	No	No	No	No
RE 11.8.3 (Bunya 1)	Field Measurement	4	100	1	68.44	5.21	5	1	2	2	20	10	50
	Reference Value	12	-	-	22	7	17	2	3	15	0	9	-
	Photo Meas.(Tech Possible)	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
	Photo Meas. (Oper Feasible)	No	No	No	No	No	No	No	No	No	No	No	No
RE 12.5.13 (Bunya 2)	Field Measurement	4	100	4	37.24	4.92	4	1	4	1	20	30	55
	Reference Value	19	-	-	32	23	23	1	4	13	0	23	-
	Photo Meas.(Tech Possible)	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
	Photo Meas. (Oper Feasible)	No	No	No	No	No	No	No	No	No	No	No	No

Med = Medium difficulty

* The assessment from photos would be technically and operationally feasible if a few scattered trees occurred in what was otherwise open grassland.

In closed woodland ecosystems (REs 11.9.6, 11.8.3 and 12.5.13), there was a low level of technical possibility of extracting the attribute information accurately and it is not operationally feasible to extract such information.

The photogrammetric method yielded the same Tree Cover Rank as the manual method for all REs (Table 3). This occurred despite substantial differences between the canopy cover percentages derived from the vertical-up photographs and the crown cover percentages recorded along the transect midline (see column 4). The relative cover percentages (measured relative to RE reference values) all fell within category 3 in Table 4 of the BAM. The Health Scores and Relative Health Scores from the photographs were similar to the field recorded values. This resulted in all REs having less than 30% of their canopies with a Health Score greater than 3.

Table 3 Estimates of Canopy Cover Attributes for Five Sample Regional Ecosystems

RE category	Method	Value Type	Crown Cover %	Health Score (%≥3)	Tree Cover Rank ¹
RE 11.3.21 (Bowenville)	Existing method	Measured	No trees in this regional ecosystem transect sample		
		Relative			
	Reference value				
	Pixcavator method	Measured			
		Relative			
RE 11.8.15 (Wainui)	Existing method	Measured	54		4
		Relative	130.1	31	
	Reference value		41.5		
	Pixcavator method	Measured	27.15		4
		Relative	65.4	33	
RE 11.9.4c (Warwick)	Existing method	Measured	66.5		5
		Relative	99.3	14	
	Reference value		67		
	Pixcavator method	Measured	40.4		3
		Relative	60.3	11	
RE 11.9.6 (Jondaryan)	Existing method	Measured	84.8		5
		Relative	101	0	
	Reference value		84		
	Pixcavator method	Measured	83.97		5
		Relative	100.0	0	
RE 11.8.3 (Bunya 1)	Existing method	Measured	61		5
		Relative	64	7	
	Reference value		95		
	Pixcavator method	Measured	64.19		5
		Relative	67.6	20	
RE 12.5.13 (Bunya 2)	Existing method	Measured	74		5
		Relative	82	0	
	Reference value		90		
	Pixcavator method	Measured	81.18		5
		Relative	90.2	0	

1. Derived from Table 4 of the BioCondition Assessment Manual

The average of the canopy cover values calculated from photographs was compared to the result of the manual method. For REs 11.9.6, 11.8.3, and 12.5.13, the values calculated from the photographs and the manual method were relatively close, i.e. 83% vs. 84%, 64% vs. 61%, and 81% vs. 74%, respectively. In these cases, it suggests that the photo imaging method is capable of producing comparatively accurate results. However, for RE 11.8.15, there is a large difference between the values: 27% vs. 54%.

4.2 Estimating Ground Cover Attributes

Table 4 provides illustrative results of two photographic methods (Photo and Gridded Photo) and the Field Method for measuring the percentage of four categories (BAM categories) of ground cover in 1m² quadrats sampled at the 45 m interval along the midline of each transect. The results indicate that it is difficult to make generalisations about the accuracy of the photo imaging methods. The percent deviation (or relative error) ranges from as low as 6% (perennial non-grass for RE 12.5.13) to a very high 167% (annual species for

RE 11.9.4c). In general, the grassland RE 11.3.21 exhibited the lowest relative error of estimation.

Table 4 Estimates and Percent Deviation (Relative Error) of Ground Cover Attributes by Three Different Methods

RE Category	Estimate Method	Ground Cover Type (Center Quadrat/Transect only)			
		Perennial grass	Perennial non-grass	Annual species	Organic Litter
RE 11.3.21 (Bowenville)	Field Estimate	45	25	0	25
	Photo Estimate	58	20	0	20
	Gridded Estimate	82	14	0	20
	Reference Value ¹	51	1.6	0	42
	% Dev Photo Estimate	29	-20		-20
	% Dev Gridded Estimate	82	-44		-20
RE 11.8.15 (Wainui)	Field Estimate	64	3	0	12
	Photo Estimate	80	0	0	5
	Gridded Estimate	87	0	0	8
	Reference Value	26.5	12	0	45
	% Dev Photo Estimate	25	-100		-58
	% Dev Gridded Estimate	36	-100		-33
RE 11.9.4c (Warwick)	Field Estimate	0	0	24	75
	Photo Estimate	0	0	50	50
	Gridded Estimate	0	0	64	36
	Reference Value	5.3	1	2	66
	% Dev Photo Estimate			108	-33
	% Dev Gridded Estimate			167	-52
RE 11.9.6 (Jondaryan)	Field Estimate	30	18	0	44
	Photo Estimate	7	15	0	73
	Gridded Estimate	21	7	0	69
	Reference Value	15	5.4	0	74
	% Dev Photo Estimate	-77	-17		66
	% Dev Gridded Estimate	-30	-61		57
RE 11.8.3 (Bunya 1)	Field Estimate	3	5	0	30
	Photo Estimate	0	0	0	60
	Gridded Estimate	0	0	0	34
	Reference Value	4.6	0.8	0	50
	% Dev Photo Estimate	-100	-100		100
	% Dev Gridded Estimate	-100	-100		13
RE 12.5.13 (Bunya 2)	Field Estimate	0	80	0	15
	Photo Estimate	0	85	0	5
	Gridded Estimate	0	85	0	12
	Reference Value	0	0.7	0	87
	% Dev Photo Estimate		6		-67
	% Dev Gridded Estimate		6		-20

¹Reference Values taken from the RE Fact Sheets (2007) published by QDERM (2010).

5 DISCUSSION

5.1 Estimating Transect and Sub-transect Attributes

This study found that while the majority of vegetation condition attributes can be estimated technically from photo imaging methods (particularly for grassland and open woodland ecosystems), their application for operational use in densely vegetated ecosystems is limited or not feasible. The following constraints were identified from this research:

- a) **Tree canopy cover and native shrub cover:** Despite the very high agreement between the manual method and photo imaging method in estimating the Tree Cover Rank (related to crown cover and health score), the photo imaging method is not feasible for operational use in densely vegetated areas. This is because (i) collecting vertical-upward photographs in dense vegetation was very time consuming due to the difficulty in moving the camera rig and its set up, and (ii) the analysis of photographs and calculations of the percent crown cover took a long time, even with experienced analysts.

- b) **Large trees:** as this involved counting the number of large (as identified by a DBH threshold) living trees per hectare within 100 x 50 m assessment area, it was not practical to take photos of the entire plot in densely vegetated areas to selectively count the large trees. In contrast, the assessment from photos would be technically and operationally feasible where a few scattered trees or shrubs occurred in what was otherwise open grassland.
- c) **Tree canopy height:** technically, it was possible to accurately estimate this parameter from photo imaging method as shown in this study. However, for densely vegetated areas, this method became impractical for routine operational use. One key reason was the difficulty in seeing the measuring poles in a perspective where the top of the canopy was captured from a distance. Understorey growth prevented the measuring poles to become visible on the photos, thus calculation did not become feasible. It was demonstrated in this study that the use of a laser range finder (e.g. “TruPulse™” 200) is more suitable for measurement of tree height than the photo imaging method. The cost of such equipment ranged from \$800-\$1,500 in 2009.
- d) **Native plant richness by lifeform group:** with the *Biocondition* method, assessment involves counting the number of native species observed in the 50 x 10 m plot, for five life-forms: trees, shrubs, grasses, forbs and others. Using the photo imaging method for densely vegetated area, the technical possibility is relatively low due to the dual need of photographing vegetation comprehensively (i.e. covering the required areal extent) and the botanical or dendrological skill needed to identify different plant species. However, for grasslands and sparse forests, the photo imaging method has operational use for determining shrub species richness, as they “stand out” from the grass matrix.
- e) **Fallen woody material:** As it involved counting the number of fallen woody logs and other debris found within the 50 x 10 m plot, it is not operationally feasible to use photo imaging methods for this purpose in densely vegetated areas. Even for grassland and open forest areas, locating and counting the number of fallen woody materials from photographs was difficult because grasses and forbs covered some of the material.

5.2 Estimating Ground Cover Attributes

In estimating organic litter and ground cover (native grasses and others) using the photographic method, grassland RE 11.3.21 produced the lowest relative error of estimation. In some REs, the accuracy of ground cover estimates is relatively high. However, the results of this study are inconclusive for regional ecosystems with dense vegetation. The range of percent deviation is large, i.e. 6% to 167%. In hindsight, it may be possible that the field estimates, considered as the “acceptable” true value, may not be the case, as they are also prone to human estimation error. It may be that the photographic method provided a more objective method. This would need further investigation. Nevertheless, this study showed that the photographic technique has a good potential for operational use in estimating organic litter and ground cover.

It appears that the two photo analysis methods (i.e. using the photographic canopy estimation method and the ground cover photographic estimation) yield results that are generally close to each other. Thus, it seems that the added time it takes to do the grid analysis procedure is not warranted. The net effect is that the manual estimate from the photos could then be done using any freely available image viewer, such as Microsoft Office Picture Manager.

5.3 Other Relevant Issues

The *Biocondition* Assessment Manual describes a procedure for measuring “canopy cover”. It consists of expressing the vertical projection of the canopy as a percentage of the midline of a selected transect. The authors of that report acknowledge that they are using the term “canopy cover” synonymously with the term “crown cover”.

In this project, vertical-up photographs taken every 5 m along the transect midline were analysed in a software (Pixcavator) that identified areas of canopies that contained a majority of leaves. For most canopies, this was less than the area which encompassed the “crown cover” of the tree. The consequence of this is that for a given tree crown of 100%, the canopy cover was well less than 100%. As a consequence, the transect canopy percentages measured by Pixcavator, from 10 vertical-up photographs along the transect midline, may be expected to be less than the comparable percentage crown cover measurement and less than the referenced amount. This may warrant a new reference value for “canopy cover” percentage as distinct from “crown cover” percentage.

The existing BAM is based on on-site physical measurement of Regional Ecosystem attributes. On-site assessment is very time consuming and labour intensive. Application of new and existing technologies may offer the potential for more rapid assessment of vegetation condition. Such technologies include ground based stereo imagery, NIR imagery and LiDAR scans, large scale aerial stereo photographs and new high resolution satellite imagery such as DigiGlobe II. The new technologies offer the opportunity to collect data about current and new attributes. It would seem appropriate to investigate the range of such new attributes that might be collected by using these emerging technologies, as well as their cost effectiveness, their relationship to the existing benchmark attributes and values, and the possible development of new RE reference values specific to these technologies.

6 CONCLUSION

This study found that the use of photo imaging methods to measure most attributes of vegetation under the BioCondition approach is technically possible. However, their application for operational use in ecosystems with closed vegetation canopies is not feasible. The estimation of the vegetation condition variables is constrained, in various degrees by, a) the difficulty in setting up and operating the camera system and related equipment in thick vegetation, b) the inexpensive, “off-the-shelf” RGB camera (vs. multi-spectral NIR-equipped) system is limited in its ability to discriminate photosynthetic and non-photosynthetic vegetation features, and c) the low-cost software system does not provide sufficient functionality to conduct the desired image discrimination tasks.

For the parameters considered in this study, the estimation of shrub species richness in grassland and open canopy forests are the only attributes that have potential for operational use. Canopy cover estimates from vertical-up photographs produced comparable Tree Cover Rank results compared to the manually based crown cover estimate method. The photographic technique also has good potential for estimating major classes of ground cover in quadrats. Canopy height can be estimated more easily by using a laser range finder than a photo imaging method. This study recommends that for the estimation of quadrat-based variables, i.e. for ground cover and litter, the use of an NIR camera be tested. These cameras have more ability to discriminate between photosynthetic and non-photosynthetic vegetation and soil. In the future, when the cost of LiDAR data acquisition becomes more economical, it should be considered for landscape scanning (horizontal application) for tree size and density and for canopy height and health.

ACKNOWLEDGEMENTS

We thank the *Condamine Alliance* for funding this project. Special thanks to Warwick Regional Council and the landowners who granted us permission to access some of the sites. We appreciate the help of Dev Raj Paudyal in field data collection.

REFERENCES

Cogger, H., Ford, H., Johnson, C., Holman, J., Butler, D. (2003). Impacts of Land Clearing on Australian Wildlife in Queensland. WWF Australia, Brisbane.

Eyre, T.J., Kelly, A.L, and Neldner, V.J. (2008). BioCondition: A Terrestrial Vegetation Condition Assessment Tool for Biodiversity in Queensland. Field Assessment Manual. Version 1.6. Environmental Protection Agency (EPA), Biodiversity Sciences unit, Brisbane.

Eyre, T.J., Kelly, A.L, and Neldner, V.J. (2006). Methodology for the Establishment and Survey of Reference Sites for BioCondition, Environmental Protection Agency, Biodiversity Sciences Unit, Brisbane.

FAO (Food and Agriculture Organization of the United Nations) (2007). National Forest Monitoring and Assessments, <http://www.fao.org/forestry/site/nfma/en/>, Accessed 29 December 2007.

Ford, H.A., Walters, J.R., Cooper, C.B., Debus, S.J.S. and Doerr, V.A.J. (2009). Extinction debt or habitat change? – Ongoing losses of woodland birds in north-eastern New South Wales, Australia, *Biological Conservation*, 142: 3182-3190.

Gibbons, P. and Freudenberger, D. (2006). An overview of methods used to assess vegetation condition at the scale of the site, *Ecological Management & Restoration*, 7, 11-17.

Gibbons, P., Zerger, A., Jones, S., and Ryan, P. (2006). Mapping vegetation condition in the context of biodiversity conservation, *Ecological Management & Restoration*, 7, 1-2.

Gilbert, J.A. and Butt, K.T. (2009). Evaluation of digital photography as a tool for field monitoring in potentially inhospitable environments, School of Built and Natural Environment, University of Central Lancashire, UK.

Maron, M. and Fitzsimons, J.A. (2007). Agricultural intensification and loss of matrix habitat over 23 years in the West Wimmera, south-eastern Australia, *Biological Conservation*, 135: 603-609.

Northup, et al., 1999, CSIRO Division of Tropical Agriculture, Davies Laboratory, PMB, Aitkenvale, QLD 4814, Australia.

NRW (Natural Resources and Water) (2006a). The Condamine Catchment (Fact Sheet), Queensland Department of Natural Resources and Water, Queensland.

NRW (Natural Resources and Water) (2006b). Land manager's Monitoring Guide, Photopoint Monitoring, National Resource and Water, Queensland Government.

Vanamburg, et at. (2005). Ground based digital imagery for grassland biomass estimation, United States Department of Agriculture, Agricultural Research Service, Great Plains System Research Unit, Fort Collins, Colorado, 80521, USA.

Semi-Automated Colour Registration and Evaluation of Digital Photogrammetry and Terrestrial Laser Scanning

Kwanthar Lim¹, Kwang-Ho Bae² and David Belton¹

1. The Institute for Geoscience Research (TIGeR), Department of Spatial Sciences, Curtin University of Technology, GPO Box U1987, Perth, WA, 6845, Australia
Phone: +61 8 9266-7604, Fax.: +61 8 9266-2703
2. Fugro Spatial Solutions Pty Ltd, 18 Prowse Street, West Perth, WA, 6005, Australia
Phone: +61 8 9282-4100, Fax.: +61 8 9322-1775

l.kwanthar@curtin.edu.au, k.h.bae@fugrospatial.com.au, d.belton@curtin.edu.au

ABSTRACT

Photogrammetry and Laser Scanning fusion is becoming increasingly common. This paper outlines the evaluation and semi-automated registration of single colour image to laser scanning point cloud data using canonical transformation and Direct Linear Transformation (DLT) registration methods. Laser scanning acquires 3D data points and intensity values but is unable to directly obtain photorealistic colour in most cases. There are instances where digital images are taken of the object of interest with the intention to merge the 3D data and image to reconstruct a photorealistic digital representation. Currently limited methods exist for the registration of multi-sensor platforms; a common method seen requires specially designed camera mounting. Another possibility is to transfer colour information from 2D images to the 3D points using photogrammetric methods. This method was inspired by the SCI method (Forkuo and King, 2005); the registration process utilises synthetic imagery calculated from laser scanning point clouds and matched with a camera image for colour registration. Evaluation is necessary as it provides a metric indication of accuracy and precision. The proposed research intends to aid in heritage and city modelling, to further feature detection methods, to provide cost effectiveness in industrial applications and to potentially improve model visualisation times.

KEYWORDS: Direct Linear Transform, Fusion, Laser Scanning, Photogrammetry, Registration, Multi-sensor, Terrestrial, Matching, Camera Simulation

1 INTRODUCTION

Photorealistic reconstructions of cities, buildings and objects are aesthetically pleasing and provide a more complete representation than a simple 3D model with just intensity values. There are different methods that may be used to provide colour information to a 3D point cloud and texture mapping is one such method. In the perspective of a spatial scientist, metric information is of a key importance, thus a method providing a metric indication of colour accuracy is presented. This paper presents a method using synthetic images, a concept presented in the paper by Forkuo and King (2005) and another using Direct Linear Transformation to address this issue. The methods also presented use a single image for the purposes of registering colour onto point clouds.

In general, multiple images are utilised for 3D construction but, in the surveying application with terrestrial laser scanners, it may be possible to utilise 3D point clouds for estimating the depth information of some parts of 2D images. Then either with the DLT alone or the exterior orientation of a camera with its calibration parameter, it is possible to reconstruct additional 3D information next to the available 3D point clouds. Furthermore, it is also possible to obtain more depth information of the 2D images, which is located next to the laser scanner. This allows us to obtain much more detailed or additional information to the 3D point cloud, not just the colour information. In this paper, we present some preliminary methodology and results of this idea. For example, in cases of mobile laser scanning systems that have been gaining in popularity, a procedure to estimate the depth information of 2D images is important since it provides a way of obtaining metric information of objects other than 3D point clouds from terrestrial laser scanners that are usually installed next to a camera system, e.g. spherical multiple cameras. Although it may be possible to extract the depth information from multiple-images from this system, in practice, it is recommended for 3D terrestrial laser scanners to be utilised or assisted for this process because we may not solely rely on the camera location from GNSS and INS.

2 BACKGROUND

2.1 Photogrammetry

Photogrammetry encompasses the use of images to measure and interpret shapes and sizes and locations of objects (Luhmann et al. 2006). It is often used to make three dimensional (3D) digital representations or graphical representations of an object (Luhmann et al. 2006). 2D capturing methods such as cameras, scanners and digitisers are the main instruments used in photogrammetry for data capture. Space resection is a procedure used to calculate exterior orientation (Luhmann et al. 2006). Exterior orientation consists of six parameters describing the spatial positions and orientation of the camera with respect to the global object coordinate system (Luhmann et al. 2006). In order to produce a 3D representation of an object via photogrammetry, a minimum of two 2D images with sufficient corresponding points are necessary.

Instruments that contain a two camera setup or a single camera and a grid projector use photogrammetric principles to obtain 3D data. 3D photorealistic metric models can be created with such instruments, but reconstruction effort may be high. Furthermore, a targetless approach for this purpose has been investigated by scientific communities for some time. Although there are some available methods in specific conditions, a general solution is to be developed.

2.2 Laser Scanning

Laser scanners are the instruments used to acquire 3D data points (point cloud). The laser scanning device automatically captures a point cloud containing millions of points within a relatively short period of time to approximately millimetre accuracy (McGlone et al., 2004). Like photogrammetry, laser scanning data may be used for creating 3D models, but have limitations in terms of accurate colour representation. Colour representation on 3D point clouds still relies on instruments such as a camera. According to Jansa et al. (2004), laser scanners have the benefits of having high spatial resolution, very good spatial coverage, moderate reconstruction effort, high 3D point density and depth accuracy, while the limitations are its colour, texture reconstruction, and high instrument costs. In terms of cameras, the two main attributes that may be used to reduce the limitations of laser scanning are the ability to provide colour information and the means for texture reconstruction.

Presently laser scanners may not achieve the same precision as some electronic distance measurement devices (EDM) used in conventional surveying, but it is presently being used in conjunction with photogrammetry in the following applications:

- Architecture and Heritage preservation applications include city modelling, heritage preservation and restoration, and art and cultural analysis.
- Engineering and Surveying applications requiring measurement of deformations and change detection, tunnel profiles and concrete tanks, all of which require high measurement precisions

- Automotive applications requiring measurement of surface design models for parts analysis, for deformation and safety testing, and the inspection of parts.
- Industrial applications including pipe and machinery location for power stations; aircraft and aerospace requiring extremely high accuracy for measurement of corner fittings and mechanically and thermally stressed objects; and forensics using photogrammetry for crime scene and accident measurements and reconstructions and can potentially use laser scanning for a more efficient means of 3D reconstruction.

2.3 Registration

Registration is the process of transforming data from different sources and/or of different types into a reference system to allow for measurements and interpretations to take place. Generally registrations encompass components that relate to scale, skew, rotations and translations (Luhman et al., 2006). An example of the registration process can be seen in Kang, et al. (2007), Zitová and Flusser (2003) and Al-Manasir and Fraser (2006). In most cases registration takes place during the post processing of data, but some systems exist that do the registration processing on the fly (during data capture within the instrument) as presented in Jansa et al. (2004), Sapkota (2008), and Kern (2001). This paper proposes a method of registration for a single image to point cloud data. The currently available methods that exist are rigid systems, where the camera is mounted onto the laser scanning device, offering limited flexibility with the camera's properties, i.e. pixel resolution, colour quality, field of view and so on.

2.4 Projective Transformations

The similarity transformation and Direct Linear Transform (DLT) are two transformation methods used in this paper. The similarity transformation is a common method used for obtaining scale, rotation and translation parameter, but requires approximate initial values. DLT was introduced in 1971 and has the advantage of not requiring initial values (Abdel-Aziz and Karara, 1971), as it uses its own transformation coefficients. However, the DLT coefficients of transformation are not as versatile as the similarity transformation being that its coefficients aren't necessarily describing scale, rotations and translations. In this field of surveying and mapping, metric precision is valued over the aesthetics of image overlays and texture mapping. It is important to have the measurement data displayed accurately and precisely, so that it may reflect reality in terms of metrics. Outlined below are the DLT projection equation (1) and its transposed matrix format (2) as follows:

$$\begin{aligned} x_i &= (H_{11} * X_o + H_{12} * Y_o + H_{13} * Z_o + H_{14}) / (H_{31} * X_o + H_{32} * Y_o + H_{33} * Z_o + 1) \\ y_i &= (H_{21} * X_o + H_{22} * Y_o + H_{23} * Z_o + H_{24}) / (H_{31} * X_o + H_{32} * Y_o + H_{33} * Z_o + 1) \end{aligned} \quad (1)$$

and

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & 1 \end{bmatrix} * \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix} \quad (2)$$

where $(x_i, y_i, 1)$ and $(X_o, Y_o, Z_o, 1)$ are the 2D image and 3D point clouds in the homogeneous notation, respectively, and H_{ij} ($i=1...4$, for the 12 parameter cases) is the component of the DLT matrix. With the least squares method and DLT, we can estimate the DLT matrix (i.e. H_{ij}) up to scale (Hartley and Zissermann, 2003).

DLT is one of the main transformations investigated in this registration project for the assignment of colour information obtained from the digital image, which will be referred to as Real Camera Image (RCI). DLT primarily uses 11 parameters, but a 12 parameter solution is also possible with further investigation whereby the 1 in the 3rd row and 4th column is treated as a parameter (Hartley and Zissermann, 2003). In order to produce a relevant 2D-3D projection matrix, the 3rd row of the DLT matrix is required to be independently investigated. A 2D-3D

projection matrix in this paper refers to a matrix that allows the transformation of the 2D coordinates into 3D coordinates.

2.5 Synthetic Camera Image (SCI)

Creation of an SCI is required to provide the platform whereby a relationship can be determined between the 3D point cloud and the RCI. The SCI is a projection of the 3D point cloud into 2D to simulate a digital image, Forkuo and King (2005) provides the original explanation. The SCI created in this paper uses the 2D camera projection model (sometimes referred as canonical model; refer to Zeisserman and Hartley, 2003) as seen below:

$$\mathbf{r}_i = \mathbf{k} [\mathbf{M} \mid \mathbf{Tr}] \times \mathbf{r}_o \quad (3)$$

where $\mathbf{k} \in \mathbb{R}^{3 \times 3}$ is the camera matrix, $\mathbf{M} \in \mathbb{R}^{3 \times 3}$ and $\mathbf{Tr} \in \mathbb{R}^{3 \times 1}$ are the rotation and translation matrix, respectively, and \mathbf{r}_i and \mathbf{r}_o are the 2D image and 3D object points in the non-homogeneous notation, respectively. The camera projection model can be expressed similarly to the DLT projection transformation when transposed into matrix format.

2.6 2D Iterative Closest Point

This rigorous method is used for correspondences of points, lines and objects. (Besl and MacKay, 1992; Yang, C., and G. Medioni, 1992) A point based approach is implemented for 2D correspondences. Simulated data used in this paper is created mathematically, therefore having no systematic errors. These are 3D datasets simulating a perfect laser scanning system. It is assumed that trials with the simulated data after transformations will obtain a singular solution. Other variants of ICP exist that are based on other than point-to-nearest-point distance, which is the method used in this paper, as mentioned by Al-Manasir and Fraser (2006), Rusinkiewicz and Levoy (2001), and Vosselman and Maas (2010).

Outlined below in Figure 1 is a brief overview of the ICP procedure represented by pseudo-code; beginning with point selection, then calculating and applying the transformation components, and repeating the process until convergence is reached. Colour registration from 2D onto 3D may be quite common, using texture mapping and other mapping techniques. However, this paper intends to allow additional properties to be registered to single images from other sensor information, presents a method using DLT to attempt to preserve 2D image metric accuracy, while registering onto 3D point clouds.

```

1: procedure ICP (Points  $X1_{ij}$ ,  $X2_{ij}$ )
2: for iteration = 1 to user specified do
3:   for j = 1 to length of  $X1$  do
4:     find closest point (minimum distance) with the kdtree
5:     if distance < threshold distance
6:       then store corresponding points  $X1_{ij}$  and  $X2_{ij}$ 
7:     else end
8:   end
9:   calculate rotations and translations with least squares
10:  apply transformation
11:  reduce threshold distance
12:  if converged && small residuals
13:    then break
14:  else end
15: end procedure

```

Figure 1: Brief summary of the conventional ICP algorithm

The ICP and its variants have been utilised for the automated registration of 3D point clouds, medical images, and pattern recognition. The most important and difficult part of the ICP algorithm is to develop metric criteria, e.g. Euclidean and Mahalanobis distances with consideration of statistical inferences, for the corresponding points in either 3D or 2D datasets.

3 METHODOLOGY

3.1 3D reconstruction with the SCI and back-projection

The initial dataset uses simulated data for the process of this registration method with the assumption that it will provide a solution with a known outcome and a controllable environment (i.e. manual input of systematic errors and distortion profiles). The simulated dataset does not incorporate any simulated errors and it is assumed that it should achieve a near-perfect solution under these test conditions. Since a simulated dataset is used, an image of the object cannot be obtained using a camera, thus the camera image (RCI) will also need to be simulated. This simulated camera image will be referred to as a Synthetic Real Camera Image (SRCI), and it is created in the same process as the SCI. The difference between the SCI and the SRCI is that some distortion parameters (principal point offsets and small rotations) are introduced when the SRCI model is created. The SRCI is treated as an image taken in the real world (RCI), see Figure 2(b) and Figure 3(a).

Arbitrary parameters are given for the initial SCI, such as rotations, translations, focal lengths and pixel sizes. However, the SCI should have the parameters reflecting the RCI. The initial test case will assume that the SCI has converged closely to the SRCI. Therefore, the parameters are set so that there will be only a slight shift in the SRCI to the SCI, as shown in Figure 2(a). In the case of an RCI under normal circumstances, exterior orientation parameters can be obtained via methods based on co-linearity equations or projective relations (Luhman et al., 2006), and calibration parameters should be applied to remove the systematic errors from the image. The correspondence process aims to bring SCI to correspond closely to the RCI. The SCI contains parameters to transform 2D back to 3D. There are several methods to attempt to bring the SCI into the same viewpoint as the RCI, such as finding exterior orientation parameters via resection process or using a least squares DLT method which will be discussed later on. In the paper by Forkuo and King (2005), the SCI image obtained contains slight differences to the RCI. These differences are assumed to be caused by the systematic errors of the laser scanner and the random errors during the transformation estimation process, such as the varying width of a light post. To respond to the SCI errors causing irregularities to the RCI/RSCI, ICP is performed to match the two datasets. The 2D ICP method used in the initial dataset is based on corresponding control points in the SCI and RCI. However other ICP methods may be implemented for non-simulated datasets to achieve better results based on the object scanned.

The reverse transformation parameters are obtained by performing a pseudo inverse on the projection matrix, i.e. Eqs. 1-3. The SCI contains the projection parameters which will be used along with the individual scale components to calculate the back-projection from 2D into 3D. Each SCI scale component will be assigned to the corresponding RCI coordinate so that colour may be obtained on performing the back-projection. The 3D point cloud obtained may have deviated from the original point cloud and the colour that has been registered onto the 3D point may not correspond perfectly to the features, but this is expected and explained above about the sources of error.

3.2 Extraction of colour using the DLT

In order to only extract colour information from 2D image, the DLT coefficients can be determined with the iterative least squares. Obtaining the DLT projection matrix simply requires the 3D points and their corresponding 2D image points for a solution (Abdel-Aziz and Karara, 1971). Notably in this method, the solution obtained will show a good match for the 3D points to the image. However, the DLT transformation coefficients will not allow the solution to revert back to the original 3D using the same parameters since the depth information of the 2D image is not always available in real cases.

To register colour to the 3D points, corresponding transformed points (2D) are assigned the colour information from the image. The point cloud data contains image coordinate (pixel) information via transforming with the DLT matrix; these coordinates will relate to the RCI coordinates. Colour information in the RCI can be assigned for each matching point pair.

4 RESULTS/ANALYSIS

4.1 Simulated Dataset

The simulated dataset presented, resembling a corner (Figure 2a), is used to create the SCI (Figure 2(b)) as well as the SRCI (simulated RCI) for testing purposes. Figure 2(a) has a line extending from the origin indicating the simulated camera direction for both the SCI and SRCI. As mentioned earlier in section 3.1, the SRCI created has a slight deviation from the SCI as shown in Figure 2(b).

ICP is the method of correspondence used to match the two similar images together, as seen in Figure 3(b). As shown, this ICP method used is able to obtain a good enough correspondence and these slight offsets will carry through when back-projected into 3D. Alternatively a different correspondence method may be used that may produce a closer and much more complete match.

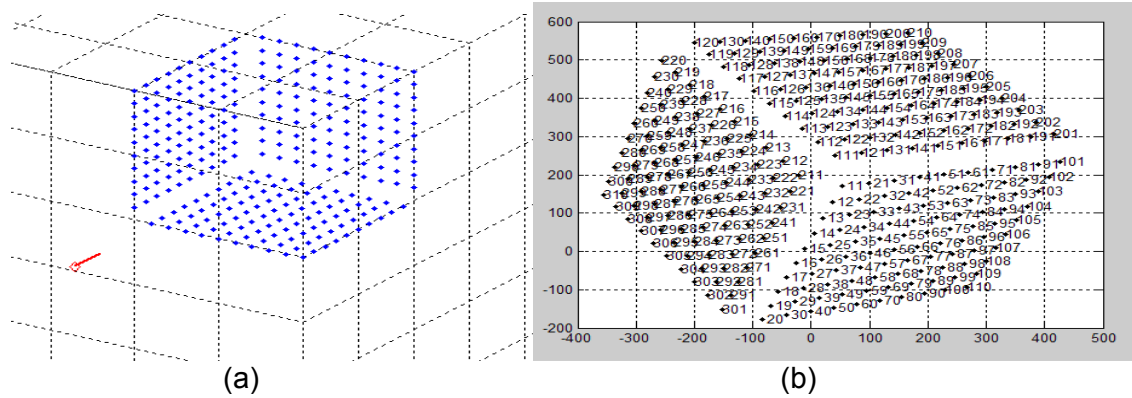


Figure 2: (a) A 3D representation of the simulated point cloud used. (b) The SRCI created from the simulated data, as an image in pixels.

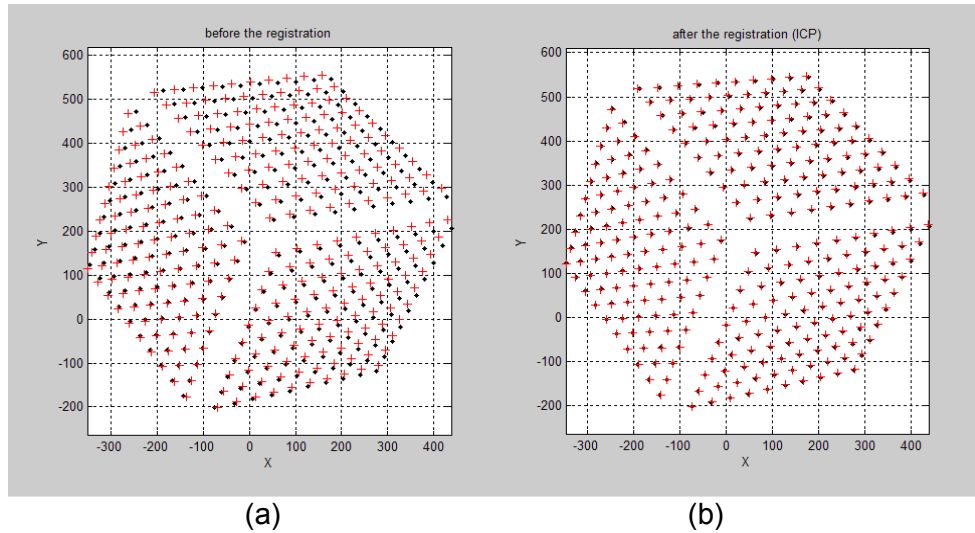


Figure 3: (a) Before performing the ICP registration (b) The updated image coordinates (dots) after the ICP.

After the ICP registration, Figure 4(a) shows the results of applying the individual scale components and the back-projection parameters to the SRCI. As expected, small deviations still exist between the SCI and SRCI, Table 1 shows the average overall error offset for the image between the SCI and the post-ICP SRCI, as well as in terms of its X, Y components. Also, the RMS error offset between the original 3D data and the back-projected data is calculated. The overall error of 1.68 pixels upon performing the back-projection will relate to 26.9cm overall error in the point cloud at a distance of approximately 60m. It is worth noting that in this simulation, the

camera resolution is not fixed like a standard camera and is a factor that will affect the result of the colour registration. Some limitations exist for this method when using non-simulated datasets in that the exterior orientation parameters have to be well defined in order to produce an SCI that represents the RCI well. Without good orientation parameters, especially translation, it would presumably affect the scale giving an erroneous solution as shown in Figure 3(b).

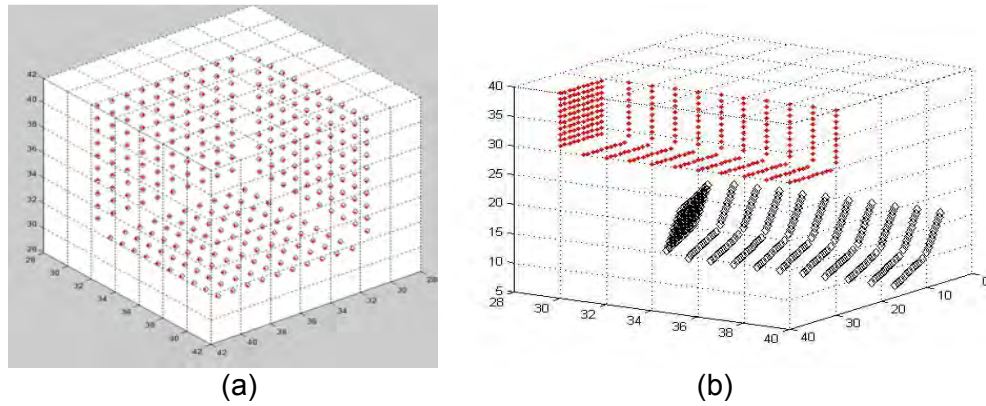


Figure 4: (a) The dots indicate the back-projected results, diamonds represent original (b) Erroneous data (diamonds) due to inaccurate parameters.

Table 1: The RMS error between the SCI and updated RCI

	RMS Total	RMS in X	RMS Y	
Image: (pixels)	1.68	1.65	1.71	
	RMS Total	RMS in X	RMS Y	RMS in Z
Point Cloud: (m)	0.0269	0.0246	0.0224	0.0325

4.2 Trial Scan Dataset

This dataset is acquired using the Leica Scanstation 3000, with roughly 80,000 points at the distance of 1.5m. The Figure 4(a), shows the image of the area with the corresponding point cloud data being mapped on using DLT. The targets used in the Figure 2(a) were a mixture of Leica HDS targets and printed black and white targets. Once the point cloud contains the image coordinate information, the RCI colour data for the corresponding point is assigned to give the result shown in Figure 4(b). The red asterisks in the image are the image target points after projection, to visually indicate the extent of the shifts due to back-projection. The DLT method projection has an error of 11.8 pixels, signifying that the colour misrepresentation in the point cloud is approximately 0.018m.

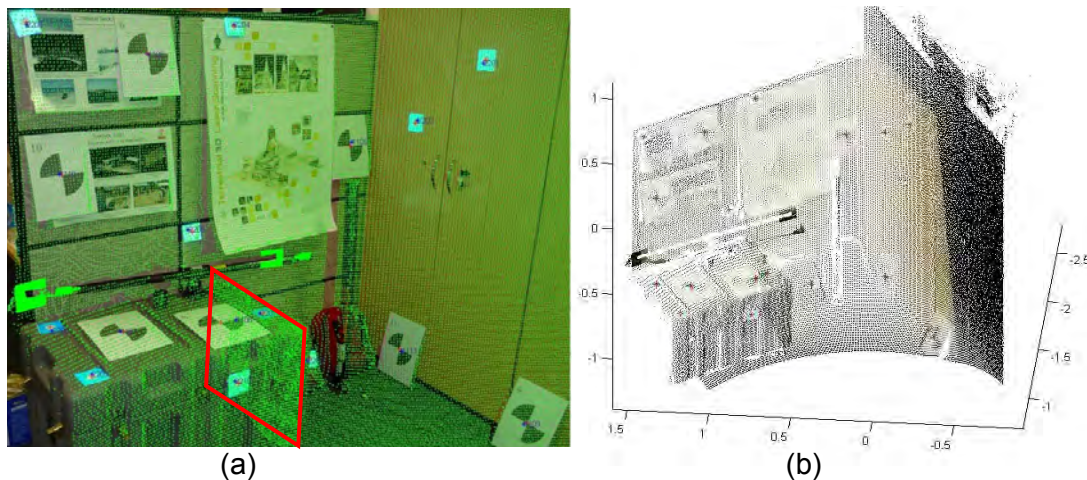


Figure 4: (a) Light-green dots on camera image represent point cloud data (b) 3D metric photorealistic point cloud data.

Table 2: RMS error between the original control points and the DLT control points.

	RMS Total	RMS in X	RMS Y
Image: (pixels)	11.8	16.5	2.5

Some apparent limitations are; erroneously stored colour data, outlined in red in Figure 4(a) where the region of the point cloud is not shown in the image, but present in the point cloud. A method suggested to avoid this occurrence is to apply some threshold parameters. To achieve good results, sufficient amount of well-located and defined control points, or a well-defined set of exterior orientation parameters for the projection matrix are needed.

5 CONCLUSION AND DISCUSSION:

Though photorealistic models can be obtained using two images and photogrammetry techniques, alternatively this method obtains a photorealistic metric model with a single image and point cloud data. As shown, it can be seen that the registration of 2D to 3D points using the DLT method achieves results that are satisfactory for photorealistic metric models. Using an SCI, RCI, ICP and DLT we can transform between 2D and 3D, whilst accounting for metric accuracy. However, there are limitations to the method at this stage, whereby this method can only be applied to small rotations and translations. Further investigation is necessary to improve the determination of the DLT coefficients, and examination into the SCI method and components are required to find a solution for its current limitations.

There are several possibilities that require further investigation to acquire the depth information for the estimation of the 3rd row of the DLT matrix, such as calculations using the 1st and 2nd row coefficients along with focal length and object scale relationship or defining the 3rd row via geometric means by finding rotation and translation elements. Factors that influence the accuracy of the registration also includes distance of images taken from scanner and pixel resolution of the images taken. In addition, we plan to implement a complete error analysis with a Gauss-Helmert model including the error model of 3D point clouds and possible 2D images after the separate calibration procedure.

6 ACKNOWLEDGEMENTS

This project was in part supported by the Australian Research Council (ARC) Linkage Infrastructure grant (LE0775672) and Prof M. Bennamoun and his team at the University of Western Australia are appreciated for their hardware support.

REFERENCES

- Abdel-Aziz, Y. I., and H. M. Karara. 1971. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry, *Proceedings of the Symposium on Close-Range Photogrammetry*. Falls Church, VA. American Society of Photogrammetry, pp. 1-18.
- Al-Manasir, K., and C. S. Fraser. 2006. Registration of terrestrial laser scanner data using imagery, *The Photogrammetric Record*, 21 (115), pp. 255-268.
- Bae, K.-H., and D. D. Lichti. 2008. A method for automated registration of unorganised point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63 (1), pp. 36-54.
- Besl, P. J., and N. D. McKay. 1992. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14 (2), pp. 239-256.
- Forkuo, E. K., and B. King. 2005. Automatic fusion of photogrammetric imagery and laser scanner point clouds, *Proceedings of the XXth ISPRS Congress*, Istanbul, Turkey, pp. 921-926.
- Hartley, R., and A. Zisserman. 2003. *Multiple view geometry in computer vision*, Cambridge University Press, 655 pages.

Jansa, J., N. Studnicka, G. Forkert, A. Haring, and H. Kager. 2004. Terrestrial laser scanning and photogrammetry - acquisition techniques, *Proceedings of the XX ISPRS congress, Istanbul, Turkey*. (<http://www.isprs.org/congresses/istanbul2004/comm5/papers/184.pdf> Accessed on August, 2011).

Kang, Z., S. Zlatanova, and B. Gorte. 2007. Automatic Registration of Terrestrial Scanning Data Based on Registered Imagery, *Proceedings of FIG Working Week*, ([http://www.gdmc.nl/publications/2007/Automatic Registration Terrestrial Scanning Data.pdf](http://www.gdmc.nl/publications/2007/Automatic%20Registration%20Terrestrial%20Scanning%20Data.pdf) Accessed on August, 2011).

Kern, F. 2001. Supplementing Laser Scanner Geometric Data with Photogrammetric Images for Modeling, *Proceeding of the XVIII. International CIPA Symposium. Surveying and Documentation of Historic Buildings - Monuments - Sites Traditional and Modern Methods*, pp. 454–461.

Leica-Geosystems. <http://www.leica-geosystems.com/> (Accessed August, 2011).

Luhmann, T., S. Robson, S. Kyle, and I. Hartley. 2006. *Close range photogrammetry; Principles, Methods and Applications*, Whittles Publishing, 510 pages.

McGlone, C., E. Mikhail, and J. Bethel. 2004. *Manual of Photogrammetry*, ASPRS (American Society of Photogrammetry and Remote Sensing), 1151 pages.

Rusinkiewicz S. and M. Levoy. 2001. Efficient variants of the ICP algorithm, *Proceedings of the third international conference on 3D Digital Imaging and Modeling (3DIM)*, (http://www.cs.princeton.edu/~smr/papers/fasticp/fasticp_paper.pdf (Accessed August, 2011).

Vosselman, G., and H. G. Maas. 2010. *Airborne and terrestrial laser scanning*, Whittles Publishing, 318 pages.

Yang, C., and G. Medioni. 1992. Object modelling by registration of multiple range images, *Image and Vision Computing*, 10 (3), pp. 145-155.

BRIEF BIOGRAPHY OF PRESENTER

Kwanthar Lim graduated from Curtin University with a degree in Bachelor of Surveying. After working as a research assistant, he has been a tutor on several Photogrammetry units for both undergraduate and postgraduate students at Curtin with Dr Kwang-Ho Bae. He proceeded to do his postgraduate in Masters of Surveying and Mapping with a focus in laser scanning. He has co-authored a paper with Christophe Weyer and Dr Kwang-Ho Bae.

Kwang-Ho Bae obtained a BSc in Physics from Korea University, Seoul, Korea and his BSc (Hons) and MSc in Physics from the University of Adelaide. He gained a PhD in Spatial Sciences from Curtin University of Technology in 2006. He has been a lecturer in Photogrammetry and Laser Scanning in the Department of Spatial Sciences at Curtin and now is with Fugro Spatial Solutions from September 2011. His research interests are Laser Scanning data analysis such as registration, segmentation, calibration and positional error analysis, Photogrammetry and 3D range cameras.

David Belton received the BS degree (1st Honours) in Computer Sciences and Mathematics from Curtin University of Technology in 2003 and completed the PhD degree in Spatial Sciences in 2008 at Curtin. In 2008, he took up a research associate position with the CRC for Spatial Information and Curtin University of Technology after submitting his PhD thesis. He also received the International Society for Photogrammetry and Remote Sensing (ISPRS) Prize for Best Papers by Young Authors at the ISPRS congress, Beijing, in 2008. He also received D. B. Johnston Award for Excellence in postgraduate studies in the spatial sciences area from the SSI and Curtin University of Technology in 2009. His research interests lie in 3D point cloud processing, error analysis, and optimisation techniques.

Compressing Images Using Contours

Gabriel Scarmana

University of Southern Queensland
Toowoomba, Australia
gabriel.scarmana@usq.edu.au

ABSTRACT

A method for the vectorisation of digital images into contour maps with subsequent conversion of the contours to a pixel format is presented. This method may offer an alternative spatial image compression which is computationally inexpensive and can be directly applied to compressing certain classes of grey-scale and/or colour (RGB, 24 bits) imagery of any size.

The feasibility of this study is based on research which shows that pixel based imagery can be sufficiently and accurately represented by their contour maps if a suitable contour model and scale selection method is used (Elder and Goldberg, 2001).

The compression process is based on filtering and eliminating those contours that may contain redundant information. Contours extracted from digital images may contain multiple redundant data (i.e. intersecting or nested contours), any of which might logically be used as a basis for discrimination or, on the other hand, used in the reconstruction of an original captured image.

As per current image compression techniques the proposed method does not require special hardware and, if combined with existing encoding schemes, it can be efficiently used for image transmission purposes due to its relatively modest storage requirements. Depending on the applications, and the amount of contour lines employed in the reconstruction of an image, the process allows for various levels of image accuracy while preserving visual integrity and reducing compression artifacts.

KEYWORDS: contouring, image compression, image reconstruction, image processing, image editing.

1 INTRODUCTION

Image compression schemes are based on the fact that any set of data can, and generally does, contain redundancies. In digital images, data redundancies can exist as repeated patterns and other forms of common pixel intensity information between multiple pixels of the image. The goal of image compression is to characterise these redundancies and code them to a new form that requires less data than the original (Weeks, 1996).

All image data compression schemes are basically twofold as they involve both a compression operation and an inverse decompression operation. The compression operation converts the original image into a compressed image data form. The decompression operation converts the compression image data back to its original uncompressed form (Weeks, 1996).

The measure of the amount of compression achieved in an image compression operation is obtained from dividing the data size of the original image by the data size of the compressed image. The result is called the compression ratio. The higher the compression ratio is, the smaller the compressed image has become. In all cases, the aim is to maximise compression ratios in

image compression operations while still meeting application requirements (Duperet, 2002). These requirements generally involve compressed image quality (for lossy compression schemes), time to compress and decompress the image, storage and the computational effort (Russ, 2007).

The type of compression scheme where the compressed data is decompressed back to its exact original form is called lossless data compression. It is devoid of losses, or degradation, to the data. The counterpart to lossless data compression is called lossy data compression. Lossy compression schemes introduce degradations to the data they compress (Gonzalez, Woods and Eddins, 2009). However, they do so in a way that is tolerable for the intended application. A lossy compression scheme which has become an industry standard is referred to as the Joint Photographic Expert Group (JPEG).

The JPEG image data compression standard handles grey-scale and colour images of varying resolution and size. It is intended to support many industries that need to transfer and archive images. This standard is commonly used in a lossy mode. However, there is also a lossless mode with reduced compression performance. The JPEG scheme is also adjustable. For instance, the amount of retained information can be changed producing variable compression ratios and inversely proportional decompressed image quality. Hence the JPEG algorithm can be fine-tuned to meet an application's requirements of compressed image data size and decompressed image quality (De Jong and van der Meer, 2004).

Other relatively new techniques for lossy image compression include the use of wavelet transforms (ECW-Enhanced Compressed Wavelets) and fractals. Compared to JPEG, these methods have not gained widespread acceptance for use on the Internet at the time of this writing. However, these techniques are very promising as they produce higher compression ratios and higher visual image quality than JPEG. A detailed explanation of lossless and lossy image compression schemes and techniques is beyond the scope of this work, and the reader is referred to Russ (2007), Gonzalez and Woods (2008) and Cunha, Zhou and Do (2006) for the theory and applications related to this important aspect of image processing.

For the purpose of this paper, the approach to solving the problem of image compression adopts techniques from the compression of contour maps. Compression here is achieved through the sequential filtering of contour data so as to remove redundant information which may have no or little effect in reconstructing an original digital image (Wang and Liu, 2006). Filtered contour data is then encoded to improve storage requirements, speed of transmission and reconstruction. As in the case of JPEG, the proposed compression technique is adjustable because, depending on the amount of discarded data, the storage capacity may vary at the expense of image quality. Two basic filtering techniques were tested in the proposed contour-based compression process, that is, string filters and contour exclusions.

String filters are designed to remove surplus vertices from contours strings. By way of a user defined tolerance, string vertices that are within the tolerance-based offset of two straight lines are removed. The tolerance value used normally depends on the data set and the job that the data is being used for. That is, the string filter removes vertices from contour lines that do not deviate by more than a specified offset tolerance from straight lines joining successive strings.

Contour exclusions can be achieved by comparing the centroid values of the area that these contours occupy to determine the shorter and longer distance to the centroids. These parameters may offer a useful constraint for deciding whether a contour is very similar to another and thereby could be removed. Also, for any group of at least three closely nested contours, it may possible to test for redundant contours by removing the middle contour and interpolating and/or diffusing between the other two. However, this process may have the disadvantage of smoothing and/or distorting particular high frequency information (sudden changes of pixel intensity values within an image) that may be necessary for a more accurate image reconstruction.

The statistical tests presented in the ensuing sections were carried out to determine the accuracy and integrity of images reconstructed from their contour representations. A comparison with the JPEG compression protocol was also conducted to evaluate the efficiency of the proposed contour compression technique in relation to storage requirements, accuracy and quality of visual details.

STANDARD DEFINITIONS

Digital image: A digital image is a discrete approximation of an image obtained by sampling points with discrete coordinates and quantising the values of each sample. It is formed by a finite number of sample elements equally spaced over a square grid with a rectangular shape. Each element is called a pixel and has an intensity value. The rows and columns of elements determine the spatial coordinates (x, y) of the pixel; and the intensity determines its grey-scale or colour value. In a grey-scale image, all pixels have shades of grey ranging from black to white. In the case of colour images, the intensity determines the colour of each pixel according to some colour model, such as RGB. See Figure 1 for an example of gray-scale continuous and digital images.

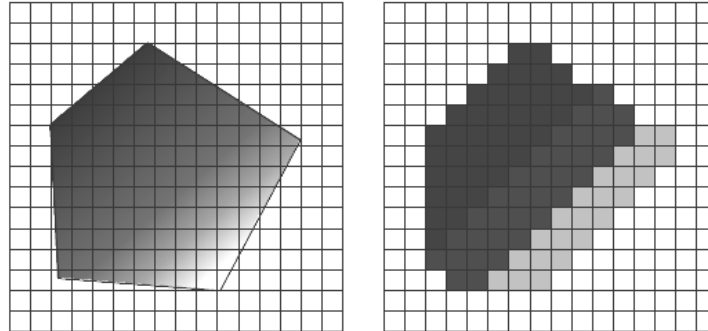


Fig. 1: Continuous image (left) and sampled digital image (right).

Pixel-based contour representations (Figure 2) require a lot of space because many pixels are needed; they are not scale independent since the number of pixels required changes with the size; and they are discrete. Pixel-based contours are generally not smooth unless numerous pixels are used. Given these downsides, a geometric representation or vectorisation provides a better way to represent contours, for example, with line segments defined by pairs of points. This representation does not have the problems associated with the discrete counterpart [3].

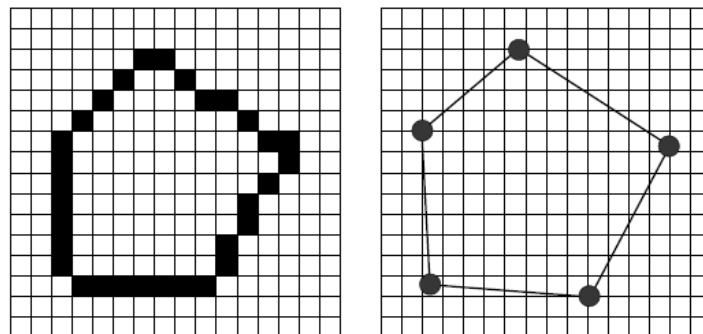


Figure 2: Pixel-based (left) and geometric-based (right) contour representations.

The definition of contour used in this work relates to the term used in cartography and surveying where a contour line joins points of equal elevation (height) above a given reference level (Watson, 1992). Note that this definition may differ from the one commonly used in image processing where contours relate to shapes, edges or lines and object boundaries (Elder and Goldberg, 2001).

CONTOUR DETECTION

The first step in the method is to generate contour lines for the gray-scale intensity values and assign coordinate values to pixels in the raster format image data. This is followed by interpolating between the pixels to find the coordinates of the points on the path of a contour with a

specific grey-scale intensity value. Hence, a digital image can be represented either as a three-dimensional surface with elevation z , or by using contours where each point (i.e. a vertex) of a given contour line corresponds to a position (x,y) with the same intensity value as the other points on the line.

Each pixel value can be assigned an x and y coordinate, which may for convenience be based on the coordinate of the centre of the pixel. Each pixel then has at least a pair of coordinates and a pixel grey-scale or intensity value. If two adjacent pixels have grey-scale values of 40 and 50 respectively, then a contour for grey-scale value 45 would lie between these two pixel centres. The coordinates in pixel space of a point on this contour can be estimated by interpolation, thus enabling the contours and/or the contour bounds to be estimated to sub-pixel accuracy.

For the class of images investigated in this work, the contour intervals or increments were selected based on the dynamic range of the image and/or on its histogram. Image histograms provide convenient, easy-to-read representations of the concentrations of pixels versus pixel intensity in an image. Using this graph it is possible to discern how much of the available dynamic range is used by an image (Russ, 2007). The dynamic range is the ratio of the highest (lightest) pixel intensity which a sensor records to the lowest (darkest) pixel intensity. The lightest pixel would correspond to the brightest intensity in an image; the darkest pixel to the deepest shadows.

In Figure 3 the pixel values of the .tif (tagged image file) image on the left, referred to as the *aerial view* (1000^2 pixels), fall between 15 and 215 (within a maximum range of 0 to 255) as shown on the right in Figure 3, with none in the other regions of the image. This indicates a relatively wide dynamic range of intensity values thus, in general, requiring a larger number of contours than an image with a narrow scale distribution (small dynamic range). The histogram in Figure 3 shows an example how many of the total number of pixels (3584 pixels) have an intensity value of 187 and the percentage (0.4%) they represent within the image of the *aerial view*.

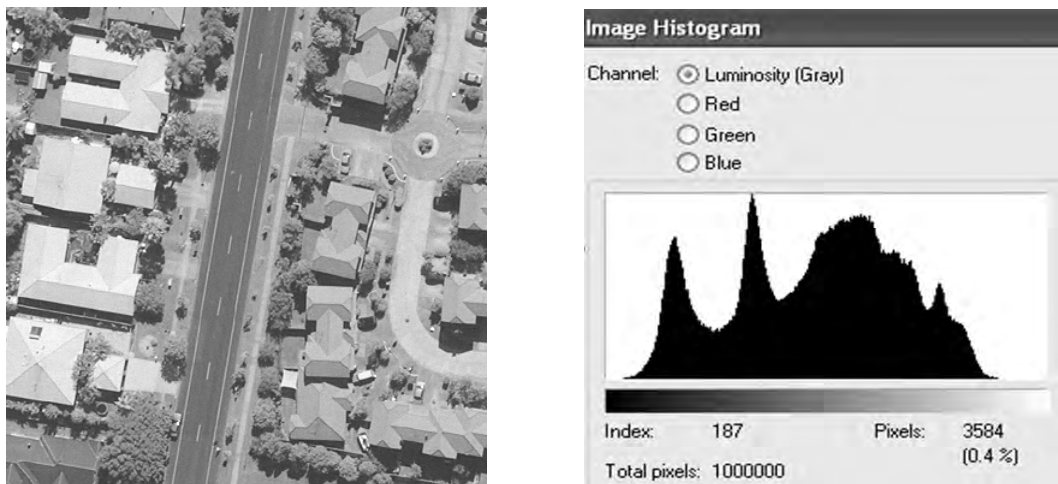


Figure 3: The histogram (right) shows that the distribution of pixel intensity values for the image (*aerial view*) on the left fall between approximately 15 and 215 within a range of 0-255.

Figure 4 illustrates a series of contour lines as produced for an array of uniformly spaced pixel intensity values. In this instance the contour increment was 4 units. The dots connecting the contour segments are referred to as vertices. There exist a number of methods and computer packages used for constructing contour lines from grid data points. The contours in Figure 4 were derived using a weighted average method of interpolation (Watson, 1992). Figure 4 (on the right) shows the contour lines for the *aerial view* in Figure 3 corresponding to the grey-scale intensity value of 100 only (full range being 0-255).

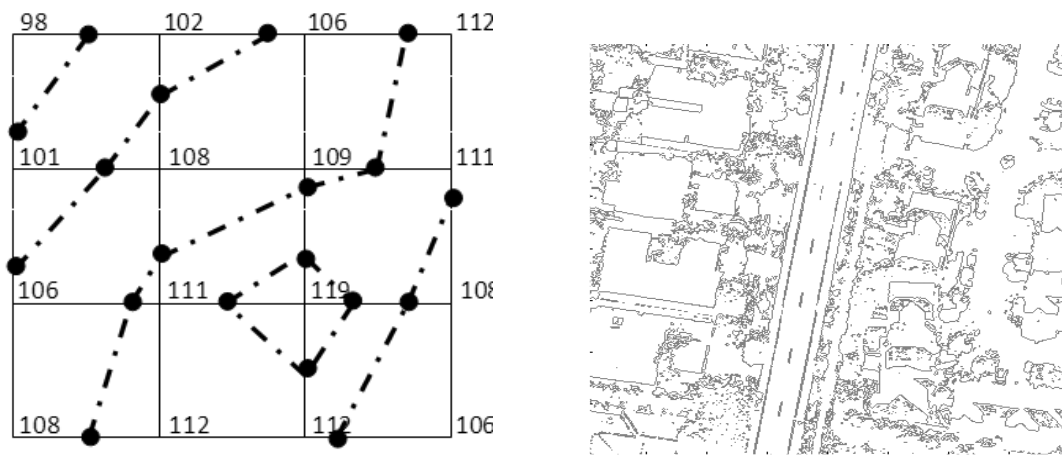


Figure 4 – Contours produced for a uniformly-spaced 4x4 grid of pixel intensity values (left) for contours incrementing in units of 4; the contour lines corresponding to the grey scale value of 100 in the *aerial* view of Figure 3 (left) are shown on the right.

IMAGE RECONSTRUCTION: FROM CONTOUR TO PIXELS

A number of tests were performed to investigate which contour increment would suffice to reconstruct images with a relatively large dynamic range such as aerial images, human faces, and landscapes. The tests were performed on a number of images by comparing the results of reconstructions via interpolating from respective contour representations with their original raster format.

Interpolating from contour lines to determine the value that would exist at the intersections of a regularly spaced grid is a classic problem in computational cartography and surface modeling. Traditional methods extend straight lines in a number of directions from the test point until these lines intersect the contour lines, the coordinates of these intersections are then used to interpolate the point of interest with a weighted average as described in Watson (1992) and Li and Heap (2006). These methods work efficiently if the contours are not nested.

However, using the above technique on nested contours usually results in an artifact appearing in the form of a cone. Since the outer contour line are longer than the inner ones, the averaging rule causes the surface between these two contour lines to droop, as if pulled down by gravity, in a scalloping or terraced effect. To avert these shortcomings, the method used in this research uses the filtered contours approach to extract bi-directional “cross-sections” in x and y at regular intervals.

To determine the pixel intensity values which would exist at the intersections of a regular grid using the randomly spaced nodes of these cross-sections, several interpolators may be used depending on the application and accuracy requirements. The interpolation technique adopted here is the Matlab function “griddata” (www.mathworks.com). This function fits a surface of the form $z = f(x,y)$ to the data in the (usually) non-uniformly spaced vectors x, y, z (i.e., the nodes of the cross sections). “griddata” interpolates this surface at specified x_i, y_i locations (a uniform grid) to produce the required z values using the Bi-harmonic Spline Interpolation method devised by Sandwell (1987). The surface always passes through the input data points.

Since the interpolation methods determine decimal values for the points at the grid intersections, the values are rounded to the nearest integer as this is the norm for representing pixels intensities within a digital image. In addition, these values should be restricted within the interval 0-255 for the case of standard (8 bits) gray scale imagery.

TESTS AND RESULTS

The first test was aimed at determining the degree of accuracy and memory storage requirements that can be expected when reproducing a digital image from its contour representation. The contours increments in this exercise varied between 2 and 10. The results are based on the *aerial* view in Figure 3. Table 1 shows the results of this test. The Root Mean Square

(R.M.S.) error (Spiegel and Stephens, 1999) refers to the error of the differences between the original *aerial view* and the same image reconstructed using the vertices of grey-scale contours.

The figures in Table 1 show a linear degradation of the R.M.S error as the contour interval is increased. The table also illustrates the amount of memory required to store the contour vertices needed for the reconstruction of the aerial view in Figure 3. As the contour interval increases, the memory requirement decreases.

Table 1: Accuracy and memory requirements needed to reconstruct the original *aerial view* using contour increments between 2 and 10.

Contour increments	Accuracy R.M.S.	Memory required (.txt)	Maximum Fluctuation value	Minimum Fluctuation value
2	3	0.50 Mb	+2	-2
4	9	0.33 Mb	+9	-11
6	14	0.20 Mb	+25	-23
8	18	0.17 Mb	+30	-32
10	20	0.13 Mb	+47	-55

By way of comparison, the .txt file needed to store the x and y coordinates of the vertices for reproducing the *aerial view* (for contour increment of 8 intensity values) used approximately the same amount of memory (0.13 Mb) of a JPEG compressed version of the same image with compression ratio 10:1, while producing a more accurate image with superior visual details (Figure 5). In this figure, the difference in visual quality between the aerial view reconstructed from contour data and the corresponding JPEG image is evident. Indeed, the blocking effects, typical of a lossy JPEG protocol, were almost completely eliminated.

In addition, upon subtracting the JPEG version of the aerial view in Figure 5 from the original aerial view in Figure 3 generated grey-scale differences fluctuating between +39 (maximum) and -44 (minimum) with an R.M.S error equal to +/-25. By contrast the contour approach produced grey-scale fluctuation values ranging between +30 and -32 respectively with an R.M.S. error equal to +/-18.

For easy extraction, processing and transmission over a digital link the contours data was encoded and saved in a .txt file. The .txt file contains a two-row matrix where each contour line defined in the matrix begins with a column that defines the value of the contour line and the number of (x,y) vertices in the contour line.



Figure 5: A JPEG version (right, section only) of the *aerial view* (1000^2) for a compression ratio of 10:1 and the same image reconstructed using the vertices of contour lines with increments equal to 8 (left, same section).

The remaining columns contain the data for the (x,y) pairs. The x, y values are stored to the first decimal place and separated by one tab space. In the coding shown below 112 is the grey-scale intensity of a selected contour and the contour line is formed by 7 vertices (Gonzalez et al., 2009). The x coordinate of the vertices are in the top row (i.e. 17.2 17.1...16.7) whereas the corresponding y coordinates are placed in the bottom row (i.e. 100 99...95.2)

```

112 17 2 17 1 17 17 16 9 16 7 16 7
7   100 99 98 97 9 97 96 95 2

```

A second test was devised to establish the effect that filtering the contour data (prior to the reconstruction phase) has on the memory required to store such data. The same *aerial view* in Figure 3 was used in this test. To generate the grey-scale intensity contours, an increment of 5 grey-scale intensity values was selected for a range of 15-215 pixel intensity values; this range is a multiple of the increment value and includes all the grey scale variations within the original image. The intervals ranging from 15 to 215 thus produced 40 grey-scale contour line values.

The resulting contours were then filtered in succession using the two filters introduced in Section 1. The first filter applied aimed at eliminating nested contours. Within any group of at least three closely nested contours, the middle contour was eliminated. Upon completing this elimination phase, the contour lines were further filtered using a string filter. Figure 7 illustrates the principle behind a string filter. Table 2 shows the results of the reconstruction upon varying the tolerance tube (shown in Figure 7) from 0.2 to 1.0 unit of intensity. The table also indicates the associated levels of memory requirements for the .txt files needed to store the contour data.

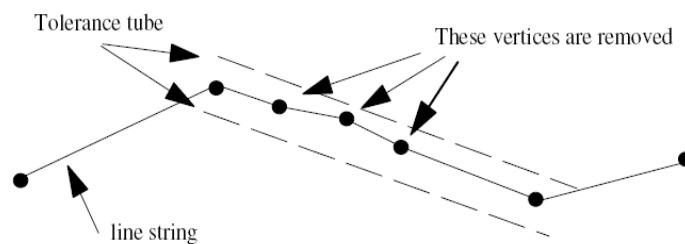


Figure 7: A string filter option is used to remove surplus vertices from contours lines.

By comparing Table 2 and Table 1 it may be concluded that a further reduction in memory requirements is obtained by filtering and eliminating contour data which may be considered as redundant and that for lower values of the tolerance tube width there are no major effects on the accuracy of the final reconstruction. For example, adding a string filter to the 5 grey-scale values contour, where the filter can have a tolerance tube size of up to 0.4, the resulting R.M.S. error (error 12, Table 2) is still lower than the R.M.S. error of the 6 grey-scale values contour (14, Table 1). The tables show that for contour increments of 5 grey-scale values, the memory required to store the *aerial view* as a .txt file may be reduced but at the expense of the accuracy of the image reconstruction process.

Table 2: The effect of increasing the tolerance tube on the accuracy of reconstruction and the memory required to store the contour data (in Mb) as a .txt file for the *aerial view* (for a contour increment equal to 5).

Tolerance tube	0.2	0.4	0.6	0.8	1.0
Accuracy R.M.S. error	10	12	18	25	36
Mem. required (.txt file) in Mb	0.30	0.17	0.15	0.11	0.10

COLOUR IMAGES

When processing colour images, the same contour and vectorisation concepts above hold true. However, instead of a single grey-scale intensity value, colour digital images have pixels that are generally quantised using three components (i.e. Red, Green and Blue). In general, all image processing operations can be extended to process colour images simply by applying them to each colour component.

Each of the three RGB components are processed, encoded and reconstructed separately as if they were three different grey-scale images. The results of the three reconstructions are then merged or fused to recreate the original colour image. In terms of image quality (R.M.S.) the results from the contouring approach, for the same classes of images (i.e. faces, landscapes and aerial views), were similar to those obtained from the grey-scale imagery described above.

On the other hand, the memory requirements to store/transmit contour data were also comparable to the memory needed to store colour imagery in JPEG format. The contour approach also provided more accurate results with no visible artefacts or blocking effects. However, the reconstruction process required to regenerate the colour image after transmission on a digital link required 2.5 times more time than reproducing the same image in JPEG format. In this context, additional research may be required to increase the speed of the image reconstruction from contour data when dealing, for example, with real time applications.

CONCLUSIONS

Notable findings encountered in this study are:

- Generating contour data from digital images involves assigning coordinates values to pixels in the raster format image data, and interpolating between the pixels to find the coordinates of points in the path of a contour having a given intensity value. This enables the contour bound to be found to sub-pixel accuracy if required.
- The conversion of certain classes of digital images into contour maps may be used to compress and reconstruct images in pixel format that are more accurate and with improved visual details than JPEG compressed versions of the same image, while requiring similar memory space for storage and speed of transmission over digital links.
- For the class of images investigated in this work, the contour approach to image compression requires contour data to be filtered and discriminated from the reconstruction process.
- Spline interpolation was used to reconstruct digital images from their contour representations. The process involves determining the pixel intensity value which would exist at the intersections of a regular grid using randomly spaced nodes of bidirectional cross-sections.
- The contour approach or vectorisation of digital images for the purpose of image compression can be applied to both grey-scale images and colour images. In the case of colour images each of the three channels (Red, Green and Blue) is processed and encoded separately. The results of the reconstruction for each colour component are then merged or fused to recreate the original colour image.
- Refinements to the proposed method are being undertaken to increase the accuracy achievable for a variety of scenes and dynamic ranges (including bi-tonal imagery).
- More research is required to assess the accuracy of the enhancement process in the presence of added random noise and/or video imagery.
- To compete with standard compression protocols for image compression additional study is required to speed up the process of reconstructing an image from its contour representations.

REFERENCES

- Bonham-Carter G. F. (2006) *Geographic Information Systems for Geoscientists: Modeling with GIS*. Vol. 13 of *Computer Methods in the Geosciences*. Elsevier, Butterworth-Heinemann. 398 pages.
- De Jong S. M. and van der Meer F.D. (2004) "Remote Sensing Image Analysis Including the Spatial Domain", Kluwer Academic Publishers, pp. 360.
- Duperet, A., (2002) "Image Improvements" Chapter 2, *Digital Photogrammetry*. In Michel Kasser and Yves Egels editors, Taylor & Francis Publishers. p.p. 351: 79-100.
- Elder J. H. and Goldberg R. M. (2001) "Image Editing in the Contour Domain". *IEEE Transactions on Pattern Analysis and Machine*. Vol 23 No. 3.
- Gonzalez R. C. and Woods R. E. (2008) *Digital image processing*. Prentice Hall, 954 pages.
- Smith M. Goodchild F. and Longley P.A. (2009) "Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools. 3rd Edition. The Winchelsea Press. pp. 560.
- Spiegel M. R., Stephens L. (1996) "Theory and Problems of Statistics, Schaum's Outline Series, McGraw-Hill Book Company. pp. 556.
- Wang S., Ge F. and Liu T. (2006) "Evaluating edge detection through boundary detection," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 1–15.
- Watson D. F. (1992) "Contouring: A Guide to the Analysis and Display of Spatial Data". Pergamon Press. New York. 321 pages.
- Weeks, A. R. (1996) "Fundamentals of Electronic Image Processing. SPIE Optical Engineering Press, Bellingham, Washington. 570 pages.
- Cunha A. L., Zhou J. and Do M. N. (2006) "The Non-subsampled Contourlet Transform: Theory, Design, and Applications" *IEEE Transactions on Image Processing*, VOL. 15, NO. 10, October.
- Elder H. J., Goldberg M. R. (2001) "Image Editing in the Contour Domain". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23 Issue 3, March 2001. IEEE Computer Society Washington, DC, US.
- Russ C. J., (2007) "The Image Processing Handbook", Published by CRC Press, ISBN 0849372542, 9780849372544, 817 pages.
- Gonzalez C. R. , Woods R. E. and Eddins S. L. (2009) "Digital Image Processing Using MATLAB" 2nd edition, Gatesmark Publishing. 827 pages.
- Li J. and Heap D. A. (2008). "A review of Spatial Interpolation Methods for Environmental Scientists". Record 2008/23. Geocat No. 68229. Geosciences Australia. Australian Government.
- Sandwell, D. T. 1987), "Bi-harmonic Spline Interpolation of GEOS-3 and SEASAT Altimeter Data", *Geophysical Research Letters*, 14, 2, 139–142.

Standardizing Australian Semi-Urban land cover mapping through FAO Land Cover Classification System

Kithsiri Perera^{1*}, David Moore², Armando Apan¹, Kevin McDougall¹

¹Faculty of Engineering and Surveying and Australian Centre for Sustainable Catchments,
University of Southern Queensland, West Street, Toowoomba 4350 QLD Australia

²Terranean Mapping, PO Box.729, Fortitude Valley, 4006 QLD Australia
perera@usq.edu.au, david.moore@terranean.com.au, apana@usq.edu.au,
kevin.mcdougall@usq.edu.au

ABSTRACT:

The application of satellite images for land cover mapping in Australia has achieved formidable developments in recent years as a result of increased data availability and improvements in software/hardware. However, the lack of standardization in land cover products creates considerable difficulty when integrating these land cover maps. To address this issue, this study investigated FAO's land cover classification system (LCCS) to map the Brisbane and Gold Coast region, as representatives of the Australian urban landscape. The surface diversification of this region in Southeast Queensland offers a reasonable land cover profile that could be used to explain most of populated Australian coastal regions.

SPOT 10m satellite data supported by extensive field investigations and 2.5m very high resolution SPOT data were utilized for the classification. The SLATS 2001-2003 (State-wide Land cover and Trees Study) data were also used in identifying training sites and for preparing legend categories. Initially, SPOT 10m Image was classified into three levels (dichotomous phase) of the FAO LCCS and then spectral features, field investigations, and SLATS data were used to generate 4th level (Modular-Hierarchical phase) land cover categories.

Results generated a very satisfactory land cover map with 15 sub-categories at 10m resolution compare to existing land cover products of SEQ. The hierarchical FAO classification approach used in this study produced a land cover map which can be used with national and global level land cover products based on FAO LCCS. Future research steps are investigating mapping rural and arid regions of Australia to combine into the national data set.

KEYWORDS: FAO LCCS, urban land cover, SPOT, *a priori* classification, Dichotomous approach, Southeast Queensland

1. INTRODUCTION

The urban population of Australia has remained a very high percentage throughout the history of the country. According to urban and non-urban population data, Statistical districts with a population greater than 100,000 people are categorised as "Urban". At 30 June 2004, three quarters of the Australian population (15.1 million people) lived in these urban areas. Some 12.8 million urban residents lived in Australia's eight capital cities. The urban population increased by 192,100 people in 2003-04 while non-urban Australia increased only by 46,500 people (Year Book Australia, 2007). At 30 June 2010, the estimated resident population of Queensland, where the present study focuses, was 4.51 million, an increase of 89,100 people since June 2009. From June 2005 to June 2010, Queensland had an average annual growth

rate of 2.5%, which was the second fastest of all states and territories, behind Western Australia (2.6%) (Australian Bureau of Statistics, 2009-10).

These increases of population have caused a direct impact on land management strategies. Hence, systematic exploration of the land resources is becoming an integral part of any environmental plan. In general terms, “Land” is a part of the earth, or the ground, not covered by water. According to some definitions in law, “Land” is described as a three dimensional space consisting of land and space below and above it (Butt, 2001). The FAO document defines the land according to its contribution to productivity. The main resource controlling primary productivity for terrestrial ecosystems can be defined in terms of land: the area of land available, land quality and the soil moisture characteristics (Di Gregorio and Jansen, 2000). The land is further characterised by its “Land Cover” and “Land Use”. The Australian Institute under the Natural Heritage Trust, the National Land and Water Resources Audit, agrees with the FAO’s definition of the land cover, which is describes as the observed biophysical cover on the earth’s surface including vegetation and manmade surfaces (Di Gregorio, 2005). Further, the National Land and Water Resources Audit defines Land Use as the purpose to which the land cover is committed (National Land and Water Resources Audit, 2007).

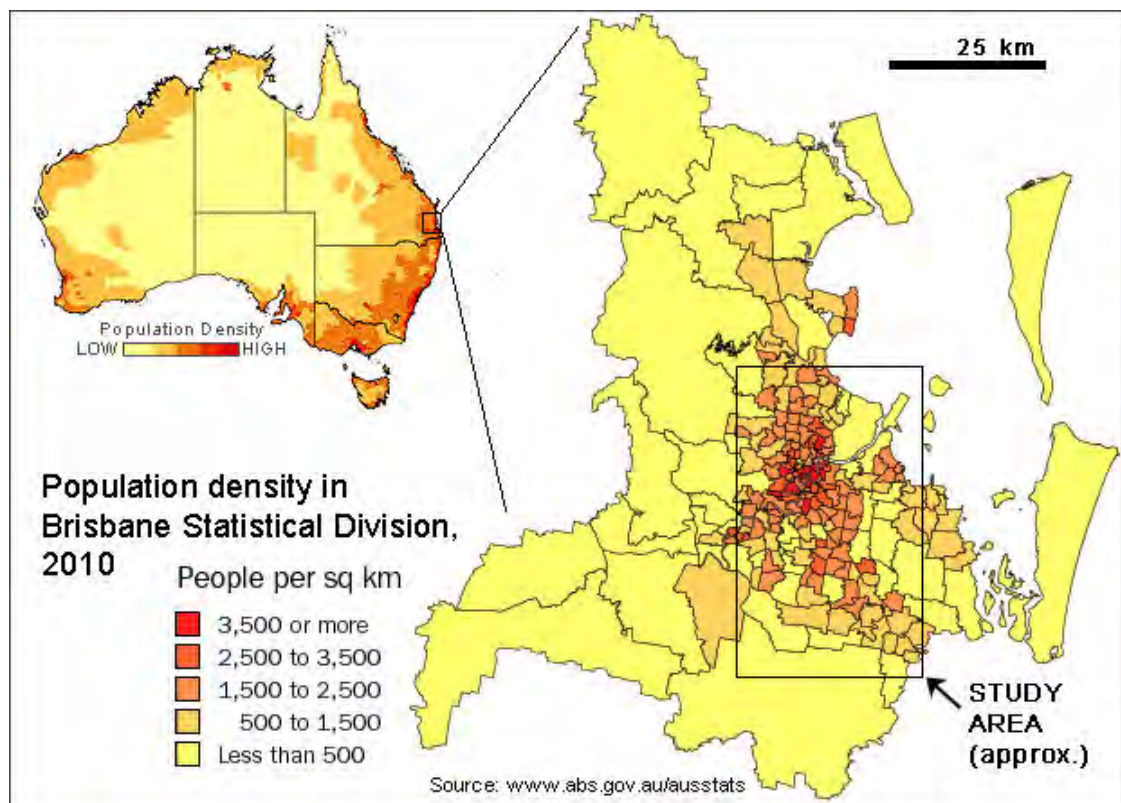


Figure 1. Population distribution in the area of interest.

Within last 50 years, the gross value of Australian agricultural sector expanded from \$4.5 billion in 1960/61 to \$46.5 billion in 2007/08 (Australian Bureau of Statistics, 2008). Due to its massive scale of agriculture and mining activities, present day Australia faces a new challenge to maintain its environment (World Resources Institute, 2006). Producing reliable and detail land cover information is vital in this background. A standardised and up-to-date land cover dataset is required to assess the condition of the natural resource base, for land use planning, to facilitate modelling of water quality, to assess soil erosion, and to facilitate the assessment and modelling of soil health and the sustainable production of food and fibre

(DAFF, 2007). Data generation must be conducted to satisfy the logic of standard land cover classification systems in order to be able to compare multi-temporal inter-state and international data. To address these issues, the Land Cover Classification System (LCCS) adopted by the FAO could be used as the standard to build a local land cover classification system for Australia.

The study presented in this paper investigates the “Urban Land Cover” aspect of the FAO LCCS. This will be complimentary to “Rural Land Cover” mapping, with the combination of both creating the potential to build state and national level land cover products. The study area of this research is south-east Queensland (Figure 1) which includes 66% of the state population and is considered to be one of the most important regions with high population in the country (Australian Bureau of Statistics, 2009-10). The term “semi-urban” in this study represents the mixed-land cover character of the Brisbane region, where man-made structures are intermixed with woody vegetation and grasslands.

2. CONSTRUCTING THE CLASSIFICATION SCHEME

Earth surface mapping developed rapidly with the introduction of earth observation satellites in 1972. Land cover and land use maps at various scales were generated to address specific needs of local areas, but none of the classification schemes became internationally recognized or standardized. This led to the FAO and UNEP gathering in 1993 to establish a land cover classification system to match the wider spectrum of global land cover types and by 2000 the FAO LCCS became fully operational. The Land cover classification system (LCCS) adopted can be considered as a truly international approach with logical definitions that apply to all land cover types of the world (Di Gregorio and Jansen, 2000).

2.1 Basics considered in FAO LCCS

The FAO LCCS system is considered as the only land cover classification approach available today which can be applied to any region of the world regardless of the economic conditions and data source. The FAO method is an “*a priori*” classification system, as it defines all the classes before the classification is conducted. The advantage of this approach is the possibility of maintaining the standardisation of classes. For this purpose, LCCS developed pre-defined classification criteria, or *classifiers*, to identify each class, instead of identifying the class itself. This concept is based on the idea that a land cover class can be defined without considering its location or its type, using a set of pre-selected classifiers. Therefore, when the user requires a large number of classes, a large number of classifiers are required.

To make the classification easier, the FAO system uses a *dichotomous* (divide into sub categories) approach in hierarchical levels and uses eight classifiers to group all land cover types at the third level. In other words, any location on the earth surface can be categorized into one of the eight classes without creating a conflict. Up to this third level, FAO used the presence of vegetation, edaphic (plant conditions generated by soil and not by climate), and artificiality of land cover for classification. Additionally, the third level of FAO classification can be considered as a concept based on visual classification, which uses the directly visible and knowledge based components of the ground.

In practice, a further breakdown of the third level eight classes must be conducted to obtain a detailed level of land cover classes. For that purpose, FAO uses a *hierarchical approach*, or the *Modular-Hierarchical Phase*, to build additional classifiers, but strictly within one of eight classes identified in the third level of the dichotomous phase. In this 4th phase, the system uses a set of pre-defined pure land cover classifiers, different from the eight classes in the dichotomous phase, as presented in Table 1.

Table 1. Dichotomous approach to build primary classes in FAO LCCS

First level	Second level	Third level
A. Primarily vegetated	A1. Terrestrial	A11. Cultivated and managed terrestrial areas
		A12. Natural and semi-natural terrestrial vegetation
	A2. Aquatic or regularly flooded	A23. Cultivated aquatic or regularly flooded
		A24. Natural and semi-natural aquatic or regularly flooded
B. Primarily non-vegetated	B1. Terrestrial	B15. Artificial Surfaces and Associated Areas
		B16. Bare Areas
	B2. Aquatic or regularly flooded	B27. Artificial water bodies, snow and ice
		B28. Natural water bodies, snow and ice

The pure land cover classifiers are defined by **Environmental Attributes** (e.g., climate, soil, and etc) or by **Specific Technical Attributes** (specific details like crop type and soil type) (Africover, 2008). In both cases, the user gets the freedom to add to these classifiers their own research interests, scale of the classification, and the physical and climatological conditions of the field. The FAO LCCS document presents a large number of classifiers to use in this level and the user can use only a selected set from the list to match with the scope of their own mapping project.

2.2 Australian land cover and its recent changes

The Australian flora and fauna is a composite of Gondwanan elements, and has evolutionary lines shared with South America. About 80% of the flora of Australia is endemic to the country and most of the species are extremely restricted in geographic and climatic range. For example, 53% of the approximately 800 species of eucalypts have climatic ranges spanning less than 3°C of the mean annual temperature (Hughes, 2005). Also, about 23% have adapted to less than 20% of mean annual rainfall changes (Pittock, 2003).

The millions of years old unique Australian landscape faced a dramatic change within last two centuries with the arrival of European settlers. The native vegetation cover or plants present in Australia before European settlement has declined to 87% of the country (State of Environment, 2006). Most of the native forest change has occurred through clearing of forests and woodlands, which originally covered 54% of the country and now covers only 42%. Although rainforests cover only about 0.3% of Australia, they contain about half of all Australian plant families and about a third of Australia's mammal and bird species (Department of the Environment, 2008). The rapid growth of urban feature expanded along the eastern coast, specifically from Brisbane to Melbourne, threatening even these limited rainforests. Population concentrations demand more energy and place more strain on already scarce natural resources, such as water and energy, causing a direct impact of overall land cover.

The restricted geographic and climatic range of the Australian plants means that the country's natural land cover is very vulnerable to the effects of climate change. Global warming is also may be affecting these flora (and fauna), since the largely flat Australian geography offers limited space to form dynamic differences in climate. Accurate and detail land cover mapping products are hence vital for any study related to these resource stress issues.

2.3 Applicability of FAO LCCS system in Australian terrain

Australian land cover is greatly influence by climate rather than its near-flat terrain with 99% of its land area below 1000m (Hughes 2003, 2005). Figure 2 compares the annual rainfall and topography of the country, which shows heavy rainfall along the east coastal areas of Queensland; this rainfall supported the development of nearly all urban centres of the state. Within Queensland, the central region receives extremely low rainfall (for example, at Birdsville, the mean annual rainfall is less than 200mm), while the northeast coast receives heavy monsoon rains (at Innisfail, the mean annual is over 3200mm) (see locations on Figure 2). Vegetation types throughout the state have adapted to these climatic variations. When classifying land cover of Australia, the *a priori* classification approach of FAO LCCS, provides a logical approach to separate land cover types. It helps to ignore differences in the land surface of Australia at the initial three levels of the classification (see Table 1). However, for the construction of the 4th level of the classification system, regional environmental features and field information must be considered.

When building the land cover map through these four levels of FAO LCCS, the near-flat terrain of Australia requires a focus on climate and soil characters rather than topology. The other elements to consider for the classification are spectral characteristics and the resolution of original image data, final mapping resolution and the quality of supporting data (including ground truth data). In this study we used SPOT 10m satellite data and a set of GIS data for the mapping. Also extensive ground surveys and SPOT 2.5m color composite images were used to build the classifiers.

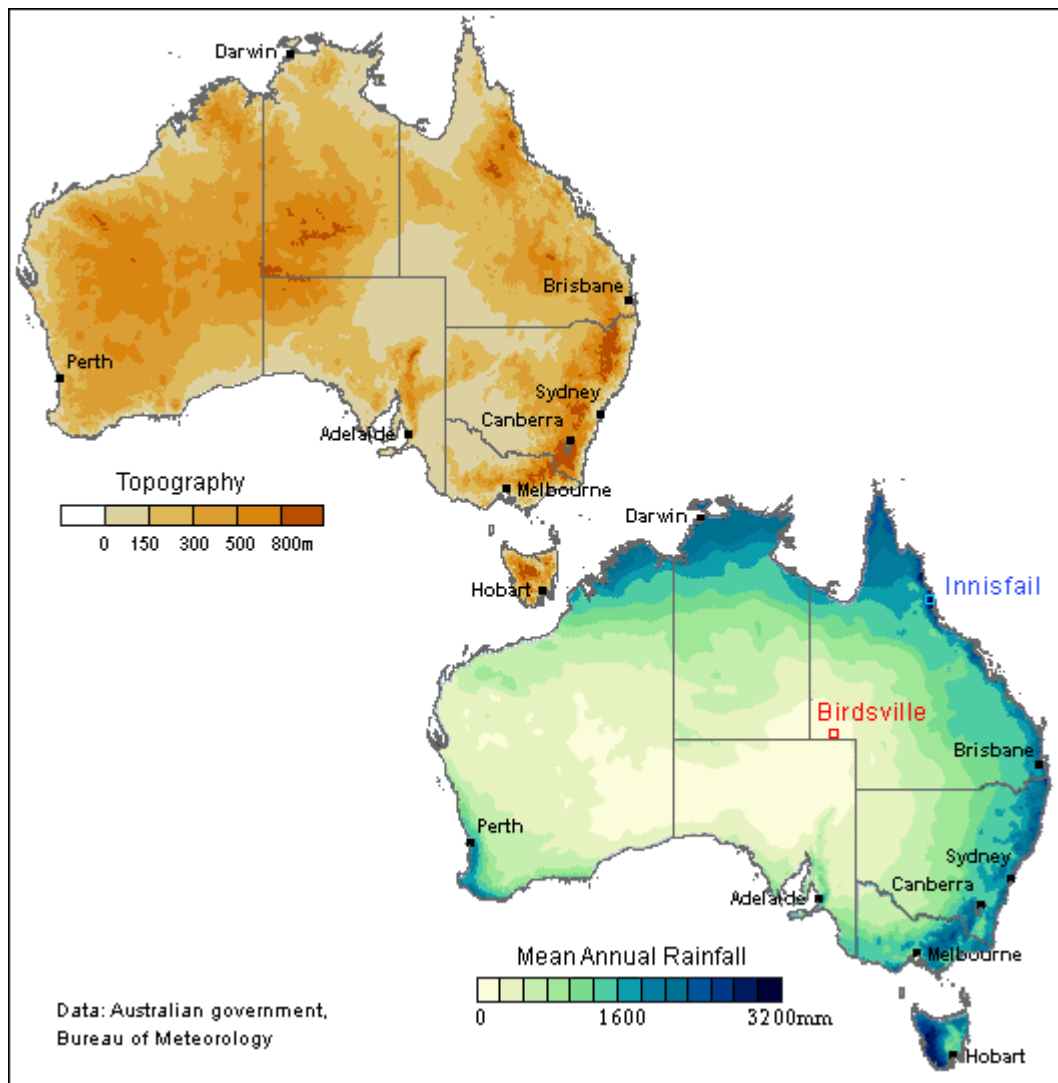


Figure 2. Elevation and annual rainfall of Australia (Climate statistics for Australian locations, 2011).

3. THE CASE STUDY

The land cover of Queensland varies from semi-desert barren lands and huge farm lands in the vast hinterland to some of Australia's largest remnant tropical rainforests including a world heritage site (Department of the Environment, 2008) and urban environments in east coast. Mapping the land cover characteristics of region associate with urban features is the prime object of this study. However, the mapped area contained a number of other common land cover types that represent significantly different land cover characteristics of the state.

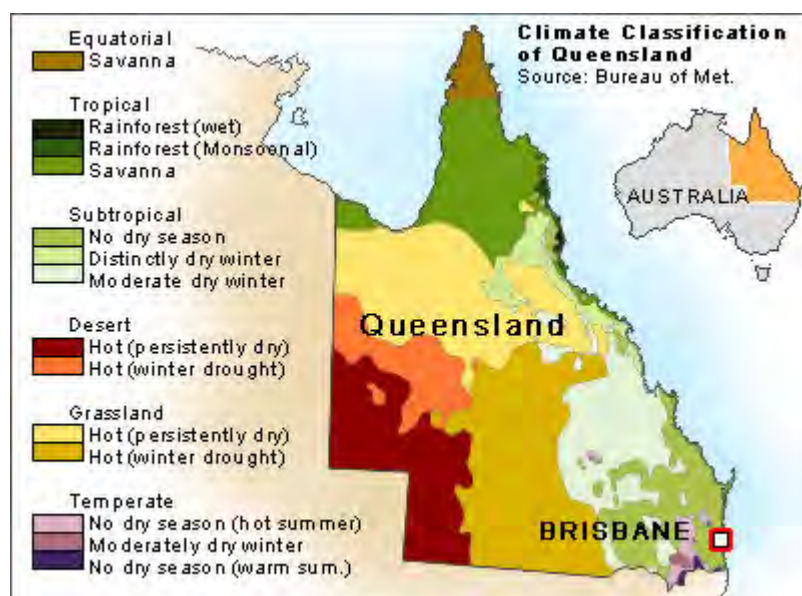


Figure 3. Locations of study areas on QLD Climatic Zone map (Bom, 2011)

The selected location of the study area falls within the subtropical climatic zone (Figures 1 and 3). The land cover classification of this region includes highly contrasting spectral characteristics making an approach developed specifically for this area suitable for application to other urban regions of Australia with appropriate modifications. Table 2 presents some of the basic information related to the study area. The Brisbane region (the study area), located in Southeast Queensland (SEQ), is the main urban cluster of the state. Its sub-tropical climate has warmer and wetter conditions with no clear dry season (Michael et al, 2005). The Brisbane region has very high population density figures compared to the national figures. For this study, we extracted a 1000 km² section from the full satellite scene.

Table 2. Main features of study areas.

Element	Brisbane City
Selected area	1000 km ²
Mean annual rainfall	1149.1mm
Mean annual maximum temperature	25.5°C (at regional office)
Climatic zone (based on Köppen)	Subtropical - No dry season
Elevation	0 – 280 m
Population density	115 (SEQ) (2004)
Main land cover features	Urban and settlements, woodlands

4. DATA AND DATA PROCESSING

4.1 Used data

Following data types were utilized in the study (Table 03).

Table 3. Used data in the study

Data type	Data set identifier	Date
SPOT 10m	390/405	06072006
SPOT 2.5m	sp5col2p5_SG5614.ecw sp5col2p5_SG5615.ecw	20052006 20052007
High resolution images in Google Earth		Latest
Statewide Land cover and Trees Study, QLD	SLATS	2001-2003
Field Survey		Apr 2008

4.2 Building the classification for study areas

Micro Image TNT software package (TNTmips 2008:74) was used for all aspects of the image processing. The construction of the first three levels of the classification was completed by strictly following the FAO LCCS structure. For these initial three levels, spectral characteristics of SPOT images and vegetation index image were used extensively. We also used different levels of spectral information to isolate broad classes at each level of the LCCS. A new set of training sites were selected for each level to perform the next level classification. Those training sites were selected with the help of 2.5m SPOT images, field investigations, different image indexes of SPOT, and general knowledge of the study area. Under the dichotomous approach (Table 1) of FAO LCCS, the accuracy of each initial level is permanently affected by the accuracy of following levels of the classification.

4.2.1 Classification Level I: A supervised classification to isolate non-vegetated lands was conducted through careful selection of training sites from 100% non-vegetated areas. Spectral values of each SPOT band and NDVI image together with 2.5m SPOT images were used to identify these training sites precisely. All other areas under different levels of vegetation (from vegetated area to a mix of bare ground and grass) were classified into vegetated areas.

4.2.2 Classification Level II: The re-classification was carried out with two classes of level I to generate four classes. After observing spectral characteristics and NDVI values, image classification was conducted through selecting training sites using the 2.5m and 10m SPOT images. Only 3 classes out of four were found, with the class A2 ("aquatic or regularly flooded areas under primarily vegetated category") (Figure 3) not found in the Brisbane region.

4.2.3 Classification Level III: At this level, FAO LCCS has 8 sub classes to represent all land surface features on the earth. The availability of the area under each class in the study area is directly dependant on the local features of land cover of the respective area. A clear example is: in a remote desert region with no human settlements or any vegetation, the classification may comprise of only one class (B16, A6: Loose and Shifting Sand) from these 8 classes. The Brisbane region has a predominantly sub-tropical climate and has substantial amounts of vegetated lands. However, due to its urban feature, the region contains a large number of sub-categories of artificial surfaces. We found six classes out of the original eight classes at this level (see Level III in Figure 3).

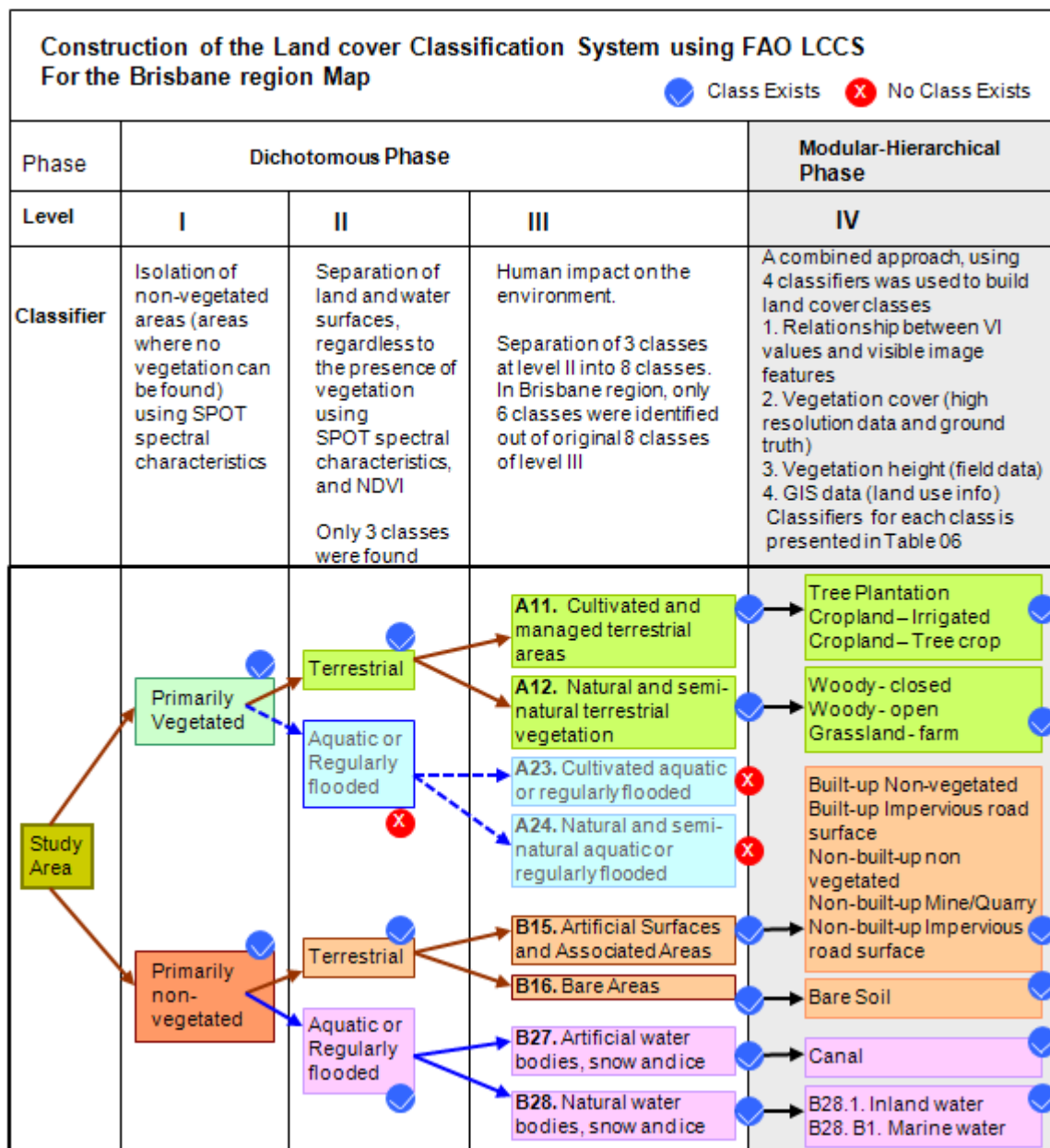


Figure 4. Building the classification scheme according to the FAO LCCS.

4.2.4 Classification Level IV: The 4th level of the classification is the challenging phase of the land cover mapping. For this level, the classification process must identify classes closer to real world land cover with clearly demarcated boundaries. As an example, even after extensive studies, the LCC for Tasmania conducted in 2006 had 14 classes at local level, but one of them, “seabird rookery complex” found no matching class name in FAO LCCS to be re-named (Atyeo and Thackway 2006). Fundamentally, the 4th level or local level class generation has to be conducted through applying more detail “classifiers” (Di Gregorio, 2005), as FAO LCCS requires. In this study we used very high resolution 2.5m satellite images and ground survey information to build classifiers for the 4th level. Additionally, spectral characteristics of SPOT 10m images played a strong role in the classification process. Figure 4 shows the simplified flow of this process which presents all four levels with regard to the Brisbane map. All classifiers used to generate each class in level IV for Brisbane map are presented in Table 4.

Table 4. Land cover classes of Brisbane area under FAO LCCS.

Class name on the map	Classifiers	FAO LCCS Classifier Codes
Tree Plantation	Visually identified training sites observed from field investigation with high NDVI values + pattern + size of the plot	A11 Cultivated and managed terrestrial Areas A1 Trees B5 Continues C1 Single crop
Cropland - Irrigated	Visually identified training sites using 2.5m data and field investigation + high NDVI value (around 0.6)	A11 Cultivated and managed t terrestrial Areas A3 Herbaceous D4 Surface irrigated
Cropland – Tree crop	Visually identified training sites observed from field investigation with moderate NDVI readings + pattern + size of the plot + associated land cover	A11 Cultivated and managed terrestrial Areas A2 Shrubs B5 Continues C1 Single crop
Woody vegetation (closed)	Visually identified training sites using 2.5m data and field investigation + high NDVI value (higher than 0.3) + closed woodlands (> 60%) + tree height is over 2.5m	A12 Semi natural terr. vegetation A1 Woody A10 Closed B1 Height 7 – 2 m
Woody open	Visually identified training sites using 2.5m data and field investigation + high NDVI value (higher than 0.3) + open woodlands (10 – 40%) + tree height is over 1m	A12 Semi natural terr. vegetation A1 Woody A21 Closed to Open B14 Height 5 – 05m
Grassland - farm	Visually identified training sites observed from field investigation with moderate NDVI value in Graminoids (grass)	A12 Semi natural terr. vegetation A6 Graminoids A10 Closed (>60-70%)
Built-up-Non-vegetated	Visually identified training sites using 2.5m data and field investigation + Special characteristics of artificial surfaces + GIS data to separate urban limits	B15 Artificial surfaces and assoc areas Urban Areas A13
Built-up-Impervious road surface	Visually identified training sites using 2.5m data and field investigation + Special characteristics of artificial surfaces + GIS data to separate urban limits	B15 Artificial surfaces and assoc areas A8 Roads - Paved
Non-built-up non vegetated	Visually identified training sites using 2.5m data and field investigation + Special characteristics of artificial surfaces + GIS data to separate urban limits	B15 Artificial surfaces and assoc areas A12 Industrial and other

Non-built-up-Mine/Quarry	Visually identified training sites using 2.5m data and field investigation + Special characteristics of the surface	B15 Artificial surfaces and assoc areas A2 Non Built Up A6 Extraction sites
Non-built-up-Impervious road surface	Visually identified training sites using 2.5m data and field investigation + Special characteristics of artificial surfaces + GIS data to separate urban limits	B15 Artificial surfaces and assoc areas A8 Roads – Paved (GIS data to separate urban limits)
Bare soil	Visually identified training sites using 2.5m data and field investigation + Special characteristics of the surface + associated land features	B16 Bare Areas B1 Dunes
Canals	Visually identified training sites using 2.5m data and field investigation + Special characteristics of water surface	B27 Primarily Non-Vegetated A4 Artificial water bodies, flowing
Inland water	Visually identified training sites using 2.5m data	B28 Primarily Non-Vegetated A4 Natural water bodies, flowing A7 Natural water bodies, stationary
Marine water	Visual identification + Spectral classification of 2.5m data and 10m data	B28 Primarily Non-Vegetated B1 perennial

5. RESULTS OF THE CASE STUDIES

This paper describes the characteristics of the Australian land surface and the application of FAO's LCCS to classify the surface-based characteristics into the land cover classes appropriate to the urban area of Brisbane and its vicinity.

5.1 Brisbane, the coastal urban region

Mapping urban environments using high resolution satellite data at 10m spatial resolution is a relatively an easy task, since spectral characteristics of artificial surfaces show a clear discrimination against vegetation. With the case of Brisbane city, it stands out clearly against the surrounding mostly vegetated areas of Queensland. The map shown here as Brisbane (Figure 5), is a subsection of a larger map that covers the whole of the southeast Queensland (SEQ) catchment, which was classified under this study. SEQ has a wide variety of land cover types and its Brisbane-Gold Coast coastal urban belt with over 1.5 million people (2007 Projected) (ABS, Australia) is the busiest urban region in the northern half of Australia. Table 4 presents land cover classes mapped in Brisbane area using the FAO LCCS approach.

An error matrix was used on the classified SEQ map to determine the percentage of land cover mapping accuracy. The error matrix used a total of 190 points and found total class accuracy for the entire SEQ Catchments area to be 90%. The selected sub-region in this study may record a different accuracy, if a separate sample is checked for the sub-region only. Specifically, the accuracy of separating the built-up areas from vegetated lands is high, compared to finding boundaries amongst vegetation types. Also, the inclusion of GIS data

layers into the final map of SEQ forced the classification process to establish some of the vegetation types as mixed vegetation classes, or to produce a separation between “Urban” and “Non-Urban” classes, which is not possible to determine through spectral classifications only.

5.2 The qualitative aspects of the new map

This study was conducted to apply the FAC LCCS system to an Australian urban and peri-urban cover classification. A full SPOT image scene was classified and a sub-region of 1000skm was presented in this study in order to present urban features clearly. To maintain homogeneity within each land cover class, classes have to be built with broad and easy to understand classifiers. A large number of classes based on micro-level local information is appropriate for local level detail mapping, and such a scheme must be organized in order to accommodate the national level land cover products.

We have used an approach based on spectral values of each image band and visual observation of the 2.5m super-resolution colour images, which are the basic needs for a classification. We then added field observation information to the training site selection and the category refinement process, which helped to generate the classifiers used to break level 3 classes into the 4th or final level classes. As explained earlier, the classification produced a satisfactory level of accuracy with the map while accommodating the classification scheme suggested by FAO approach. Figure 6 visually compares a small section of the new map with the existing SLATS 2001-2003 data set to indicate the improvement in land cover data when high resolution images were used.

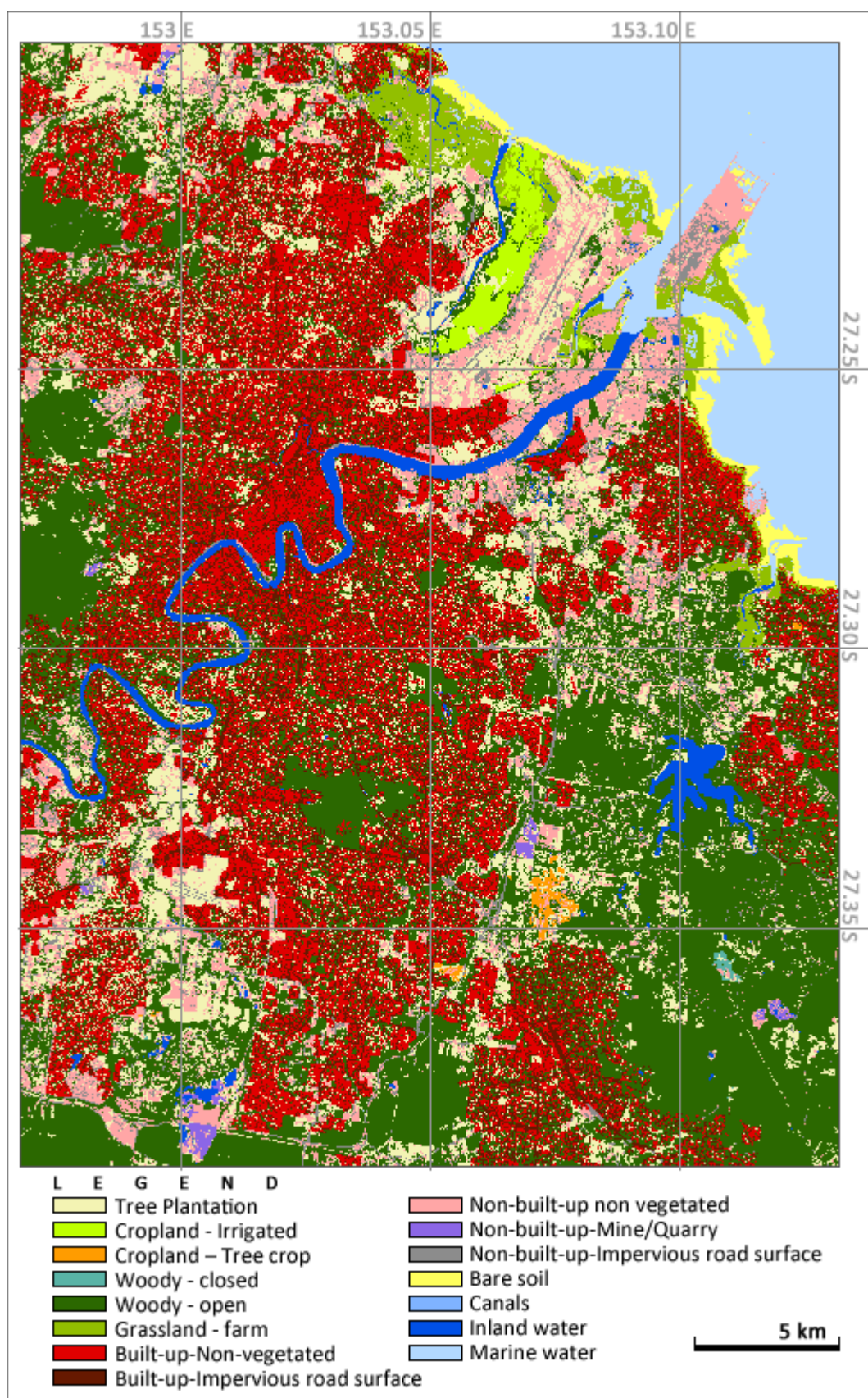


Figure 5. Land Cover Map of Brisbane vicinity.

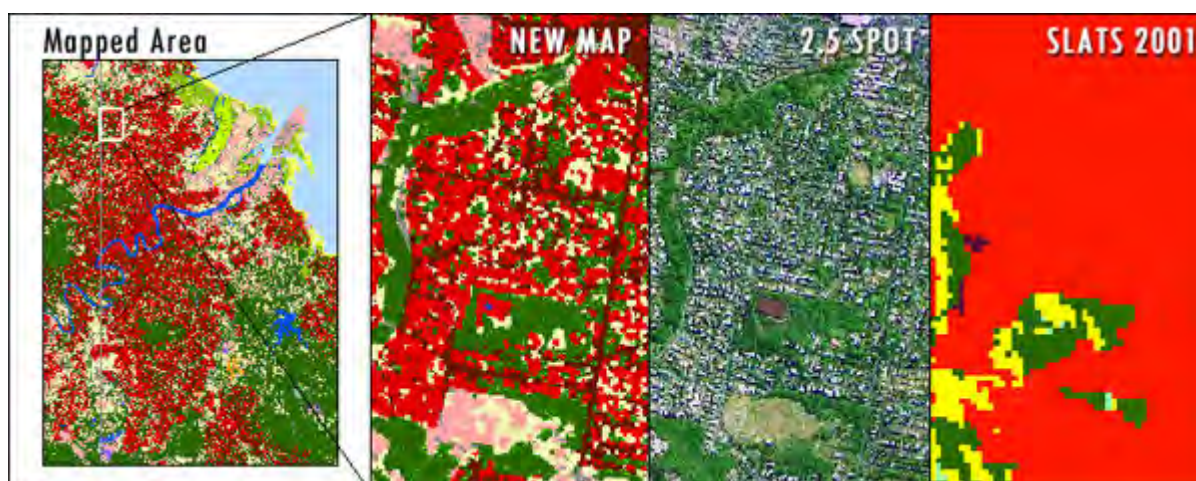


Figure 6. Comparison of land cover details of a sample area, where red indicates built-up areas.

6. CONCLUSIONS

Australia's agriculture and mining based economy requires an accurate assessment of land use and land cover. However, mapping the country at 10m or finer resolution has just started and over 90% of the country is yet to be mapped. This study classified the busiest urban land cover plot, Brisbane, from Queensland, Australia.

The prime objective of the study was to build the classification system common for the region using the fundamental approach of FAO Land Cover classification system. The FAO system has three initial class levels based on a *priori* (pre-defined) classification approach and the 4th detail level or the Modular-Hierarchical Phase. A careful observation of the spectral information against super resolution satellite data and ground survey information enabled classifiers for 4th level to be selected. Through the process, a number of different urban land cover types was identified apart from the typical woody vegetation and grasslands categories. Some classes ended with same name and same class identifier when the classifiers were similar to each other (i.e.; A12.2. woody open class).

The results show a promising outcome for mapping urban land cover under the FAO method. Enhancing the classification methodology with FAO LCCS was the prime object of the study; however, the map was completed with a significant accuracy proving the 10m spatial resolution as a useful data source for regional and national level land cover planning.

REFERENCES

- Africover LCCS, FAO, (2008). online document, <http://www.africover.org/LCCS.htm>
- Atyeo C. and Thackway R., 2006, Classifying Australian Land Cover, Australian Government, Bureau of Rural Sciences
- Australian Bureau of Statistics, (2008). <http://www.ausstats.abs.gov.au/ausstats/>
- Australian Bureau of Statistics, (2009-10). 3218.0 - Regional Population Growth, Australia
- Bom, (2011). Australian climate zones - all climate classes
www.bom.gov.au/climate/environ/other/kpn_all.shtml
- Butt, P., (2001). Land Law, Law Book Co of Australasia, 2001
<http://www.teamlaw.org/LandDef.htm>

- Climate statistics for Australian locations, (2011). Weather data, www.bom.gov.au/climate/averages/tables/cw_040214.shtml
- DAFF (2007). Department of Agriculture, Fisheries, and Forestry, Australia, <http://www.daff.gov.au/>
- Department of the Environment, (2008). Water, Heritage and the Art, Gondwana Rainforests of Australia, (2008). <http://www.environment.gov.au/heritage/places/world/gondwana/information.html>
- Di Gregorio A., (2005). FAO Land Cover Classification System, Classification concepts and user manual, software version 02.
- Di Gregorio A., and Jansen L. J. M., (2000). Lands cover classification system (LCCS), FAO.
- Hughes L., (2003). Climate change and Australia: trends, projections and impacts. *Austral Ecology* 28, 423-443.
- Hughes, L., (2005). Impacts of climate change on species and ecosystems: an Australian perspective, *The International Biogeography Soc.*, Summer 2005 Newslet.: Vol. 3, No. 2
- Michael F.H., Sue M., Richard J. Hobbs, Janet L. Stein, Stephen G. and Janine K., (2005). Integrating a global agro-climatic classification with bioregional boundaries in Australia. *Global Ecology and Biogeography*, 14, 197-212
- National Land & Water Resources Audit, (2007). Australian land cover mapping, <http://lwa.gov.au/files/products/national-land-and-water-resources-audit/pn21311/pn21311.pdf>
- Pittock, B., (2003). Climate Change: An Australian Guide to the Science and Potential Impact. Australian Government agency on greenhouse matters, 94-101 pp
- Soil map information <http://www.cazr.csiro.au/connect/resources.htm>
- State of the Environment, (2006). Independent Report to the Commonwealth Minister for the Environment and Heritage. Australian State of the Environment Committee. Land and Vegetation section, www.environment.gov.au/soe/2006
- World Resources Institute, (2006). Agro data; *World Resource Institute*, 2006, <http://earthtrends.wri.org/text/agriculture-food/country-profile-9.html>
- Year Book Australia, (2007). Urban and non-urban population, <http://www.abs.gov.au/>

A dense surface modelling technique for foot surface imaging

Jasim Ahmed Ali AL-Baghdadi, Albert K. Chong

Faculty of Engineering and Surveying, University of Southern Queensland
Toowoomba, QLD. Australia

jasim76@gmail.com, chonga@usq.edu.au

Kevin McDougall¹, Duaa Alshadli¹, Peter Milburn², Richard Newsham-West²

¹Faculty of Engineering and Surveying, University of Southern Queensland,
Toowoomba, QLD. Australia

²School of Physiotherapy and Exercise Science, Gold Coast Campus, Griffith University,
QLD, Australia

Kevin.mcdougall@usq.edu.au, duaa.alshadli@usq.edu.au, p.milburn@griffith.edu.au,
r.newsham-west@griffith.edu.au

ABSTRACT

Automated 3D point cloud generation of an object surface from images using a Dense Surface Modelling algorithm is a reliable technique. Recently, this technique has been applied in numerous mapping applications such as the human face, historical building facades, digital archaeological artefact recording and forensic investigation. In this paper, the technique is applied to the mapping of the dorsal and plantar aspect of a human foot during weight-bearing, which is considered a difficult surface for 3D mapping. The purpose of the research is to develop an approach that provides low-cost, high-quality 3D surface models which can be used to study the dynamics of the foot during slow-gait. The objective of this paper is to present the techniques used and the results of this investigation.

The research results show that the total gaps in the generated 3D plantar surface, was less than 0.1 percent. However, these gaps did not reduce the anthropometric mark's positional measurement accuracy as these marks could be clearly identified in the 3D model. The 3D representation of the dorsal surface of the foot during walking exhibits significantly fewer holes than the plantar surface at about 0.02 percent. All the defined anthropometric landmarks appear clearly on the dorsum of the foot's 3D surface, thus making digital measurements on the surface an easy task. Light rays coming from the plantar surface must pass through a 12 mm tempered glass and, depending on the camera's position, some of the light rays suffered refraction and reflection, making the gaps in the plantar surface reconstruction unavoidable. However, the overall accuracy of the developed photogrammetric measurement technique is approximately 0.3mm for all the generated surfaces.

KEYWORDS: *human surface texture, 3D surface model, photogrammetry, dorsal and plantar surface*

1 INTRODUCTION

The human foot consists of many components working together to provide the body with support, balance, and mobility. The foot is known to suffer from abnormalities due to genetic defects, accidental injury, or improper footwear associated injury. To protect the foot from injury, footwear should be designed and manufactured to fit the dynamic contours of the foot, but most often, they are designed from static models of the foot. Thus, it is necessary to know more than

just the static dimensions of the foot, and include its shape, flexibility and load during standing, walking gait or running. Accurate 3D modelling of the dorsal and plantar surfaces is one element that may allow researchers to investigate the loading dynamics and flexibility of the foot and guide footwear design. However, capturing accurate, dynamic 3D dorsal and plantar surfaces of the human feet during gait has been a challenge for measurement scientists. The improvement in second generation laser scanning devices is still not a suitable tool for fast gait as these high-speed devices still require a few tenths of a second to complete a scan. During walking gait, the 3D dorsal and plantar surface changes shape rapidly, and thus a fast-rate 3D data capture technique is required. Hawke et al. (2006) states that customised foot orthotics have shown to be more effective at reducing pain and redistributing pressure than standard “off the shelf” orthotics. The authors cite the need for dynamic data of the foot to design customised foot orthotics. Consequently, the capture of 3D dorsal and plantar surface data at slow (walking) gait will provide more benefits than those collected from a stationary foot in the design of footwear and orthotics for general mobility. The aim of the research is to develop an optical imaging technique to capture 3D surface data of the human plantar during low-speed gait.

Currently, digital electronic sensors, optical imaging sensors, flat-bed scanners, and laser scanning devices are available for 3D dorsal and plantar surface capture. Krauss, et al. (2008) investigated sex-related differences in foot morphology using a 3D-footscanner where 847 individual feet were scanned in a static posture. The purpose of their research was to determine whether sex-related differences were significant and thus provide footwear manufacturers with data for designing appropriate shoes for men and women. Tessutti et al. (2008) used a digital electronic sensor to investigate the in-shoe plantar pressure distribution to determine the relationship between chronic injury and surface type during running. Participants ran on two common surfaces: (1) asphalt and (2) natural grass at medium speed. Their investigation found that the type of surface used in running contributed significantly to the greater loads recorded on the rearfoot and forefoot due to differences in surface compliance. The authors concluded that running on grass is likely to cause less injury than running on asphalt over medium and long distance runs. Coudert et al. (2006) developed an anthropometric measurement system which determines the change in foot shape during walking to determine the parameters needed to develop footwear that can adapt its inner volume to foot shape changes. They found that when the forefoot is on the ground, the width of the foot increases by about 5 mm, or about 5%. Zhao et al. (2008) used 3D data of the foot captured by laser scanner to determine girth measurements such as the ball girth, instep girth, waist girth, long heel girth, and short heel girth. These measurements were essential dimensions for the design of personalized footwear. The authors recommended the use of digital 3D feet in footwear manufacturing CAD/CAM systems, thus allowing customized footwear design and manufacturing. Martedi et al. (2009) developed a technique to produce a 3D surface of the plantar using the 2D surface captured by a flatbed scanner. While the authors argued that the error was below 1.0 mm, there were drawbacks such as the generation of accurate sole shape reconstruction. In general, the above-mentioned techniques were not capable of capturing accurate digital 3D surfaces of the dorsal and plantar surface during gait. Therefore, determining the suitability of optical video imaging systems and the photogrammetric technique for obtaining dynamic surface contours of the foot was the objective of this investigation.

2 EQUIPMENT AND SOFTWARE

2.1 Imaging Platform and Synchronising Device

A stable glass-top step-on platform for the human subject to stand on during imaging and a camcorder mounting platform were constructed in order to capture the dynamic surface contours of the foot. The camera mounting platform was designed to provide 100 percent three-image overlap of the plantar surface area (Table 1). In addition, a low-cost electronic synchronising device was constructed to provide video-frame synchronization for the multiple camcorder arrangement.

Table 1, Video camera type and imaging sensor specifications.

Brand	Range of FL (mm)	Effective format size (mm)	Pixel Count (pixel)	Pixel Size (mm)
2x Sony HDR SR-10	3.1-46.5	Still: 2.765x1.555 Video: 2.304x1.296	Still: 2304x1296 Video: 1920x1080	0.0012 0.0012
1x Sony HDR-XR150	2.5-62.5	Still: 2.375x1.782 Video: 2.142x1.253	Still: 2048x1536 Video: 1920x1080	0.0011 0.0011

2.2 Photogrammetric Control

A set of 10 to 30 retro-targets with known coordinates was required for accurate 3D scaling using the photogrammetric technique. The coordinates of these points were determined using a set of 12 convergent (four vertical and four rotated 90° to the left and four rotated at 90° to the right) images taken with a calibrated fix-mount lens camera. The process was repeated three times. A high-precision Invar scale bar was used in the calibration that was calibrated by the manufacturer with a factory-calibration accuracy of $\pm 11 \mu\text{m}$ at 15° C.

2.3 Photogrammetric Software

Australis® photogrammetric software was used mainly for the self-calibration of the imaging sensors using a bundle adjustment technique which is based on the mathematical equations 1 and 2 below. The software is also suitable for the determination of the high-accuracy retro-target coordinates used in photogrammetric control. In general, the coordinate system of the computed coordinates is based on the geometry of the orientation device used while the scaling is based on the Invar scale-bar used.

$$\Delta x' = \Delta x'_0 - \frac{x'}{PD} \Delta PD + K_1 x'^2 + K_2 x'^4 + K_3 x'^6 + P_1 (r'^2 + 2x'^2) + 2P_2 x' y' - C_1 x' + C_2 y' \quad (1)$$

$$\Delta y' = \Delta y'_0 - \frac{y'}{PD} \Delta PD + K_1 y'^2 + K_2 y'^4 + K_3 y'^6 + 2P_1 x' y' + P_2 (r'^2 + 2y'^2) + C_2 x' \quad (2)$$

where

$\Delta x', \Delta y'$: axis-related correction values for imaging errors,
 $\Delta x'_0, \Delta y'_0$: small corrections for perspective centre (X_p , and Y_p),
 ΔPD : small correction for principal distance,
 K_1, K_2, K_3 : lens distortion parameters,
 P_1, P_2 : lens decentring parameters,
 C_1, C_2 : affinity and shear parameters, and
 r' : radial distance.

3D surface models were created using *PhotoModeler Scanner®* software. The software is capable of scanning image pairs to produce dense point clouds and meshed surfaces using the Dense Surface Modelling (DSM) technique. The software allows the export of point clouds and triangulated meshes in many formats, thus allowing further data-cleaning, data manipulation and analysis.

3 METHODOLOGY

3.1 Camera Calibration and Test-Subject Imaging

The selected high definition (HD) video cameras were calibrated individually using a custom-made camera calibration testfield at an object distance of 1000 mm. This object distance was similar to the gait platform design specifications for imaging the plantar surface. In the calibration, eight sets of convergent video clips of the testfield were captured; individual frames were extracted from the clips; and the frames were processed using *Australis*® off-the-shelf camera calibration software. Subsequently, the cameras were mounted on the prototype imaging platform which was positioned under the glass-top step-on platform. Spot lights were placed around the platform to ensure the brightness was uniform across the sole of the foot. When the subject was ready to start walking on the step-on platform, the camcorders were started and the LED synchronisation device activated.

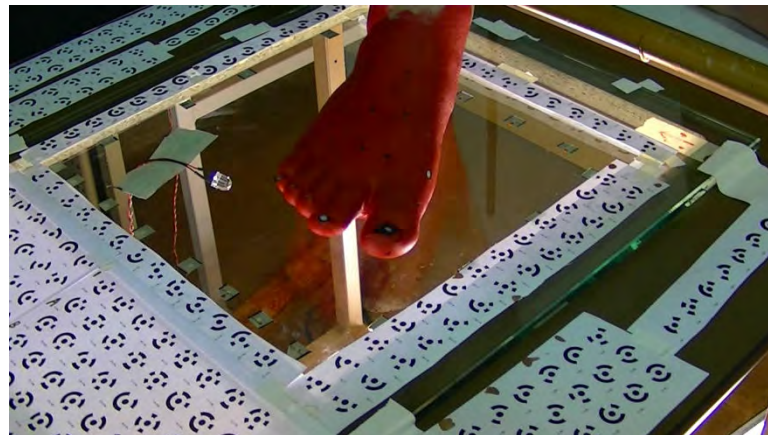


Figure1: An extracted image frame showing: 1) image registration target array; 2) the LED synchronization device.

3.2 Skin-Texture Enhancement and Anthropometric Landmark Marking

Tests were carried out to determine the best skin texture enhancement technique for human subjects of varying skin-colours. Test results showed that uniform accuracy of the captured 3D surface of the sole can be attained if the skin was painted with a matte red 'face and body' paint, thus creating a uniform matte colour surface. In this investigation an off-the-shelf matte red-colour face and body paint was applied to the subject's foot. The paint was lightly applied to the sole ensuring no streaks or lumpy appearance and allowed to dry before signalized targets were placed on anthropometric landmarks on the foot. Also, image non-signalised registration marks were placed in the optimal location using black-ink marker pens.

3.3 Gait Method

The test-subject performed a set of slow steps over the glass-top of the step-on platform. In the first activity, the subject stepped on the glass-top and, raised the heel to simulate a gentle step-forward act. In the second activity, the subject placed his foot to the glass-top and gently raised the front of the foot. To minimize error introduced by the disparity of the individual video camera timing error, the LED synchronization device was activated before each task.

3.4 Image Processing in PhotoModeler®

The captured images of each gait movement were processed in the *PhotoModeler*® software using the DSM technique. Image-pairs were uploaded and relatively orientated. The imaged coded targets were used mainly to expedite the orientation process. It was found that black/white circular targets were more accurately digitized by the software. Initially, a depth-range was assigned to

expedite the search for a good match in the subsequent images. This setting was particularly important for this project as human skin surface has weak texture. Once the total error was less than 1.0mm, the orientation was considered satisfactory. As the dorsal and plantar surfaces were smooth, a medium density rate (medium sample rate value) was applied because a high density resulted in wavy and rippled model appearance and a low density resulted in insufficient details on the model.

4 RESULTS AND ANALYSES

4.1 Video Camera Calibration

Three video cameras were calibrated in the laboratory and the average of four independent sets of the computed lens parameters is provided in Table 2. The table shows the manufacturer's stated focal length in brackets and the calibrated principal distance (PD). The Sony XR150 video camera gave the largest variation between the two sets of values while the Sony SR10 had the least variation. Because the differences were approximately +3.5%, +1.2 % and +2.0%, respectively, the approximate scaling error for the three-camera system of +2.2 %, based on the manufacturer's focal lengths, was adopted in the image processing. In other words, at an object distance of 1000 mm, the computed 3D distance of a 350 mm long human foot would be approximately 8 mm longer if calibration was not carried out. Also, in Table 2, the presented values for parameter K_2 were very small and their contribution to the computed 3D surface error was estimated to be less than 0.05 mm in the real-world dimension. As the radial lens distortion parameter K_3 , the decentring distortion parameters and the dynamics fluctuation parameters (Eq. 1 and 2) were less likely to be significant, they were removed in the calibration and not included in Table 1. Because the impact of the principal point offset and radial lens parameters (K_1) were significant, the video cameras should be calibrated to achieve high-quality photogrammetric measurements.

Table 2. Video camera calibration results

Brand	PD (mm)	X_0	Y_0	K_1	K_2
Sony XR150	2.691 (2.6)	-0.026	-0.002	-0.0036	0.00027
Sony SR 10	3.038 (3.0)	0.049	0.067	0.0021	0.00019
Sony SR 10	3.059 (3.0)	0.044	0.063	-0.0018	-0.00022

4.2 3D of the Plantar Surface

The 3D surfaces generated from the captured images of the plantar surface (left) and the generated 3D surface using the three captured images (right) are presented in Figure 2. The generated 3D surface clearly showed the compressed overlaying fat pad at the heel. The total amount of gaps in the generated 3D surfaces was less than 0.1 percent and did not reduce the anthropometric mark positional measurement as these marks were clearly identifiable. Mapping the individual toes presented a minor problem as the toes were very close together. However, because the plantar surface was the main interest, this problem was considered minor and hence there was no significant drawback in using this mapping technique. Figure 3 shows the 3D surface of the dorsal aspect of the foot being raised above the glass-top platform. The proportion of gaps was significantly less (0.02 percent) than the 3D plantar surface (Fig. 2) due to light rays from the plantar surface passing through a 12 mm tempered glass, and potentially experiencing refraction and/or reflection. In addition, the lighting around the plantar surface created bright and dim spots on the plantar surface of the foot as a result of the glass plate. Figure 4 shows the dorsal surface of the foot when the heel is raised above the glass plate. This 3D surface has the least number of holes as the forward tilt of the foot allowed the light to provide an even brightness of the imaged surface. All the marked anthropometric landmarks appeared clearly on the 3D surface, thus making digital measurements on the surface straightforward.

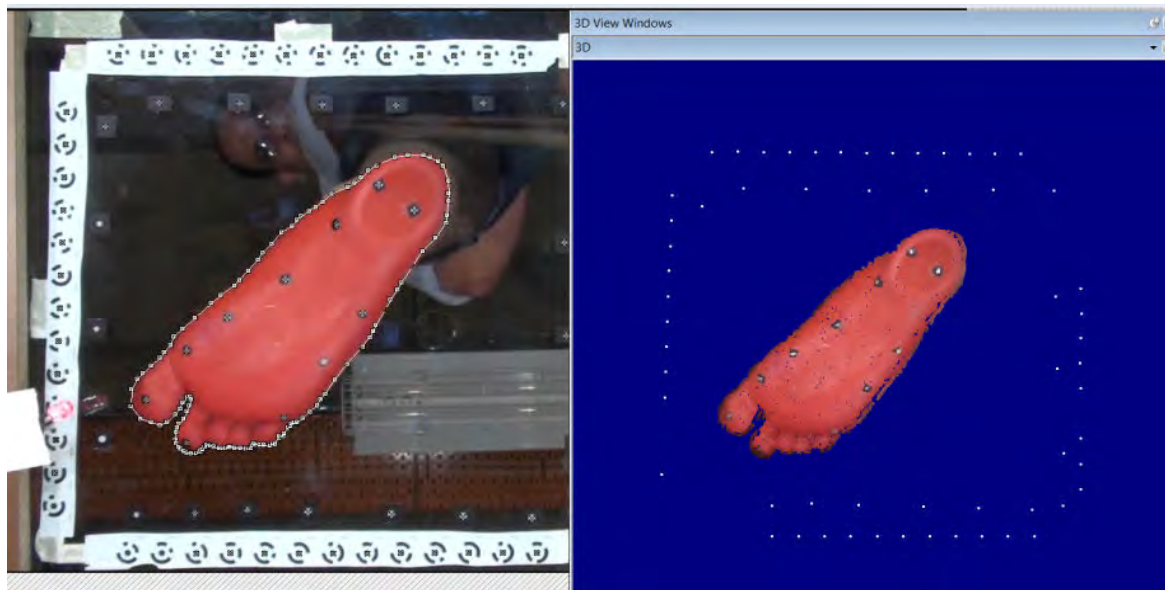


Figure 2. (Left) Raw image from video clip; (right) the processed 3D surface.

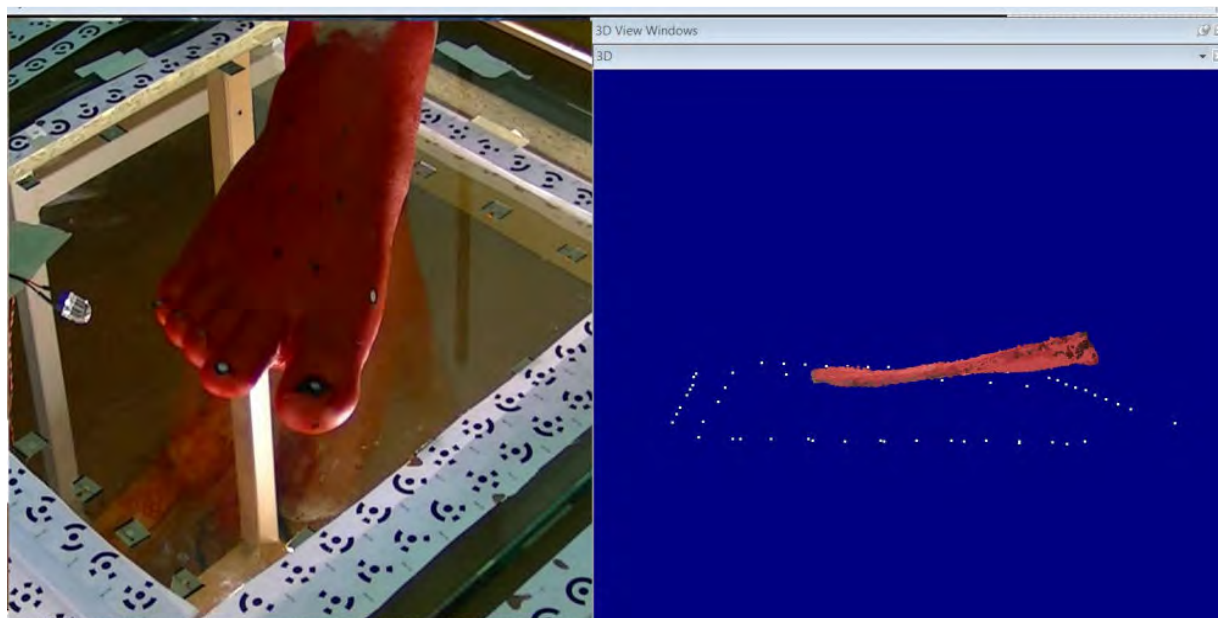


Figure 3. (Left) Video image of front of foot raised; (Right) Front of foot raised above glass-top.

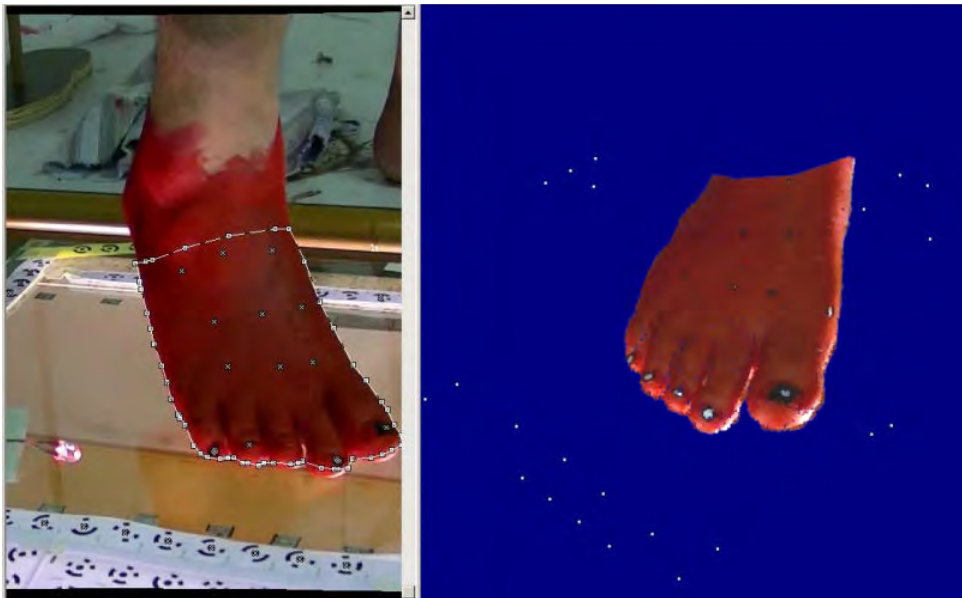


Figure 4. (Left): Video image of raised heel; (right) processed 3D surface model

4.3 3D Surface Modelling Accuracy

The 3D distances between signalised-targets were used for accuracy evaluation and were calibrated with a precision of ± 0.015 mm. Based on the differences between the calibrated and the measured 3D distances of the 3D surface model, the technique used achieved an accuracy of 0.250 mm, thus confirming the high-quality of the computed 3D surface model.

5 BENEFITS AND DRAWBACKS OF THE DSM PHOTOGRAMMETRIC TECHNIQUE

It is apparent that there are a number of benefits and drawbacks in the developed photogrammetric technique for the generation of a 3D surface model. The benefits include:

1. Dynamic characteristics of the dorsal/plantar surface can be mapped at 25 (SECAM/PAL) or 30 (NTSC) Hz which is sufficient for 3D recording of plantar motion during walking gait,
2. Image capture using standard video cameras has few restrictions; are easy to operate, and video clips can be processed easily.
3. The overall achievable accuracy of the generated 3D surface of the dorsal/plantar surface is approximately 0.3 mm,
4. Low cost video cameras (approximately \$US 200) are suitable for the application,
5. The camera mounting platform, a simple glass-top step-on platform and the photogrammetric control can be fabricated relatively cheaply, and
6. In general, little or no regular maintenance is required.

The drawbacks are:

1. Face and body paint is required to obtain suitable images of different skin-colours,,
2. A small percentage of holes may appear in the generated 3D surface, particularly along the edges, unless the captured images include the side of the foot,
3. The technique is not suitable for individual toe mapping,
4. The system as presented here is suitable for walking gait only as the video frame-rate is low.
5. High definition video cameras are essential for good quality imaging, and
6. The imaging processing software requires DSM functionality for rapid generation of the 3D surface. The cost of typical DSM software such as PhotoModeler® for commercial purposes is in excess of US\$ 5000.

6 CONCLUSION

The objective of the research was to develop a low-cost, high-quality 3D surface model which could be used to study the dynamics of the plantar surface of the foot during walking gait. A three video camera mounting platform, a glass-top step-on platform and a set of photogrammetric targets were constructed and used for imaging and the captured images were processed using PhotoModeler® Scanner.

The total proportion of gaps in the generated surface of the plantar surface of the human foot was approximately 0.08 percent. However, these holes did not reduce the anthropometric mark positional measurement as these marks could be clearly identified. However, individual toe mapping presented a minor problem as the toes were close together and difficult to map individually. The 3D surface of the dorsal surface of the foot showed significantly less holes than the plantar surface (0.02 percent) due to light not having to pass through a 12 mm tempered glass and experiencing light refraction and/or reflection. All the anthropometric landmarks appeared clearly on the dorsum of the foot's 3D surface thus making digital measurements on the surface convenient. The overall accuracy of the developed optical measurement technique was approximately 0.3 mm.

The results show that the developed technique using off-the-shelf HD video cameras is suitable for the 3D surface mapping of the plantar and dorsal surfaces of the foot. This information provides a unique insight into the dynamic characteristics of the foot during gait which can be incorporated into the design of footwear and orthotics.

REFERENCES

- Coudert T., Vacher P., Smits, C., Van der Zande, M., 2006. A method to obtain 3D foot shape deformation during the gait cycle, *Ninth International Symposium on the 3D Analysis of Human Movement*, June 28-30. Valenciennes, Frances, <http://www.univ-valenciennes.fr/congres/3D2006/Abstracts/117-Coudert.pdf>
- Hawke, F., Burns, J., Radford, J.A., du Toit, V., 2008. Custom-made foot orthotics for the treatment of foot pain. *Cochrane Database Systematic Reviews*, 16(3): CD006801
- Kraussa, I., Graua, S., Mauchb, M., Maiwaldc, C., Horstmannnd, T. 2008. Sex-related differences in foot shape. *Ergonomics*, 51(11): 1693-1709.
- Martedi, S, Saito, H., 2009. Shape measurement system of foot sole surface from flatbed scanner image. *MVA2009 IAPR, Conference on Machine Vision Applications*, 20-22nd June. Keio, Japan [http://www.hvrl.ics.keio.ac.jp/paper/pdf/MVA2009_sandy.pdf].
- Tessutti, V., Trombini-Souza, F., Ribeiro, A. P., Nunes, A. L., Neves Sacco, I.N., 2010. In-shoe plantar pressure distribution during running on natural grass and asphalt in recreational runners. *Journal of Science and Medicine in Sport*. 13: 151–155.
- Zhao, J., Xiong S., Bu, Y., Goonetilleke, R.S., 2008. Computerized girth determination for custom footwear manufacture. *Computer & Industrial Engineering*, 54:359-373.

Automated Matching of Segmented Point Clouds to As-built Plans

David Belton^{1,2}, Brian Mooney², Tony Snow² and Kwang-Ho Bae³

¹CRC for Spatial Sciences

²Department of Spatial Sciences, Curtin University
GPO BOX U1987, Perth, Western Australia, 6845, Australia

³Fugro Spatial Solutions Pty Ltd,

18 Prowse Street, West Perth, WA, 6005, Australia

d.belton@curtin.edu.au, brian.mooney@student.curtin.edu.au, t.snow@curtin.edu.au,
k.h.bae@furgospacial.edu.au

ABSTRACT

Terrestrial laser scanning (TLS) is seeing an increase use for surveying and engineering applications. As such, there is much on-going research into automating the process for segmentation and feature extraction. This paper presents a simple method for segmenting the interior of a building and comparing it to as-built plans. The method is based on analysing the local point attributes such as curvature, surface normal direction and underlying geometric structure. Random sampling consensus (RANSAC), region growing and voting techniques are applied to identify the predominant salient surface feature to extract wall and vertical segments. This information is used to generate a 2D plan of the interior space. A distance weighted method then automatically locates the corresponding vertices between the different datasets to transform them into a common coordinate system. A traditional survey was performed alongside the 3D point cloud capture to compare and validate the generated 2D plans and the comparison to the existing drawings. The accuracy of such generated plans from 3D point clouds will be explored.

KEYWORDS: Terrestrial Laser Scanning, Segmentation, 3D Point Cloud, Matching

1 INTRODUCTION

Laser Scanners enable users to sample dense three-dimensional (3D) point data from surfaces with high spatial resolution, which allow for the representation and modelling of salient features. This raw point data is commonly referred to as point clouds, and with post-processing can be converted into model or vector format and salient information can be resolved. Terrestrial Laser Scanning (TLS) has gained popularity in areas such as 3D-reconstruction of terrain (Frank et al., 2007), deformation monitoring (Gordon, 2005), building segmentation (Miliareisis et al., 2007) industrial modelling (Rabbani and van den Heuvel, 2004), and other conventional surveying applications.

The need for reconstruction of building environments has lead to the research in developing automation techniques. Examples of such techniques are based on recognising geometric structure (Vosselman and Dijkman, 2001), sweeping vertical plans to detect wall features (Budroni A. and Böhm, 2009), random sampling and Hough transforms for detecting elements (Tarsha-Kurdi et. al., 2007), 2D density histogram for detecting the occurrence vertical features (Okorn et. al., 2010) and region growing on the surface normal direction (Rabbani et. al., 2006).

This has lead to investigations on the comparative ability of TLS with traditional techniques and practices. The systematic errors for TLS are, in principle, the same as that of an EDM instrument, which is extensively described by Rüeger (1990). Even with their increased usages in the surveying industry, there are still questions concerning the spatial quality of the information

captured by Terrestrial Laser Scanners, especially in terms of accuracy. Studies on the positional accuracy of terrestrial laser scanners have been conducted in the past (Boehler et al. 2003, G. Mólnar et al. 2009, Lichti 2007). There has been few studies on the comparison between traditional surveying and laser scanners (e.g. Lichti et al. 2005, Froehlich and Metternleiter 2004), especially in the area of studies with respect to as-constructed surveys.

This paper aims to automatically extract 2D drawings, and analyse conventional total station surveys and Terrestrial Laser Scanning surveys compared to as-constructed plans. This is done in order to demonstrate that the Terrestrial Laser Scanner can be used as a comparable surveying method for indoor applications.

2 METHODOLOGY

The test site used for the procedures outlined in this paper was located on the third level of the Department of Spatial Sciences at Curtin University. Workflow consisted of three stages; acquisition of the data and control network by survey and TLS, processing the 3D data to produce a 2D drawing, and aligning and comparing the 2D drawing with the control survey and as-built plans in terms of accuracy. The control survey allows for generating as-built plans using traditional techniques for comparison, alignment of the point cloud to the plans, and geo-referencing the different aspects to a global coordinate system, which in this case was MGA94 and AHD94 for the easting and northing components, and the elevation component, respectively. Processing of the TLS point cloud data is aimed at automatically producing a 2D line drawing representation of the data, and defining vertices and control points that are common between the control and as-built survey, and the 2D plans. An automated method of comparing the network of control points and vertices is used by examining the distance matrix between the different elements, and finding the most likely corresponding elements between the datasets.

2.1 Control survey

While not a necessary component of the project, to allow for geo-referencing of the point cloud data, a control network was established by traversing between two local survey marks and using the two shaft method (Anderson and Mikhail, 1998). This method involves having two wires separated by the longest distance possible, in this case 110m, with known coordinates established from adjacent surface control. A traverse was then carried out between the two wires using control points placed in strategic locations of the third floor of the building. This method was chosen over other shaft plumbing methods, e.g. the Weisbach triangle method, because it produces a lower orientation error (Anderson and Mikhail, 1998; Schofield and Breach, 2007).

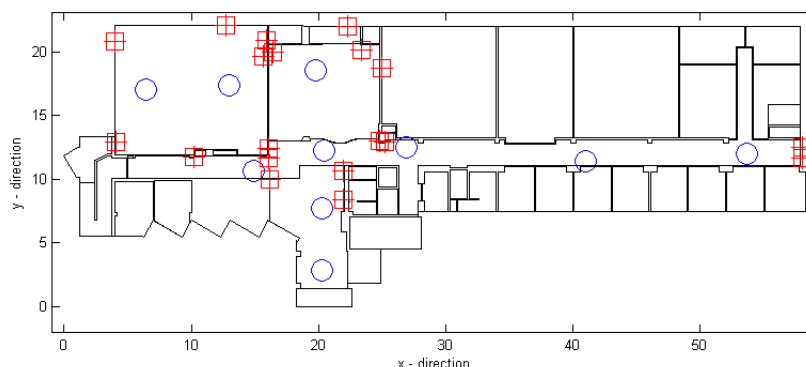


Figure 1: 2D plan with scanner setups represented by blue circles and geo-referenced targets denoted by red crosses.

This method allows the placement of a network of 19 targets for registering the separate scans into a local coordinate system, and to align the registered point cloud to a global coordinate system. Figure 1 shows the placement of these targets relative to the building scans. These 19 geo-referenced targets (black and white Leica paper targets) were placed and positioned utilising a Sokkia total station (Set530RK3; Sokkia, 2011). The measurements were carried out in reflector-

less mode, and measured 5 times with the results averaged to and adjusted. Targets were acquired within an accuracy of $\pm(3\text{mm} + 2\text{ppm})$ for the distance and 5" for the angle. To align and compare the point cloud, control survey and the 2D plans together in a common coordinate system, a control survey was also conducted to generate vertices representing the corners and intersection of walls.

2.2 Point cloud acquisition

The point cloud of the test site was captured using a Leica Scanstation (Leica 2011). Due to the need to traverse through narrow corridors, 10 setups were used to ensure adequate coverage and capture of the interior. The locations of these stations are represented in Figure 1 by the blue circles. The network of 19 geo-referenced targets outlined in the previously section were used to register the separate scans setup into a single point cloud which is tied into the absolute coordinate. 3D point cloud post-processing software provides a set of semi-automatic tools for this registration procedure. In this case, Cyclone (Leica, 2011) was utilised which can detect both reflective and paper-based targets. Once all the targets are detected and labelled, the Cyclone software utilises an iterative processing method to find the optimal transformation solutions for each set of point clouds, e.g. the ICP and its variants (Bae and Lichti, 2008).

Table 1: Statistics on the differences between control points by total station (geo-referenced) and registered point cloud.

	$\Delta\text{Easting (m)}$	$\Delta\text{Northing (m)}$	$\Delta\text{Height (m)}$
Mean	-0.2564×10^{-4}	-0.5128×10^{-4}	-0.2564×10^{-4}
Standard Deviation	0.001724	0.001605	0.001953

Although the Cyclone software can successfully detect targets in an automated manner, there can be instances where features present in the point cloud are incorrectly identified as targets. As such, a manual inspection or outlier detection method is used to identify and remove such occurrences. The final registration error is presented in Table 1, which shows the mean registration error is in the order of 2mm. This is well within the expected positional error for the Leica Scanstation. The registered point clouds for this experiment are presented in Figure 2.

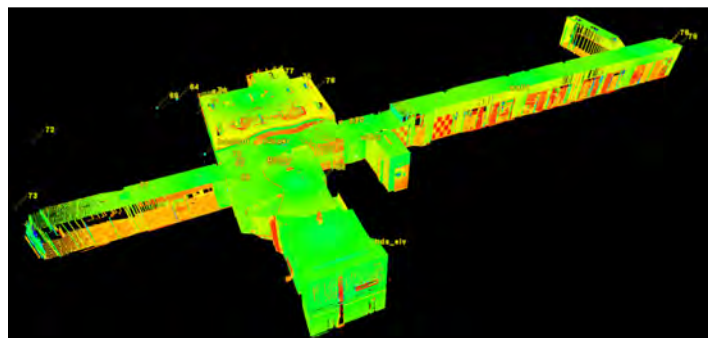


Figure 2: Registered point cloud of corridors on level three of the spatial sciences building.

In order to compare the geo-registered 3D point clouds and the CAD model, a 2D line model must be derived from 3D point cloud data. This transformation calculated between the x and y coordinates of the 2D drawings and the Easting and Northing components of the 3D point. Automated methods for performing this are outlined in the following sections.

3 AUTOMATED SEGMENTATION AND FEATURE IDENTIFICATION

The manual extraction of features is a labour intensive task. Therefore, much research has focused in the development of automated methods of segmenting and extracting features from

point clouds. When the objective is to derive the 2D plan, the process becomes much simpler as extraction methods are focused mainly on identifying and isolating vertical surfaces that correspond to walls, windows, doorways and other vertical features relevant to 2D drawings. The method outlined in this section is divided into two stages. The first stage is to identify all points that are potentially sampled from a vertical surface, and segment them into continuous regions. The second stage is to model these regions into 2D elements, and identify features that will comprise of vertices in a 2D drawing. These will allow for a comparison to be made with the as-built plans and the control survey by matching corresponding features between the different data sources.

3.1 Data processing and Region growing

Since the aim is to identify points that are candidates associated with walls, a method to classify such points is based on the local surface normal and surface curvature. A simple method used to derive this information is by fitting a planar surface to the local neighbourhood surrounding each point using principal component analysis (PCA) (Johnson and Wichern, 2002), and using the surface variance to determine the local curvature. Applying PCA produces a covariance matrix for a local neighbourhood such that:

$$C = \frac{1}{k-1} \sum_{i=1}^k (p_i - \bar{p})(p_i - \bar{p})^T \quad (1)$$

where p_i is defined as the vector form of the position of the i^{th} point in the neighbourhood containing the nearest k points and \bar{p} represents the centroid of the neighbourhood calculated as the mean of the neighbourhood. The covariance matrix can be represented by eigenvalue decomposition such that the real positive eigenvalues, λ_0 , λ_1 and λ_2 , along with the corresponding eigenvectors e_0 , e_1 and e_2 form an orthogonal basis of the neighbourhood in R^3 (Golub and Loan, 1989). The covariance matrix C can be decomposed as follows:

$$C = \sum_{i=0}^2 \lambda_i e_i e_i^T \quad (2)$$

where $\lambda_0 \leq \lambda_1 \leq \lambda_2$. Note that eigenvectors e_i represent the principal components, with the corresponding eigenvalues λ_i denoting the variance in these directions (Golub and Loan, 1989). For a local neighbourhood of a point cloud, e_0 approximates the local surface (Pauly et al., 2002). This result is equivalent to the first order least squares plane fit (Shakarji, 1998). From this, an approximation for the surface curvature can be specified, as presented in Pauly et al. (2002), by the following:

$$\kappa \approx \sigma_n^2(p) = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \quad (3)$$

Using this information, a point can be classified as a likely candidate to have been sampled from a wall if the surface normal is approximately orthogonal to the vertical direction and the surface is locally flat (i.e. the surface curvature is nominally zero). Thresholds can be applied to test for this attributes. Figure 3(a) shows such classified point with the following thresholds:

$$\begin{aligned} \kappa_i &\approx \sigma_n^2(p) < 0.001 \\ \theta_i &= |90 - \arccos(n_i[z])| < 5^\circ \end{aligned} \quad (4)$$

where $n_i[z]$ is the vertical component of the normal direction for point p_i . To refine this information, a technique such as the random sampling consensus method (RANSAC) can be applied (Fischler and Bolles, 1981). This method iteratively selects random sub-samples from the local neighbourhood and fits a surface (in this case a plane) to find a set of points that are within a set tolerance of the points (termed the consensus set). The plane with the largest number of points in the consensus set is selected as the best fit, and the associated points will have the surface attributes of the fitted plane. If points do not belong to any consensus set, it can be said that they are not sampled from a locally planar surface.

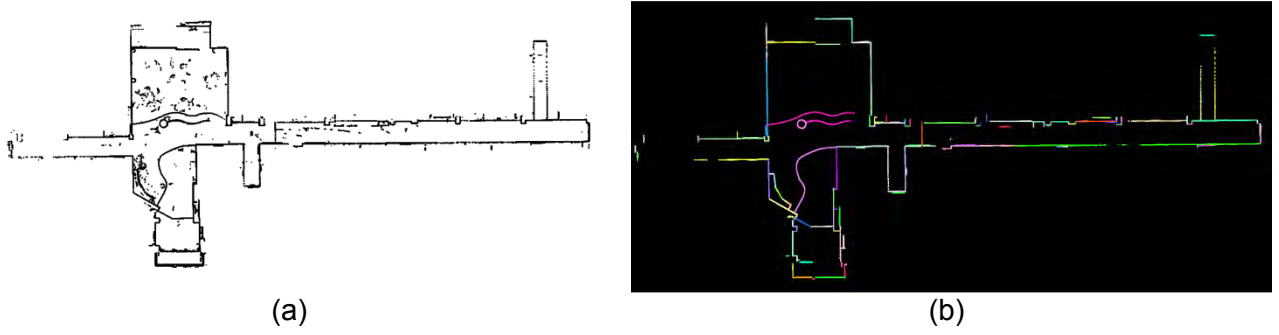


Figure 3: (a) the classified points that are likely sampled from vertical surfaces. (b) the segmented points of different vertical surfaces. Different colours denotes different surface segments.

From these classified points, continuous surface segments can be extracted to determine the sampled points that belong to each vertical wall feature. There are several options to segment classified points. Some of these include clustering or analysing in the normal orientations (Budroni and Böhm, 2009), Hough transformations (Vosselman and Dijkman, 2001), RANSAC for shape and region extraction (Schnabel et. al., 2007). In this paper, a simple region growing method was applied on the normal and curvature values to segment the points into regions representing smooth surfaces (Rabbani et. al., 2006).

The region growing process starts by selecting a seed point from the previously classified points. Points in the local neighbourhood (within a certain range) around the point of interest are examined based on the difference between normal directions and the distance between them in the normal direction. Those points in the neighbourhood where both attributes are within a set tolerance to those of the point of interest (nominally zero) are added to the same segment as the point of interests. A new point of interest is selected from the added points until all such points have been examined, and no additional points can be used. A new seed point is selected and the procedure repeated until every point has been visited. Segments with an insignificant amount of points are ignored. Figure 3(b) shows the results of the procedure applied to the test data set. Refinements to the segmentation process can be made by including curvature, principle directions and boundary conditions (Belton and Lichti, 2006).

3.2 2D plan extraction and Feature identification

Because the point cloud has been orientated to the horizontal plane, the vertically aligned features will be of importance to generating the 2D plans (for example, walls doorways, windows, pillars, etc). As such, the previously classified points and segments can now be examined in the 2D horizontal domain (the x-y axis or easting and northing). If the raw points are examined, then the walls can be extracted by using voting techniques on the normal direction and point location (Knuth, 1999). This can be done by examining the angular direction (α_i) and distance (δ_i) from the origin for every point, which is defined by:

$$\begin{aligned}\delta_i &= \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \\ \alpha_i &= \arctan\left(\frac{n_i[x]}{n_i[y]}\right)\end{aligned}\tag{5}$$

where x_i and y_i are from the coordinates of the i^{th} point, x_0 and y_0 are from the coordinates of the origin, and $n_i[x]$ and $n_i[y]$ are the corresponding x and y elements of the normal direction respectively. A clustering algorithm can then be applied to find concentration of points with similar parameters, similar to the Gaussian sphere approach in 3D (Várady, 1998). These clusters of points can then be modelled as a continuous line segment. Similarly Hough transformations (Vosselman and Dijkman, 2001) can be applied to the 2D point coordinates, similar to those mentioned in the previous sections (Budroni and Böhm, 2009), as well as a region growing process on the attributes described in Equation 5. Since the points have already been segmented, these procedures are not necessary in this instance. However, they are mentioned because if the sample density of the points is too sparse in the 3D domain, using the described methods in this section

may produce more robust results as the sample density in just the 2D horizontal domain will be higher than in the 3D domain.

For each of the segments, a line is fitted to the 2D points using PCA on the x and y coordinates, to get a mean point to represent the centre (c_j) of the line and with the associated direction (d_j). If the RMS value for the line fit is significantly large, then the segment can be categorised as a non-straight line segment. In this case, other line types can be fitted such as circles, arcs or splines, using least squares. For a point i belonging to a straight line segment j , the distance along the line segment can be specified as:

$$s_i = (x_i - c_j[x])d_j[x] + (y_i - c_j[y])d_j[y] \quad (6)$$

The points with the largest and smallest values of s_i denote the extents (p_1 and p_2) of the straight line segment such that:

$$\begin{aligned} p_1 &= c_j + s_{min}d_j \\ p_2 &= c_j + s_{max}d_j \end{aligned} \quad (8)$$

These points can then be used to create the line segment for the 2D plan drawing, as shown in Figure 4 (a). To extend these line segments, the extents of a line segment are tested to determine if it is close to another line segment. If it is, then it is likely that the extent can be modified to extend and intersect to the other nearby line segment. These cases are illustrated in Figure 5 (a) and (b). If the directions of the lines are nominally aligned, the two segments likely belong to the same element and are merged together to create a new segment, as shown in Figure 5 (c). For the case where the line extents are close to one another (within a specified tolerance), but the directions differ significantly, then the intersection of the two lines can be calculated. This intersection can then define the new extents of both lines. Figure 4(b) shows the results of applying this procedure to the initial line segments from Figure 4(a).

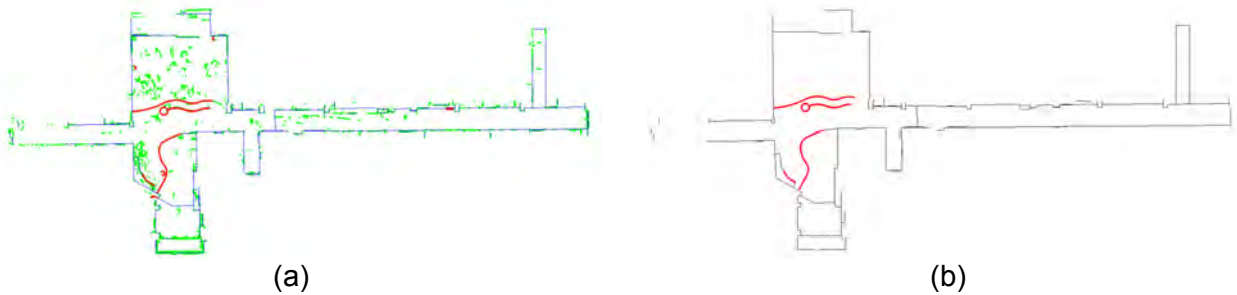


Figure 4: (a) the initially extracted 2D straight lines (blue), extracted splines for non-straight segments (red), and the raw points (green). (b) the final extracted 2D drawings, with straight lines (black), and more complicated sections modeled by splines (red)

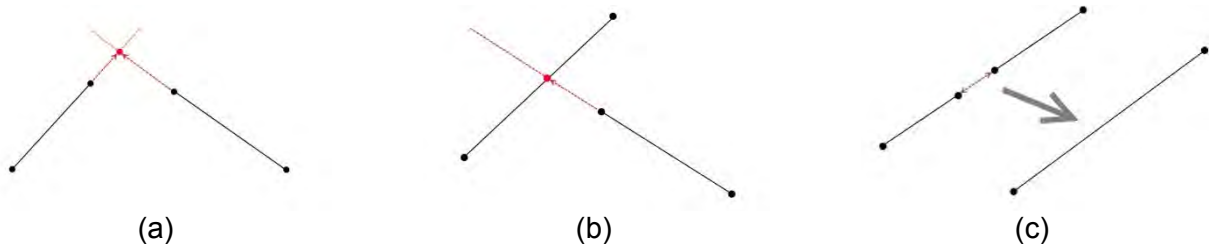


Figure 5: (a) Extending two lines to intersection, (b) extending the extent of a line to another line segment, and (c) merging two nominally aligned sections

These intersections can be matched to corresponding vertices in the as-built plans and used to align the three representations (point cloud, survey and as-built drawings) together for comparison. As can be seen in Figure 4, there are missing elements from the 2D drawings, and not all the information is captured. One reason for this is that the corridor consisted of large glass

sections which could not be captured. Another reason is that, due to the narrow corridors, the average point density was quite sparse in some areas. Additional setups and a finer scanning resolution will help achieve a greater level of detail.

4 AUTOMATED CORRESPONDENCE

In this section, a method for aligning the different datasets is presented utilising a method based on the voting algorithm (Knuth, 1999). The procedure is similar to those employed in calibration and registration techniques for aligning separate overlapping point clouds. The method comprises of several steps, the first is to isolate points that will be used for correspondence between the data sets. In this case, such points will comprise of the surveyed targets and the corner vertices representing the intersection between walls. An initial correspondence between the points is then found by examining the relationship between points in a data set, based on the distances between them. Once the vertices have been match, the different data sets can be transformed into a common coordinate system for comparison.

4.1 Correspondence matrix and Vertex Matching

Different properties, such as geometric invariant properties (Sharp et al., 2002), (e.g. curvature, moment and spherical harmonics invariants), can be used to find correspondence between the data. The only requirement is that the attributes are observable between different data sets. Similarly, the vertices can be chosen using different geometric or spectral properties to extract feature points, such as using mean and Gaussian curvature to define points of local minimums and maximums (Beinat et. al., 2007). In this case, the vertices used are the surveyed targets and the extracted and identified vertices representing corners of walls. The distances between the points within a dataset are used to find correspondence by searching for similar distances occurring in multiple data sets.

The first step, given the dataset $S^{(1)}$, is to calculate a distance matrix $D^{(1)}$ where the i^{th} row and the j^{th} column in the matrix contains the Euclidean distance between the vertices p_i and p_j such that:

$$D^{(k)}_{i,j} = \|p_i - p_j\|, \quad \text{given } (p_i, p_j) \in S^{(k)} \quad (9)$$

The distance values in $D^{(1)}$ and $D^{(2)}$ are compared to determine if the values between two observed points in $D^{(1)}$ and two observed in the $D^{(2)}$ are within a specified tolerance. A voting method is used where vertices between datasets, if the same distance to another vertex is being observed, then the likely correspondence is increased by one. The follow algorithm explains the procedure for producing the correspondence matrix C :

Procedure for Correspondence Matrix

1. Read in set $S^{(1)}$ comprising of n points;
2. Read in set $S^{(2)}$ comprising of m points;
3. Create $D^{(1)}$ and $D^{(2)}$;
4. Create zero n by m matrix C ;
5. For $i1 = 1$ to n
6. For $j1 = 1$ to n
7. For $i2 = 1$ to m
8. For $j2 = 1$ to m
9. If $(D^{(1)}_{i1,j1} - D^{(2)}_{i2,j2}) < \text{tol}$
10. $C_{i1,j1} = C_{i1,j1} + 1$;
11. End
12. End
13. End
14. End
15. End

Algorithm 1: The voting method for determining the correspondence matrix

This produces a correspondence matrix between the plans and the point cloud for the test site as specified in Figure 6. If there is a high value between the i^{th} row and the j^{th} column in C , then this indicates the likelihood of a high correspondence between the point p_i in set $S^{(1)}$, and the point p_j in set $S^{(2)}$. For each p_i in set $S^{(1)}$, the p_j in set $S^{(2)}$ is initially selected as the matching vertex if it has the highest correspondence value in row i^{th} , and the value is significantly large. If a point p_j in set $S^{(2)}$ is associated with more than one point in set $S^{(1)}$, then the point pair with the highest correspondence entry is kept, and the others removed. These matching pairs are then used to determine the transformation parameters used to align the datasets together. The matching pairs are presented in Figure 7. While this example is for 2D data, the process can be easily used for 3D data as well, since the correspondence is a simple distance metric.

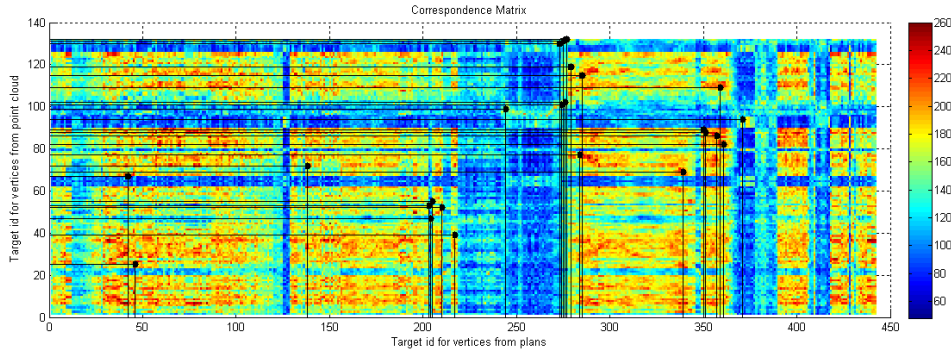


Figure 6: The correspondence matrix for the test site. Red values indicate a high correspondence between the vertices in the point cloud (vertical axes) and the vertices from the as-built plans (horizontal axis).

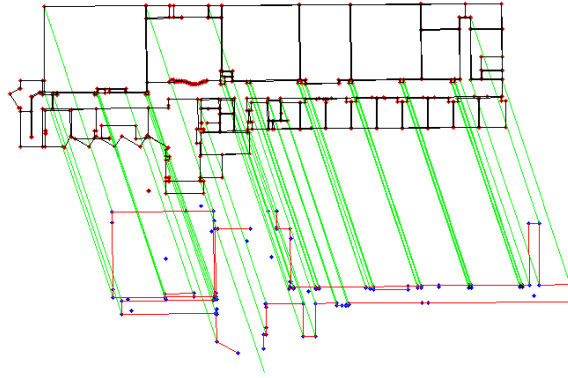


Figure 7: The initially matched vertices selected from the correspondence matrix.

4.2 Alignment and Refining Matching

To find the alignment, the problem is defined as a 2D rigid body transformation. In this case, the parameters will represent a rotation (θ) and two translations (x_t and y_t) such that:

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u_i \\ w_i \end{bmatrix} + \begin{bmatrix} x_t \\ y_t \end{bmatrix} \quad (10)$$

The parameters are solved using least squares to determine the alignment between different data sets. After an initial solution is found, errors in the match vertices are detected as outliers if they are outside a specified confidence interval. Such errors are removed from the set of matched vertices. Similarly, there may be matches between vertices that were not detected by the method outlined in the previous section. These inliers can be added to the set of matched vertices if the differences between the vertices are within the specified confidence interval. The adjustment is carried out in successive iterations as outliers are removed and inliers are added to the set of observed data points until the set of matching vertices remains unchanged. The parameters are then used to transform the datasets into a common domain and the information can be overlaid for comparison, as shown in Figure 8.

While this method presents the case for 2D data, the method can be extended to fully 3D data. In this case the Gauss-Helmert model can be used to define the transformation of the 3D data (Mikhail and Ackermann, 1976). A more rigorous refinement method, such as ICP can be applied to refine the alignment (Bae and Lichti, 2008).



Figure 8: The three data sets overlapping. Black represents the original drawings, blue was captured by the survey, green is the raw point cloud and red is the extract plans from the point cloud.

5 COMPARISON TO PLAN AND CONTROL SURVEY

In the presented case, the plans and point cloud were transformed into the common coordinate system defined by the survey data. For the matching of the plans to the survey and point cloud data, initially 24 vertices were matched with a standard error of 0.0595m. After the outliers and inliers detection method was applied, a total of 56 matches were isolated with a standard error of 0.0227m (a maximum of 0.0977m). The large values in the residuals were observed due to the deviations between the plans and the actual construction. Such large discrepancies can be seen in Figure 8. The point cloud to the survey dataset had much lower residuals, with a final standard error of 0.003m (a maximum of 0.005m). These values are close to the model accuracy and specification of the total station in reflector-less mode.

In general, the raw point cloud data conforms well to the survey data within the specified point uncertainties for the scanner ($\pm 6\text{mm}$). For extracted modelled line segments from the point cloud, the average error was approximately 0.035m (with a maximum of 0.141m). These values were closer to the accuracy of the scanner (average of $\pm 0.002\text{m}$, less than 0.01m) where the wall comprised of a flat section with no features. The larger error values were caused by small features not being identified and removed from the surface segments, such as pictures, shallow recessed doors and windows, electrical outlets, etc. These off-plane features could be removed by increasing sampling density to detect such small features. However, if the observed differences between these features and the wall are less than the positional uncertainty of the scanner points, they may not be identifiable.

The assumption that the walls are orientated to the vertical direction can also cause larger errors. A 3D planar fit to an example wall resulted in a standard deviation of $\pm 0.00197\text{m}$, within scanner accuracy. When tested, the vertical alignment of the wall was out by an angle of $5'11''$. This resulted in a 2D line fit having a standard error of $\pm 0.00380\text{m}$, larger than the true 3D planar fit. In general, the survey data and the point cloud are comparable to each other, but larger differences were seen when compared to the existing plans. This is mainly due to the changes being made after the drawings were submitted. These are easily detected from the point cloud data.

The last concern is for regions of missing data, as previously highlighted. One contributing factor was the narrow corridors creating sparse sampling regions and occlusion of sections of the wall. A higher sampling resolution and an increase number of setups may alleviate this problem. The other factor was that large sections of the walls in the test site are constructed of glass, which resulted in no sampled points being observed in these regions.

6 CONCLUSION

This paper presented a method for the automatic extraction of 2D drawings from point cloud data captured with a Leica ScanStation, and the comparison with a total station survey and as-built plan of the test site. The point coordinates from the total station were considered the ground truth and were utilized to gauge the accuracy of the laser scanner. Where the 2D line was successfully extracted, the laser scanner was illustrated to closely match the accuracies of the total station using the reflector-less mode measuring system. Problems with the automatic extraction procedure occur where there was insufficient point resolution, or when small features could not be clearly isolated. In addition, both 3D point clouds and the information from the total station differed slightly to the original design in some areas of the test-site. This was mainly due to the building structure being built quite different from the original plan.

ACKNOWLEDGEMENTS

This project was in part supported by the Australian Research Council (ARC) Linkage Infrastructure grant (LE0775672) and Prof. M. Bennamoun and his team at the University of Western Australia are appreciated for their hardware support. In addition, this work has been in part supported by Curtin University of Technology and the Cooperative Research Centre for Spatial Information, whose activities are funded by the Australian Commonwealth's Cooperative Research Centres

REFERENCES

- Anderson J. M. and Mikhail E. M. (1998), *Surveying: Theory and Practice*, McGraw-Hill, 1167 pages
- Bae K.-H. and Lichti D. D. (2008), "Automated registration of three dimensional unorganised point clouds", *ISPRS Journal of Photogrammetry and Remote Sensing* 63(1) 36-54
- Beinat A., Crosilla F., Visintini D. and Sepic F (2007), Automatic non parametric procedures for terrestrial laser point clouds processing. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI(part 3/W40B)*, 1–6
- Belton D. and Lichti D. D. (2006), Classification and segmentation of terrestrial laser scanner point clouds using local variance information. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI(part 5)*, 44–49
- Boehler W., Vincent M. B. and Marbs A. (2003), Investigating laser scanner accuracy, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIV(Part 5/C15), 696-701
- Budroni A. and Böhm J. (2009), Toward automatic reconstruction of interiors from laser data. In *Proceedings of 3D-ARCH Trento, Italy*
- Fischler M. A. and Bolles R. C. (1981), Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24 (6), 381–395
- Frank T., Tertois A.-L. and Mallet J.-L. (2007), 3D-reconstruction of complex geological interfaces from irregularly distributed and noisy point data, *Comput. Geosci.*, 33, 932–943
- Froehlich C. and Metternleiter M. (2004), Terrestrial laser scanning – new perspectives in 3D surveying, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI(Part 8/W2), 7 – 13
- Golub G. H. and Loan C. F. V. (1989), *Matrix Computations* (2nd ed.). Baltimore, MD: Johns Hopkins Press.
- Gordon S. J. (2005), *Structural deformation measurement using Terrestrial Laser Scanners*, Curtin University of Technology, Perth, Australia, PhD dissertation, 209 pages
- Johnson R. A. and Wichern D. W. (2002), *Applied Multivariate Statistical Analysis* (5th ed.). New Jersey, USA: Prentice Hall.
- Knuth D. E. (1999), *The Art of Computer Programming*, Addison Wesley, 896 pages
- Leica-Geosystems (2011), <http://www.leica-geosystems.com/> (Accessed August, 2011)
- Lichti D. D., Franke J., Cannell W., and Wheeler K. D. (2005), "The Potential of Terrestrial Laser Scanners for Digital Ground Survey", *Journal of Spatial Science*, Vol. 50(1), pp. 75-89
- Lichti D. D. (2007), Error modeling, calibration and analysis of an AM-CW terrestrial laser scanner system, *The ISPRS journal of Photogrammetry and Remote Sensing*, 5, 307-324
- Mikhail E. M. and Ackermann F. (1976), *Observations and Least-Squares*, IEP, 497 pages

- Miliaresis G. and Kokkas N. (2007), Segmentation and object-based classification for the extraction of the building class from LIDAR DEMs, *Comput. Geosci.*, 33, 1076–1087
- Mólnar G., Pfeifer N., Ressler C., Dorninger P., and Nothegger C. (2009), Range calibration of terrestrial laser scanner with piecewise linear functions, *Photogrammetrie, Fernerkundung, Geoinformation* 2009 (1), 563-576
- Okorn B., Xiong X., Akinci K. and Huber D. (2010), Toward automated modeling of floor plans. In *Proceedings of 3D Data Processing, Visualization and Transmission (3DPVT) 2010*
- Pauly M., Gross M. and Kobbelt L. P. (2002), Efficient simplification of point sampled surfaces. In *VIS '02: Proceedings of the conference on Visualization '02*, Boston, Massachusetts, 163–170. IEEE Computer Society.
- Rabbani T. and van den Heuvel F. (2004), 3D Industrial Reconstruction by Ritting CSG Models to a Combination of Images and Point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXV*.
- Rabbani T., van den Heuvel F. A. and Vosselman G. (2006), Segmentation of point clouds using smoothness constraint. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI(part 5)*, 248–253
- Rüeger J. M. (1990), *Electronic Distance Measurement*, Springer-Verlag, Berlin, Germany
- Schnabel R., Wahl R. and Klein R. (2007), Efficient RANSAC for point-cloud shape detection. *Computer Graphics Forum* 26 (2), 214–226
- Schofield W. and M. Breach M. (2007), “*Engineering Surveying*”, Elsevier, 622 pages
- Shakarji C. M. (1998), Least-squares fitting algorithms of the NIST algorithm testing system. *Journal of Research of the National Institute of Standards and Technology* 103 (6), 633–641.
- Sokkia (2011), <http://www.sokkia.com/> (Accessed August, 2011)
- Tarsha-Kurdi F., Landes T. and Grussenmeyer P. (2007), Hough-transform and extended RANSAC algorithms for automatic detection of 3d building roof planes from lidar data, in *Proceedings of the ISPRS Workshop on Laser Scanning 2007*, 407-412
- Várady T., Benkő P. and Kós G. (1998), Reverse engineering regular objects: simple segmentation and surface fitting procedures. *International Journal of Shape Modelling* 4, 127–141
- Vosselman G. and Dijkman S. (2001), 3D building model reconstruction from point clouds and ground plans. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIV (part 3/W4)*, 37–44

BRIEF BIOGRAPHY OF AUTHORS

David Belton is a research fellow with the CRC for Spatial Information at Curtin University, in the Department of Spatial Sciences. His research interests lie in laser scanning, 3D point cloud processing, error analysis, and optimisation techniques.

Brian Mooney received his BSurv degree from Curtin University in 2009. He is currently employed with Newmount Mining, Australia

Tony Snow is a Senior Lecturer and the undergraduate coordinator for surveying in the Department of Spatial Sciences, at Curtin University. His interests are in surveying principles and applications of emerging techniques and technology.

Kwang-Ho Bae works for Furgo Spatial Solutions as a laser scanning specialist. He was a lecturer in Photogrammetry and Laser Scanning in the Department of Spatial Sciences at Curtin between 2008 and 2011. His research interests are Laser Scanning data analysis such as registration, segmentation, calibration and positional error analysis, Photogrammetry and 3D range cameras.

Assessing Spatial Information Access, Use and Sharing for Catchment Management in Australia

Dev Raj Paudyal, Kevin McDougall, Armando Apan

University of Southern Queensland
Australian Centre for Sustainable Catchments
Toowoomba, Australia

devraj.paudyal@usq.edu.au; kevin.mcdougall@usq.edu.au, armando.apan@usq.edu.au

ABSTRACT

Spatial data plays an important role in many social, environmental, economic and political decisions and is increasingly acknowledged as a national resource essential for sustainable development. One of the potential areas where spatial data can make a positive impact is for improved decision making to support catchment management. Reliable spatial data infrastructure (SDI) is needed to record the environmental, social and economic dimensions of catchment management. By building an appropriate SDI, disparate spatial data can be accessed and utilised to facilitate the exchange and sharing of spatial data between stakeholders across catchment communities. The aim of this paper is to identify the factors/variables contributing to spatial information access, sharing and use across catchment management areas and evaluate the current status of spatial information access, sharing and use among Australian states from a catchment management authority perspective. A survey method was used to collect primary data from 56 regional Natural Resource Management (NRM) bodies responsible for catchment management in Australia. Descriptive statistics method was used to show the similarities and differences among Australian states. The key factors which influence sharing and access to spatial information are also explored. We found there is significant for spatial information access, use and sharing to contribute to SDI development.

KEYWORDS: catchment management, natural resource management, spatial information, spatial data infrastructure

1 INTRODUCTION

Spatial data plays an important role in many social, environmental, economic and political decisions and is increasingly acknowledged as a national resource essential for sustainable development. Catchment management is potentially one area where spatial information can be used for improved planning and decision making process. Traditionally, state government organisations were the custodians of spatial information necessary for the catchment decisions whilst catchment management authorities/regional Natural Resource Management (NRM) bodies were just the users of spatial information. Now, the situation has changed and the regional NRM bodies are also collecting a significant amount of large scale spatial information and state agencies are also interested in gaining access to this spatial data/information. The access, use and sharing of spatial information between state government agencies and regional NRM bodies is therefore becoming more important. This circumstance has opened new opportunities for managing spatial information and developing spatial data infrastructure (SDI) in catchment management sector. The

big drivers are the easily accessible spatial products like Google Earth and hand-held navigation systems which is easily accessible by community and grass-root groups (Folger, 2011).

Spatial data infrastructure is an infrastructure for sharing and use of geospatial information (UN Geospatial Information Working Group, 2007). SDI can facilitate access to the spatial data and services through improving the existing complex and multi-stakeholder decision-making process (Feeney, 2003; McDougall and Rajabifard, 2007). Moreover, it can facilitate (and coordinate) the exchange and sharing of spatial data between stakeholders within the geo-information (GI) community. By building an appropriate SDI, disparate spatial data can be accessed and utilised to facilitate the exchange and sharing of spatial data between stakeholders across catchment communities.

In Australia, all three levels of government play an important role for sustainable catchment outcomes. The Commonwealth Government provides the policy and economic framework and support for intergovernmental coordination (The Parliament of the Commonwealth of Australia, 2002). State and territory governments have the legislated responsibility for natural resource management within their boundaries. Catchment management agencies/regional NRM bodies develop, advise and co-ordinate the implementation of catchment strategies and action plans to achieve sustainable catchment outcomes. Many initiatives have commenced for natural resource information access and sharing at different levels, with spatial information being just one of the components of natural resource information. At the national level, the Australian Natural Resource Information Infrastructure (ANRII) was initiated to facilitate the access and sharing of natural resource information. As the state agencies are the custodians of spatial information, it is important to understand their role with respect to spatial information access, use and sharing. The concept of spatial data infrastructure is already well established, however its effectiveness for the management of spatial information which cross administrative boundaries has been limited (Paudyal et al. 2009). Natural resource information does not understand the artificial jurisdictional boundaries that exist across natural catchments and landscapes.

The aim of this paper is to identify the factors/variables contributing towards spatial information access, sharing and use across catchment management areas and assess the current status of spatial information access, sharing and use among Australian states from a NRM perspective. A survey was distributed to the 56 regional NRM bodies/catchment management authorities (CMAs) across Australia to assess various dimensions of spatial information sharing and access. Further, this paper explores the key factors for developing spatial data infrastructure for better catchment outcomes. The first part of the paper provides an introduction to catchment management in Australia and the importance of spatial information and spatial data infrastructure (SDI) to achieve sustainable catchment outcomes. The primary data collection methods and survey areas are then described. Finally, the paper summarises the findings from questionnaire and comments on key factors which influence spatial data access, use and sharing in catchment management.

2 CATCHMENT MANAGEMENT IN AUSTRALIA

2.1 Historical Perspectives

Australia utilises the whole of catchment approach to the management of natural resources including land and water (The Parliament of the Commonwealth of Australia, 2002). Since European settlement of Australia in 1788, the occupation has resulted in the eutrophication of waterways, extensive land clearing, and rising salinity which have impacted on the natural environment. Catchment management as an approach to the management of Australia's agricultural lands began in the early 1900s (AFFA, 2002). During the 1970's, the catchment management movement collected impetus with numerous community projects being implemented and endorsed by government agencies. By the late 1970s environmental degradation caused by agricultural and other land use practices had been recognised and soil conservation agencies moved towards taking a whole of catchment approach to control erosion and better land management forming group conservation areas (Central Coast Regional Catchment Committee, 1999; The Parliament of the Commonwealth of Australia, 2002).

The recognition of the existence of a significant problem requiring action at a national level led to a series of targeted legislative interventions and activities by national and state governments. The philosophy of catchment management became more comprehensive and state government organisations embraced community-government partnership for sustainable management of natural resources on a catchment basis. The two state agencies particularly concerned with catchment management were those charged with soil conservation and water resources management (Laut et al., 1989). The catchment management approach has enjoyed widespread community support since 1990. The current approach to catchment management relies upon the mix of the three tiers of government and various community initiatives. All three levels of government play an important role; however state government agencies has major role working closely with catchment management authorities for natural resource management.

2.2 Catchment Management Arrangements in States and Territories

Catchment-based management is the approach used for land and water resource management in Australian states and territories (The Parliament of the Commonwealth of Australia, 2002). This management approach is implemented through the creation of partnerships between the different levels of government, community groups, industry groups and academia. All states/territories have some form of catchment management authorities or natural resource management groups under their jurisdiction. There are 56 regional NRM bodies/CMAAs which are responsible for catchment management in Australia as shown in Figure 1.



Figure 1: Regional NRM Bodies/Catchment Management Authorities showing the boundary

State government organisations are primarily responsible for catchment management activities and natural resource management. In each state/territory, there is a principal state government organisation which is responsible for natural resource management. The catchment management authorities are different in their name, corporate structure, catchment management philosophy and function and responsibilities. Some are governed by members of the community and some are established by government. Those which are established by the state government have the statutory responsibilities (Ryan et al., 2010). The jurisdictional models in the states are

either statutory (NSW, SA and VIC) or non-statutory (community based) (WA, QLD, TAS). The Territory models are evolving towards independent boards but are still heavily dependent on Territory government structures and processes. There is also inconsistency between states in the name given to the regional NRM bodies. They are termed catchment management authorities in NSW and VIC, Catchment Councils in WA, Natural Resource Management Boards in South Australia, Regional NRM Groups in Queensland and Regional committees in Tasmania (Table 1).

Table 1: State NRM framework characteristics modified from (Pannell et al., 2008)

State	Title of regional body (number)	Status	Catchment Philosophy	Functions and accountability	Key State Agency
NSW	Catchment Management Authorities (13)	Statutory (CMA act 2003)	TCM	Support property vegetation plan under native vegetation act 2003 Board reports directly to Minister	Department of Environment and Climate Change
VIC	Catchment Management Authorities (10)	Statutory (CALP act 1994)	ICM	Beds, banks and floodplains of rivers Board reports to agency head	Department of Sustainability and Environment
WA	Regional Catchment Groups or Catchment Council (6)	Non statutory	ICM	Functions decided by the groups Report to stakeholders	Department of Agriculture and Food
SA	Regional NRM Boards (8)	Statutory	ICM	Water allocation planning, pests and weeds, soil conservation & biodiversity Board reports to Minister	Department of Land Water and Biodiversity Conservation
QLD	Regional Committees, Groups or Organisations (14)	Non statutory	ICM	Functions decided by the groups Report to shareholders and stakeholders	Department of Environment and Resource Management
TAS	Regional NRM Committees (3)	Statutory	ICM	Required to nominate member to NRM Council and report annually to parliament	Department of Primary Industries and Water

2.3 Importance of Spatial Information for Catchment Management in Australia

Spatial data underpins decision-making for many disciplines (Clinton, 1994; Gore, 1998; Longley et al., 1999; Rajabifard et al., 2003) including catchment management. It necessitates the integration of spatial data from different sources with varying scales, quality and currency to

facilitate these catchment management decisions. However, the institutional arrangements for catchment management do not easily align with the SDI development perspectives as multiple stakeholders work to achieve multiple goals with government organisations, often guiding many catchment decisions.

SDI can facilitate access to the spatial data and services through improving the existing complex and multi-stakeholder decision-making process (Feeney, 2003; McDougall and Rajabifard, 2007). Moreover, it can facilitate (and coordinate) the exchange and sharing of spatial data between stakeholders within the geo-information (GI) community. A preliminary step toward in achieving decision-making for catchment management has been the increasing recognition of the role of SDI contributes towards generating knowledge, identifying problems, proposing alternatives and defining future courses of action (Paudyal and McDougall, 2008). In recent years, many countries have spent considerable resources on developing their own National Spatial Data Infrastructure (NSDI) to manage and utilise their spatial data assets more efficiently, reduce the costs of data production and eliminate duplication of data acquisition efforts (Masser, 2005; Rajabifard et al., 2003).

These initiatives have been traditionally highly government dominated and generally based on the administrative/political hierarchy of the country's government. However, catchment management issues cut across political-administrative boundaries and do not follow the rules of political-administrative hierarchies. Hence, there is a need to consider SDI development across catchments differently, particularly understanding the spatial information access, sharing and use for catchment management in state jurisdictions. This understanding from community perspectives will assist in building catchment spatial data infrastructure (SDI) and hence it's contribution to national spatial data infrastructure (NSDI).

3 SURVEY OF NRM BODIES

The survey of NRM bodies was conducted with all 56 regional NRM bodies responsible for catchment management in Australia. There are 14 regional NRM bodies in Queensland (QLD), 13 in New South Wales (NSW), eight in Victoria (VIC), eight in South Australia (SA), six in Western Australia (WA), three in Tasmania (TAS), one in the Australian Capital Territory (ACT) and one is in Northern Territory (NT). The local government boundaries straddle catchment boundary however do not straddle state government boundaries so, we compared the spatial information access, use and sharing between regional NRM bodies and state government organisations. The objective of the questionnaire was to assess the current status of spatial information access, use and sharing for catchment decisions among Australian states.

3.1 Data Collection and Analysis Methods

All 56 regional NRM bodies provided a response to the survey. The questionnaire survey was undertaken between June 2010 and September 2010. Before the questionnaire was finalised, the draft questionnaire was checked with QMDC, (Queensland Murray Darling Committee) one of the regional NRM bodies in Queensland. The quality of online questionnaire was also tested before distribution. The questionnaire was distributed in two stages and targeted for two groups of regional NRM bodies. First, the questionnaires were distributed to regional NRM bodies which belong to Murray Darling Basin Authority (MDBA) and later distributed to rest of the NRM bodies. The feedback and experience from the first stage was used to assist in the second stage of the survey and hence the high response rate.

The statistical analysis of the survey results was undertaken in the SPSS Statistics package. The profile of respondents has been tabulated in Figure 2, with the largest group of respondents being GIS officers, with other respondents including staff who were directly or indirectly involved with spatial information management or the GIS operations of the NRM body.

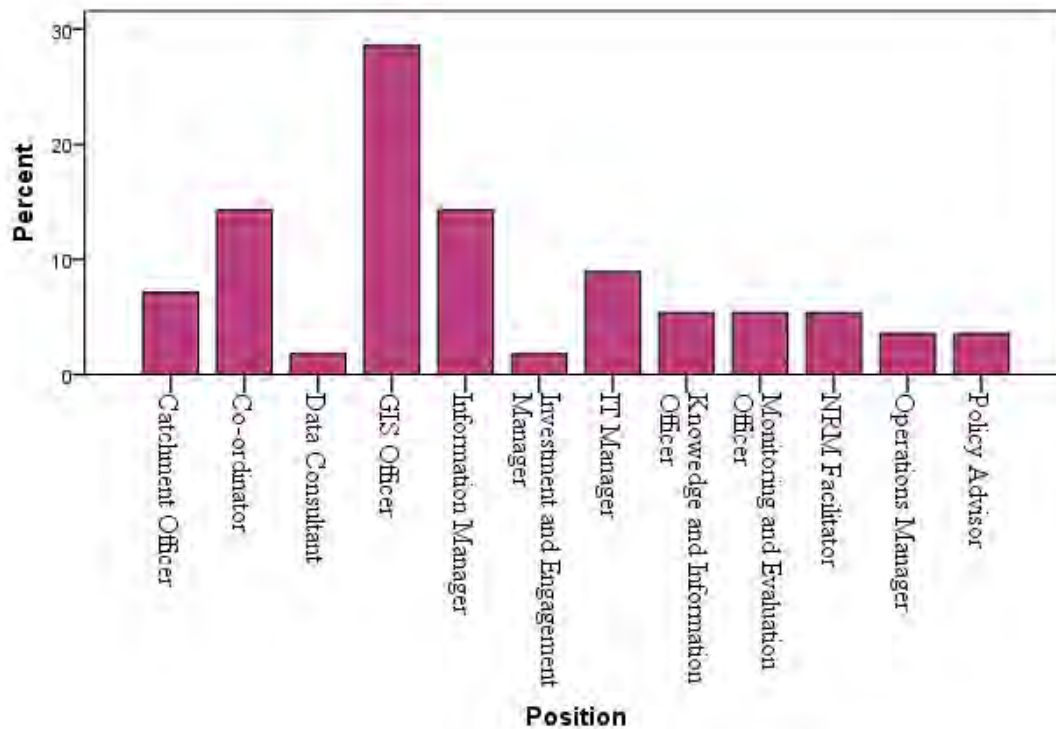


Figure 2: Profile of respondents (by position)

3.2 Areas and Factors Contributing Spatial Data Infrastructure Development

Though there are different views regarding definition and the components of spatial data infrastructure, the common terminology they have used is that spatial data infrastructure (SDI) is the infrastructure to facilitate spatial information access, use and sharing (UN Geospatial Information Working Group, 2007; Rajabifard et al., 2003; EU, 2006). These three areas (spatial information access, use and sharing) were selected and the variables contributing towards these areas were identified to explore the status of each of spatial information use, access and sharing in catchment management sectors.

The spatial information access among regional NRM bodies in Australian States was assessed using variables such as the ease of access, restriction, impact of restriction, affordability of current pricing and spatial information access medium. The spatial information use among regional NRM bodies was assessed using variables such as the type of organisation, spatial information used by staff, GIS maturity and GIS activities and spatial information receiving medium.

The spatial information sharing and networking activities were assessed with various variables including collaborative arrangement, networking, use of open source models and social media, spatial policy, funding sources, importance of spatial data provider, spatial information integration issues and data sharing agreement arrangement.

The three key areas and the contributing factors are provided in Table 2.

Table 2: Areas and contributing factors for spatial data infrastructure (SDI) development

Areas	Contributing factors
Spatial Information Access	Ease of access, restriction, impact of restriction, affordability of current pricing, spatial information access medium
Spatial Information Use	Type of organisation, spatial information used by staff, GIS maturity, GIS activities, spatial information receiving medium
Spatial Information Sharing	Collaborative arrangement, Networking, use of open source models and social media, spatial policy, cost of spatial data, funding sources, importance of spatial data provider, spatial information integration issues, data sharing agreement

4 RESULTS

Spatial Information Access

In respect to the ease of accessing obtaining spatial data, about half (48%) the organisations responded that it was moderately easy to access (Figure 3).

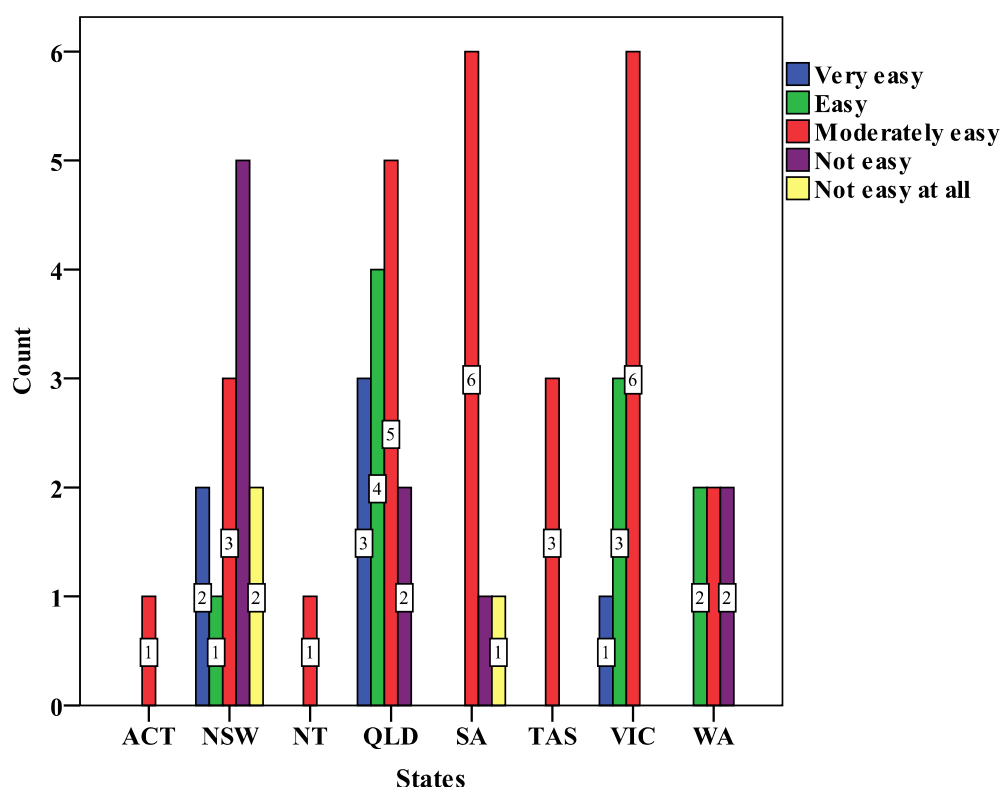


Figure 3: Ease of access to obtain spatial data from spatial data provider

However, in NSW, the majority of organisations advised that it was more difficult to access spatial information. It was possible that this was due to the restriction placed by the spatial data provider on their use of spatial data. The 6 out of 13 NSW NRM bodies advised that there were always restrictions on the use of spatial information provided by spatial data providers which limited their ability to undertake GIS activities. However, the majority of regional NRM bodies located in QLD, SA, VIC and NT advised that the restriction was not limiting their GIS activities. The majority of regional NRM bodies of TAS and WA advised that the restriction impacted on their ability to undertake GIS activities. In respect to pricing of spatial data, the majority (60%) of the organisations advised that the pricing of spatial information is affordable for their organisation and the most accepted pricing arrangement was the cost of transferring data. The majority of regional NRM bodies received their spatial information using ICT technology and digital media which also impacted spatial information access.

Spatial Information Use

The regional NRM bodies also produce spatial information which provides a strong base to develop spatial data infrastructure (SDI) at catchment level. The majority of regional NRM bodies including QLD (93%), NSW (85%), VIC (80%), WA (67%), NT (100%), and TAS (100%) identified themselves as both spatial information providers and users. However, half of regional NRM bodies in SA and regional NRM body of ACT identified themselves as spatial information users only. With respect to the use of spatial information by regional NRM bodies' staff, 40-60% of the total staff in NSW, QLD, VIC and WA used spatial information for catchment management activities. However, 60-80% of total staff in TAS and SA used spatial information for catchment management activities. The regional NRM bodies in Queensland, New South Wales and Victoria identified themselves as mature GIS organisations using spatial information for 5-10 years. However, the regional NRM bodies of Tasmania and Australian Capital Territory indicated limited experience of their spatial information use as shown in Figure 4.

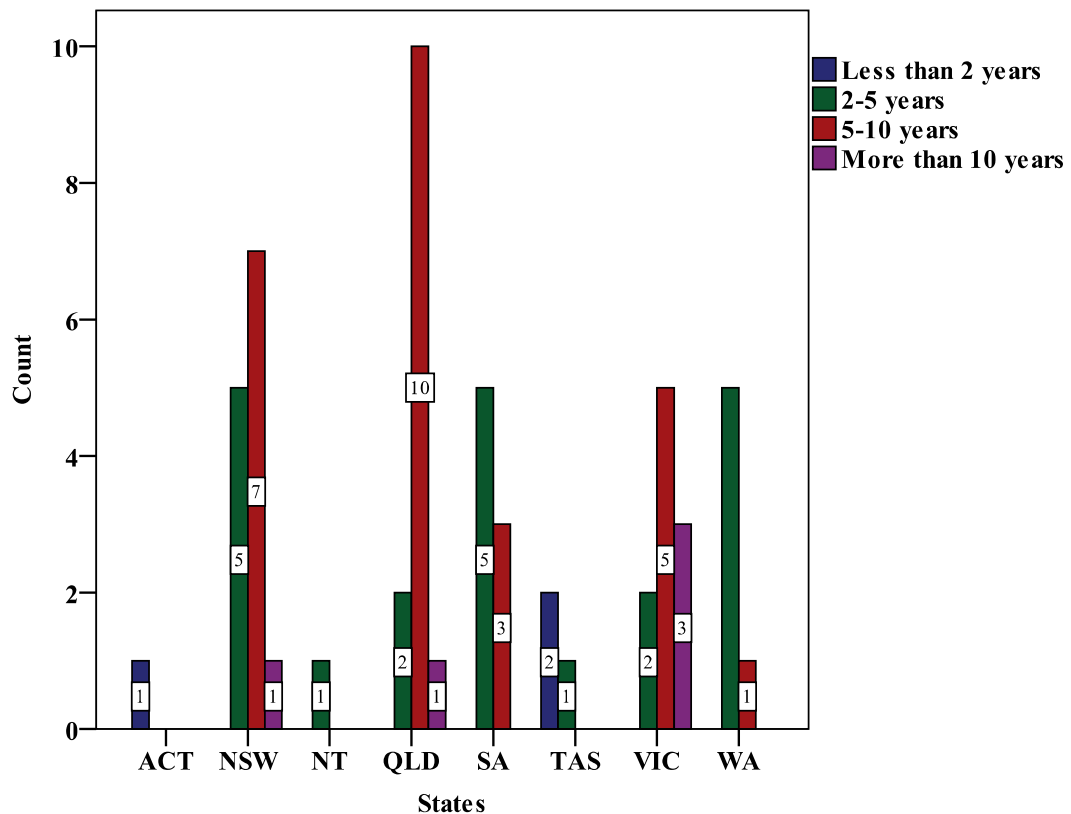


Figure 4: Length of time using GIS

The majority of regional NRM bodies use a combination of in-house and outsourcing GIS activities. The majority of regional NRM bodies/CMAAs in NSW, QLD, SA, VIC, NT and TAS are outsourcing some of their GIS activities. However, about half of the regional NRM bodies in WA are undertaking GIS activities completely in-house. The Western Australian Land Information System (WALIS) appears to have significantly influenced on WA regional NRM bodies to build capacity in-house GIS capacity. ACT regional NRM body have limited resources capacity to undertake GIS activities and outsource all GIS activities.

Spatial Information Sharing and Networking

The collaborative arrangements with other organisations with respect to the exchange of resources, skills and technology were examined. The majority (83%) of the regional NRM bodies/CMAAs advised that they have a collaborative arrangement with other organisations. Regional NRM bodies of QLD identified themselves having the most collaborative arrangements (Figure 5). There are some variations in the areas of collaboration among the regional NRM bodies. It was found that data sharing and spatial information management were the main areas of collaboration in most of the states. However, in TAS, the main area of collaboration was knowledge transfer. The next most important area of collaboration in most of the states was knowledge transfer.

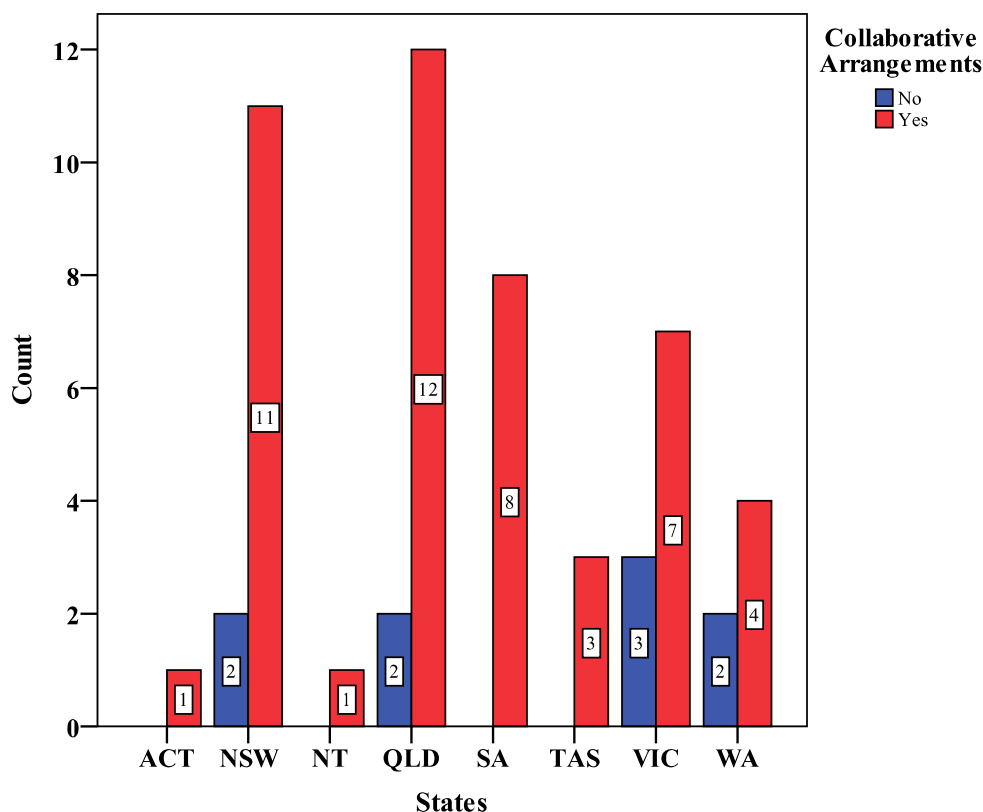


Figure 5: Collaborative Arrangements

The main partners for this collaboration and networking activities were state government organisations. Community organisations including other regional NRM bodies/CMA's were the second most common.

The majority (95%) of the regional NRM bodies/CMA's advised that they were aware of freely available/accessible spatial products e.g. Google Maps, OpenStreetMap, Wikimapia for their work needs. However, the utilisation of these products for catchment management activities is infrequent. The majority of regional NRM bodies of ACT, SA and WA were in the favour of using social media and open models for spatial information management. However, the other states were neutral on this issue. There is a growing utilisation of these new open models and social media for spatial information sharing and exchange at community level. However, there are some of good examples of the use this technology for information exchange and knowledge sharing. Due to the security, privacy and confidentiality, the regional NRM bodies are not yet very comfortable using these products.

5 DISCUSSION

The regional NRM bodies are not only spatial information users; they are also spatial information providers. The main users of spatial information generated or value-added by regional NRM bodies/CMA's are the community organisations like Landcare, Watercare, Birdwatch and land owners and indigenous groups. Government organisations, private sectors and academia research institutions are less frequently utilising spatial information managed by regional NRM bodies/CMA's. However, there is significant interest in state government organisations to have access of community owned data. This has opened a new perspective on management of spatial information and development of spatial data infrastructure (SDI) in the natural resource management sector. Spatial information use, access and sharing has significance for SDI development in the catchment management sector.

Regional NRM bodies obtain spatial information from state government agencies and the access mechanism varies amongst Australian states. However, the majority of organisations responded that the spatial access mechanism was relatively straight forward. The access policy of state government organisation has impact on spatial information access. For example, the catchment management authorities in NSW had expressed some concerns with access to information. This may be due to the access policy and the restriction placed by NSW government. In NSW, all CMAs have access to similar state wide data and the main source is from an Enterprise database held at Parramatta. In respect to pricing of spatial data, the majority of regional NRM bodies advised that the pricing of spatial information was affordable for their organisation. However, the pricing do not affect the access of spatial information.

There are limited variations among Australian states regarding spatial information use. With respect to the use of spatial information by regional NRM bodies' staff, the majority of staff are aware of spatial information and using it. The regional NRM bodies of NSW, QLD, VIC and WA are quite mature in comparison to other states. The half of WA regional NRM bodies undertaking their GIS activities in-house. There is great inspiration of state government organisation to build in-house GIS capacity of regional NRM bodies.

Information sharing, technology sharing, knowledge sharing, human resource sharing were the main areas of collaboration for spatial information management. The data sharing and spatial information management were the main areas of collaboration. Another emerging area for collaboration in NRM sector is knowledge sharing. The knowledge management and knowledge sharing is well-practised in NRM sector. The spatial data infrastructure could be one the component of spatial knowledge infrastructure and the main collaboration partners could be state government agencies and community organisations.

Based on the above explanation, there is potential for spatial information access, use and sharing to contribute to SDI development. By building the spatial data infrastructure, disparate spatial information can be accessed and used to facilitate the sharing of spatial information between stakeholders across catchment communities and it will support sustainable catchment outcomes. Figure 6 indicates how spatial information access, use and sharing facilitate SDI development and can improve natural resource management outcomes.

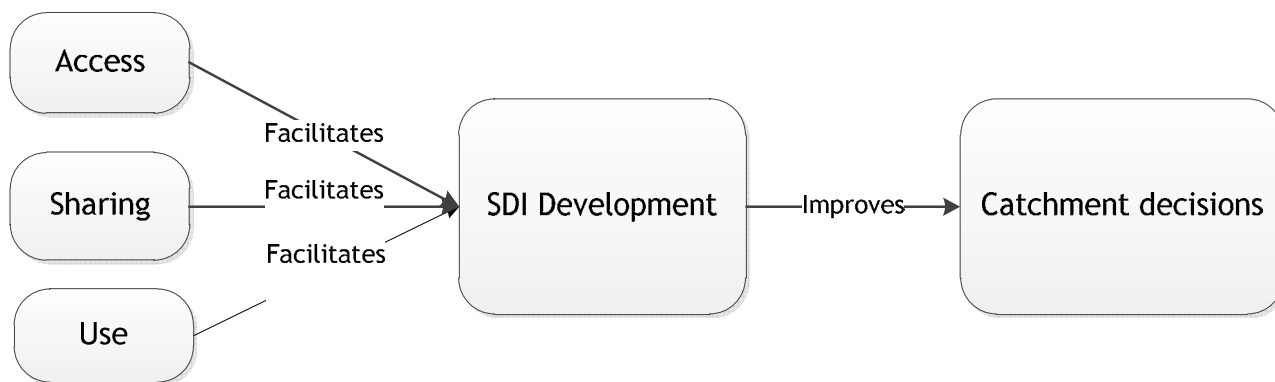


Figure 6: Contribution of access, sharing and use for SDI development

The spatial information access, sharing and use facilitates spatial data infrastructure (SDI) development in natural resource management sector which improves the catchment decisions as shown in figure 6. Thus, the SDI development strategy can be influenced by the policies on spatial information access, use and sharing.

6 CONCLUDING REMARKS

The traditional thinking about custodians of spatial information essential for catchment management has changed. The regional NRM bodies collect a significant amount of spatial information from large areas and there is significant interest from state government organisations

in gaining access to this spatial data/information. This circumstance has opened a new way of collaboration between regional NRM bodies and state government organisations for better NRM outcomes. The access, use and sharing of spatial information between state government agencies and regional NRM bodies is therefore becoming more important for the development of spatial data infrastructure (SDI) at catchment level. Particularly, the SDI development strategy can be influenced by the policies on spatial information access, use and sharing.

The access policy of state government organisations has impact on spatial information access the state. The catchment management authorities in NSW had expressed some concerns with access to spatial information in comparison to other states. This may be due to the access policy and restriction placed by NSW government. Regarding the spatial information use, the regional NRM bodies of NSW, QLD, VIC and WA are quite mature in comparison to other states. The state government organisations appear to have significantly influenced regional NRM bodies to build capacity for spatial information management. The Western Australian Land Information System (WALIS) is a very good example of GIS capacity used to build WA regional NRM bodies.

We found the data sharing and spatial information management is the main areas of collaboration between state government organisations and regional NRM bodies. Another emerging area of collaboration is knowledge sharing. There is also growing utilisation of open models and social media for spatial information management and knowledge sharing at community level. The spatial data infrastructure could be the one of the component of spatial knowledge infrastructure. However, this is a new area and how spatial knowledge infrastructure and spatial data infrastructure could be inter-linked for better catchment outcomes need to be further explored.

REFERENCES

- AFFA. (2002). *Scientific Advice to the Parliament of the Commonwealth of Australia*: The Department of Agriculture, Fisheries and Forestry, Australia.
- Central Coast Regional Catchment Committee. (1999), *A submission to the Standing Committee's inquiry into Catchment Management*. Available: www.aph.gov.au/house/committee/environg/cminq/sub115-e.pdf, accessed 15 September, 2011.
- Clinton, W. J. (1994), *Coordinating geographic data acquisition and access to the National Spatial Data Infrastructure*. Executive Order 12096. Available: <http://govinfo.library.unt.edu/npr/library/direct/orders/20fa.html>, accessed 15 September 2011.
- EU. (2006). *Assessing the Impacts of Spatial Data Infrastructure*. Brussels: Joint Research Centre EU.
- Feeney, M.-E. F. (2003), 'SDIs and Decision Support' In *Developing Spatial Data Infrastructures From concept to reality* (Eds, Williamson, I. P., Rajabifard, A. and Feeney, M.-E. F.) Taylor & Francis, London, UK pp. 195-210.
- Folger, P. (2011), *Geospatial Information and Geographic Information Systems (GIS): An Overview for Congress*. Congressional Research Service, R41825, USA
Available: <http://www.fas.org/sqp/crs/misc/R41825.pdf>, accessed 15 September, 2011.
- Gore, A. (1998), *The Digital Earth: understanding our planet in the 21st century*. Digital Earth, Wikipedia. Available: http://en.wikipedia.org/wiki/Digital_Earth, accessed 15 September 2011.
- Laut, P., Taplin, B. J. and CSIRO. Division of Water Resources. (1989), *Catchment management in Australia in the 1980s*, CSIRO Institute of Natural Resources and Environment Division of Water Resources, Canberra.

- Longley, P. A., Goodchild, M. F., D.J., M. and Rhind, D. W. E. (Eds.) (1999) *Geographic Information Systems*, John Wiley & Sons, New York.
- Masser, I. (Ed.) (2005) *GIS Worlds, Creating Spatial Data Infrastructures*, ESRI Press, Redlands, California.
- McDougall, K. and Rajabifard, A. (Eds.) (2007) *A Mixed-Method Approach for Evaluating Spatial Data Sharing Partnerships for Spatial Data Infrastructure Development*, ESRI Press Redlands, California, USA.
- Pannell, D. J., Ridley, A., Seymour, E., Regan, P. and Gale, G. (2008) In *SIF3 Working Paper 0809, CRC for Plant-Based Management of Dryland Salinity* Perth.
- Paudyal, D. R. and McDougall, K. (2008), *Building Spatial Data Infrastructure to Support Sustainable Catchment Management*. Paper presented at the Queensland Spatial Conference 2008, 17-19 July, 2008 Gold Coast, Queensland.
- Rajabifard, A., Feeney, M.-E. F. and Williamson, I. (2003), 'Spatial Data Infrastructures: Concept, Nature and SDI Hierarchy' In *Developing Spatial Data Infrastructures: From Concept to Reality* (Eds, Williamson, I., Rajabifard, A. and Feeney, M.-E. F.) Taylor & Francis, London, UK, pp. 17-40.
- Ryan, S., Broderick, K., Sneddon, Y. and Andrews, K. (Eds.) (2010) *Australia's NRM Governance System: Foundations and Principles for Meeting Future Challenges*, Australian Regional NRM Chairs, Canberra.
- The Parliament of the Commonwealth of Australia. (2002). *Co-ordinating Catchment Management: Report on the Inquiry into Catchment Management*. House of Representatives Standing Committee on Environmental Heritage.
- UN Geospatial Information Working Group. (2007). *Strategy for Developing and Implementing a UN Spatial Data Infrastructure*. New York City: United Nations.

Shove over Gen Y: Gen Z is almost here

Gita Pupedis; Chris Bellman

School of Mathematical and Geospatial Sciences
RMIT University
124 Latrobe Street Melbourne
Australia

gita.pupedis@rmit.edu.au; chris.bellman@rmit.edu.au

ABSTRACT

In recent years, the surveying and spatial science industry has grappled with the issue of attracting more qualified staff into the industry. One aspect of the problem is declining numbers of students studying tertiary programs in these disciplines. Many initiatives have resulted from the recognition of this problem, some based on comprehensive market analysis and others on instinct and experience. The marketing gurus made it clear that Generation Y was the target and the message needed to resonate with them. But the game has changed and the gurus now tell us that Gen Y is moving on and Gen Z (digital natives; net generation) is taking over.

So what was the message and how does it need to change, if at all? This paper presents views from the perspective and perception of students who have recently chosen to enrol in a program in spatial science or surveying. It draws on earlier work that evaluated student motivations for selecting programs in the spatial sciences and reports on detailed qualitative studies based on student focus groups. The paper compares students from Gen Y with those on the cusp of the Gen Y/Gen Z divide and seeks to identify the differences and common elements in their motivations, desires and needs.

KEYWORDS: Education, student recruitment, skills shortage, generational change

1 INTRODUCTION

There is growing awareness within the Spatial Science community that we continue to face a shortage of skilled labour in the industry. This seems to be a worldwide phenomenon and has created pressure on tertiary programs, in an environment of increasing competition for resources. Industry has responded well to the challenges this has presented and see recruitment as an industry problem, not just a problem for universities. However, nothing stands still for long and as fast as the Spatial Sciences industry adapts to the marketing needs of the time, the target for this marketing is also evolving and changing. Demographers tell us that the next generation (Gen Z) source their information and are influenced in their decision making quite differently to previous generations. The next generation of spatial scientists will come from this group and we need to promote careers in our industry in a way that resonates with them.

2 BACKGROUND

In 2006, the Australian Spatial Information Education and Research Association (ASIERA), through a "Heads of Department" meeting, raised concerns with both industry and government about student recruitment into Geospatial Science programs in Australia (McDougall et al. (2006)). There was clear evidence of a decline in enrolments in these programs, and a growing gap between supply and demand for labour in the discipline. This was coupled with a change in the university sector to a more demand driven model and a greater emphasis on minimising costs, often achieved through restructures that amalgamated small departments into larger units

(McDougall et al. (2006)). ASIERA was suggesting an impending crisis that might well threaten the viability of most higher education programs in the geospatial sciences. Some responses to this are addressed in previous work (Pupedis & Bellman, 2009).

Since then, much has changed in the university sector. There have been further amalgamations of university departments, to the point where very few specialist geospatial science departments have survived as independent entities. Nearly all are now part of larger departments or schools but most have managed to maintain the critical mass required to offer specific programs in geospatial science.

Further changes to the university sector will emerge in 2012, with the first year of the implementation of a demand driven model for student funding of universities, based on recommendations from the Review of Australian Higher Education (Bradley et. al., 2008). It is still unclear exactly how this model will be implemented but this could further threaten the viability of programs and courses in the spatial sciences if demand in these fields of study is not sufficient.

Much has changed in the marketplace as well. Spatial has become ubiquitous through the use of location based services, Web 2.0, smart phones and other advances in technology. The spatial industry has also begun responding to the challenge of recruiting and training staff through a number of initiatives designed to promote the career opportunities offered in the spatial sciences.

2.1 Initiatives

The Spatial Education Advisory Committee, operating under the auspices of the Australian and New Zealand Land Information Council, commissioned several studies into the issue of supply and demand of labour in the sector. These studies were underpinned by workshops held in 2006 and 2008 with stakeholders from government, a variety of industry sectors and academia. A clear and tangible product of this was a workforce plan that, among other things, identified a deep and long-term gap between supply and demand of skilled labour in the sector (SEAC, 2007). They also commissioned a study of university programs and courses available (SEAC, 2008) that identified considerable inconsistencies in the way programs were promoted and marketed and a lack of clarity in the pathways a prospective student may take to enter a career in geospatial science.

A recent series of discussion papers have emerged from The Skilled Workforce Development Initiative of the Qld Surveying & Spatial Industry (Lyons & Davies, 2011) addressing similar issues.

In other initiatives at a jurisdictional level, the Surveying Taskforce in Victoria was formed to promote careers in Surveying. While this was a collaborative effort involving professional bodies, the private sector, government and academia, a significant and important characteristic of the Taskforce was that they sought professional marketing advice. This required significant start-up capital and all participants contributed generously to this initiative. The result was a comprehensive study of the market place, a review of the “image” of surveying from a very wide perspective and the development of a “brand” that underpinned the development of the marketing message. One key aspect of work was identifying the important factors influencing Gen Y decision makers. The desire for opportunity, lack of constraints and lifestyle were all messages that came through clearly and these were embodied in the marketing program with slogans such as “... a Life Without Limits ...”. Modern, crisp graphics that were simple, distinctive and easily recognisable were also an important part of the campaign, along with a new website and a video. This created a consistent brand that was identifiable in all forms of marketing material and presentations. This model has now been adopted in South Australia and negotiations are continuing with representatives from New South Wales for the use of the marketing collateral.

Another initiative to emerge in the last few years is Destination Spatial (www.destinationpatial.org), a joint activity of the Surveying and Spatial Sciences Institute (SSSI) and the Spatial Industry Business Association (SIBA). This is a website aimed at raising the awareness of careers in Surveying and Spatial Science across the community.

2.2 Other Factors

Reviews of subjects studied at Victorian Certificate of Education (VCE) level have shown a considerable decline in the study of Geography (Hincks, 2008) and higher level Mathematics (Leung, 2007). In Victoria, this trend has not abated in recent years and the proportion of students studying these subjects continues to decline, as shown in Figure 1.

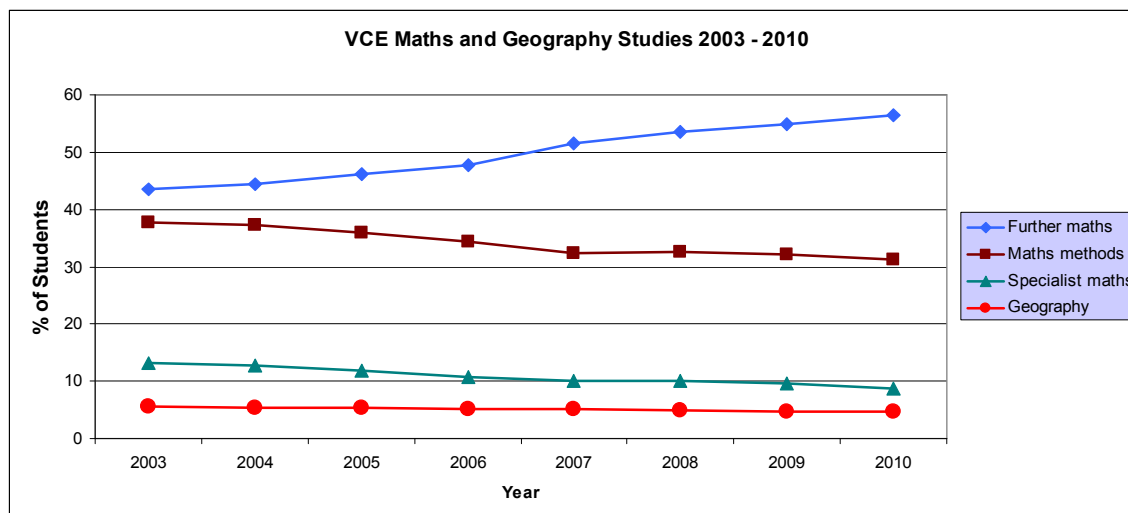


Figure 1 – VCE enrolments in Mathematics and Geography

The issues driving subject choice in the final years of high school are complex but are not helped by the broad range of choices offered to students. In 2010, VCE students had 102 VCE subjects and an additional 50 VCAL (Victorian Certificate of Applied Learning) subjects from which to choose their five or six study units.

In addition, in Victoria, female students accounted for 54% of VCE completions in 2010, yet make up less than 10% of university entrants into Spatial Science programs.

Recruitment of students into the Spatial Sciences is further compounded by the low public profile of our industry and the fact that about 150 other university programs required an advanced level of mathematics as a prerequisite for entry into their programs.

There is an abundance of studies that identify the problems faced in recruiting students and significant competition for what appears to be a diminishing pool of students. It is therefore critical that we understand the marketplace both now and into the future to ensure our marketing efforts are well targeted and well received.

3 GENERATIONAL CHARACTERISTICS

All of the marketing effort in the spatial sciences in recent years has, quite rightly, addressed the needs of Generation Y. Although this generation will continue to enter our programs for some years to come, the younger members of this group are exhibiting some characteristics of the next generation. Also, the older members of the Generation Z group are already at the point of considering subject choices that will impact on their pathways to university. It is of critical importance that we understand the needs of these groups and evolve our marketing messages and methods to meet the changes in the generational demographic.

3.1 Generation Y

This generation (those born approximately between the years 1980 and 1994) are generally considered to be highly technologically literate.

Raines (2003) describes the characteristics of this group as: sociable, optimistic, talented, well-educated, collaborative, open-minded, influential and achievement oriented. Howe and

Strauss (2000) also point out that they: gravitate toward group activity, believe “it’s cool to be smart”, are focused on grades and performance, are busy with extracurricular activities, identify with parents’ values, are respectful of social conventions and institutions, have a fascination for new technologies and are racially and ethnically diverse.

McCrindle and Beard’s (2007) studies of consumers and the factors that influence their decision making summarise these key points about Generation Y:

- They are socially connected and their decisions are likely to be influenced by their friends.
- They like things to be fun and entertaining, yet they are hard to engage.
- They like cool and socially desirable things, but not “try-hards”.
- They want life-enhancing experiences.
- They want new and innovative things, yet are unable to articulate exactly what will meet their needs.

It is also suggested that Generation Y, when compared to previous generations, exhibits different characteristics in terms of requirements and expectations of their learning environment. Oblinger & Oblinger (2005) observe that today’s learners are: digital, connected, experiential, immediate and social and that their learning preferences tend to be: teams or peer-to-peer, by engagement and experience, visual and kinesthetic.

An understanding of Generation Y’s characteristics, influences and learning preferences is not only important in terms of marketing, but also with respect to their orientation and transition activities and teaching and learning. At the same time it is necessary to be aware that although the generation starting university now is likely to be more representative of these characteristics (Skene et al., 2007), making general assumptions about them, particularly with respect to their digital background, risks overlooking the diversity within this group of students (Kennedy et al., 2008).

3.2 Generation Z

We are rapidly approaching the time that those entering higher education from high school, will be those from Generation Z – people born from 1995 onwards (McCrindle, 2010). Are there any differences between the Gen Ys and the Gen Zs? Is this really a different generation of learners and do we need to market ourselves to them in a different way?

Generation Z is the world’s first 21st Century generation. They are true digital natives and have not known a life without mobile phones, personal computers, the internet and gaming systems (Mueller, 2011). They are both comfortable with and dependent on technology. Socialising is not necessarily about physically being in the same location as your friends, it can be on-line, engaged in a collaborative pursuit (McAneny, 2010). At the same time, this generation is not tied down to one desktop computer in one location. They carry their technology around with them and are often perpetually hooked up, often doing multiple tasks at the same time (Mueller, 2010). McAneny (2010), states that this dependence on technology and need for instant gratification has a darker side, termed by some mental health experts as “acquired attention deficit disorder”. As people are so accustomed to a constant stream of digital stimulation, they feel bored when it is absent.

McCrindle (2011) states that Generation Z is also growing up faster, is in education earlier and is exposed to marketing younger. While many are vying for their attention, Gen Zs consumer decisions are most often made through social networks (Ross, 2010). However, it is not necessarily using the popular Facebook or Twitter, whose user’s average age is in the mid to late thirties (Ross, 2010).

4 STUDENT PERCEPTIONS SURVEY 2011

In May of 2011, 21 first year students and 30 final year students in Geospatial Science at RMIT University completed a student perception survey, covering a broad range of topics. These included: their work experience prior to commencing study, how they became aware of the geospatial sciences, what marketing activities and other factors had affected their decision making, the use of social media for marketing and their perceptions of what Generation they belong to. 71% of the first year students who participated were under the age of 21 and 97% of the final year

students were over the age of 21, with nearly all being in Generation Y. A few mature age students belong to Generation X.

Eight students volunteered to participate in a follow-up focus group, to discuss the topics raised in greater depth.

Although there were more similarities than dissimilarities between the two year groups, some differences were quite striking.

- **Work experience before commencing study**

76% of first year students had undertaken work experience or worked in the geospatial sciences (mainly surveying) before commencing their studies, compared to 40% of the final year students. Although a significant number from both groups had had spent a year or more working in the field, the numbers that had undertaken a week or two of work experience was much higher in the 1st year group.

- **First awareness of careers in the geospatial sciences**

A significant majority of both groups first became aware of these careers in the final three years of high school (62% of first years and 63% of final years). A significant number from both groups found out about the career possibilities while working.

- **First awareness of program of study**

There were significant differences between the year groups.

Year 1		Year 4	
Source	No. of students	Source	No. of students
VTAC Guide	6	Careers advisor/teacher	9
Parents	4	I knew someone in the program	2
RMIT website	3	Friends	2
Friends	3	RMIT website	2
Career/study expo	3	School careers night	2
RMIT Open Day	2	RMIT Open Day	2
Other/no response	0	Other/no response	11

Table 1 – Source of First Awareness of Geospatial Sciences

The most significant difference is the apparent decline of the influence of careers advisors/teachers. During the focus group discussions, it became evident that careers advisors were a very positive influence when they understood what Geospatial Science was about. One student pointed out that their careers advisor had been to a Geospatial Science seminar at RMIT and had been very articulate about how the program would suit them. The vast majority of students had sought assistance from advisors, but had not found them useful, with comments such as:

"Mine was not much help"

"I don't know what she got paid for!"

This indicates that a raised awareness of our industry by careers advisors is a critical factor.

- **Influences when selecting program of study**

Year 1				Year 4			
Influence	1st	2nd	3rd	Influence	1st	2nd	3rd
Family	9	1	1	Family	6	1	2
Career or employment prospects	3	1	1	Career or employment prospects	4	3	1
Work experience	3	2	0	Careers advisor/teacher	4	3	0
RMIT Open Day	2	1	3	Career/study expo	3	2	0
Members of the profession	1	4	1	Course brochure	2	1	3
Careers advisor/teacher	1	2	3	Members of the profession	2	1	2
Other/no response	2	10	12	Other/no response	9	19	22

Table 2 – Influences in selecting a program of study

Family and career prospects were the major influence for both year levels, with members of the profession and careers advisors ranking in the top six influences for both.

- **Most frequently visited websites**

(students were asked to list the three websites they visit most frequently).

Year 1		Year 4	
Website	% of students who visit	Website	% of students who visit
Google	86%	Facebook	77%
Facebook	71%	Google	60%
YouTube	43%	RMIT	53%
RMIT	38%	Hotmail	27%
Hotmail	33%	YouTube	10%
Other	29%	Other	63%

Table 3 – Most frequently visited websites

Interestingly, there is a smaller percentage of younger students using Facebook frequently than older students and there is a large increase in the use of YouTube by younger students.

- **Use of Facebook or other social media by Geospatial Science at RMIT**

Year 1		Year 4	
Yes, you should	62%	Yes, you should	57%
No, you shouldn't	14%	No, you shouldn't	33%

Table 4 – Use of Social Media (Note: Totals do not add up to 100% as many students did not answer this question.)

Many of the first year students pointed out that an unofficial Geospatial Facebook page already exists. Final year students thought it could be useful for promoting seminars, careers expos and for advertising job opportunities.

A number of final year students thought that a social media site would be inappropriate, as it isn't seen to be professional.

- **Use of industry websites and other promotional events**

Year 1		Year 4	
Activity	% of students	Activity	% of students
"A life without limits" website	57%	"A life without limits" website	20%
"A life without limits" DVD	29%	"A life without limits" DVD	7%
Surveying taskforce at career events	29%	Surveying taskforce at career events	10%
Surveying taskforce industry presentation at school	-	Surveying taskforce industry presentation at school	3%
Destination Spatial website	-	Destination Spatial website	-
Geospatial Revolution project (Penn State University)	-	Geospatial Revolution project (Penn State University)	3%
None of these	38%	None of these	63%

Table 5 – Percentage of students that knew of promotional activities in Geospatial Science

These figures reflect increasing penetration of both the "A life without limits" website and contact with potential students through careers events.

- **New sources of information that would help in choosing program of study**

The main information source identified by both groups of students (15 across both years) was YouTube. Typical comments were:

"YouTube video – a quick outline of various surveying career options"

"YouTube videos of the subjects that students do in the course".

This topic was further discussed in the focus group session and numerous ideas were suggested by students across year levels. These include:

- *Show application of Geospatial Science using short, sharp examples.*
- *Geospatial Science and surveying lack glamour and prestige. You should compare the career to other high profile careers (such as dentists/doctors) and highlight the advantages.*
- *You should show people in our field using cool technology like laser scanning, GIS, GPS, thermal scanning.*
- *A 90 second clip of a year in the life of a surveyor (with lots of shots of the different types of work they do), or a clip about how essential surveyors are (without them, no roads get built, no tunnels get dug etc).*

5 DISCUSSION

Our survey is a very limited attempt to understand how the student marketplace is evolving. It is too early to identify significant trends or changes and the study group does not contain a cohort that is truly Gen Z. However, some interesting observations can be made from the survey responses and from some of the literature available on the changing generations.

All of the students in focus groups were quite suspicious of the generational labels. While they could identify some different trends, they didn't feel there is a significant difference in the way they do things or in their attitudes.

Both Gen Y and Gen Z argue that they don't like being marketed to directly and are quite cynical about many marketing attempts. In Gen Z, this trend is perhaps more pronounced and they are much more influenced by their peers and a loose collective view. Perhaps paradoxically, many of the students were aware of the promotional activities that have been undertaken and did not appear cynical about these. This was particularly apparent with the year 1 cohort.

Social media is also not the dominant influence that we thought it might be. While many students use social media, they do not see this as part of their professional or academic life and

have little desire to have formal social media sites available. They would prefer an informal approach, with social media sites emerging and disappearing as their individual needs change.

However, there was a clear divide on the use of YouTube and the Year 1 students placed much greater emphasis on its use, influence and importance. Students suggested the industry should make greater use of short, sharp, entertaining videos on mediums such as YouTube, as it is an increasingly popular source of both entertainment and information. Marketing campaigns designed around “viral” strategies seem to fit well with the idea of being influenced by peers.

Perceptions of the prestige of the industry were also raised. Students felt that along with the low awareness of the industry, prior to entering the program, there was little evidence of prestigious and successful careers. None of the students understood the diversity of career paths available and in some cases, claimed this was the greatest surprise that they encountered. Again, they pointed to video presentations as a good method of conveying this message.

While the use of technology was an important means of communication, many students were also influenced by more traditional means of promotion, in particular, work experience with someone in the industry and school visits by professionals, students and academics. These raise awareness of both the area of study and the careers available in the discipline, not just to students but also to teachers.

6 CONCLUSION

The marketplace for recruiting students into tertiary programs continues to evolve. As we move into the next generation of recruits, our marketing message and methods need to adapt to the changing requirements of the marketplace.

Opportunities now exist for the clever use of technology in raising awareness of our industry and the careers available in the spatial sciences. However, this is also an increasingly crowded space and attention spans are short. While the efforts made to date have had an impact, there is little room for complacency. New approaches will need to sit beside more traditional methods to expand the multi-pronged marketing strategies that seem to be necessary in today's world.

There is no doubt that a continued and concerted effort will be required in the future to ensure the message about careers and opportunities in the spatial sciences is heard loud and clear.

REFERENCES

- Bradley, D., Noonan, P., Nugent, H., & Scales, B. (2008). *Review of Australian Higher Education Final Report*. accessed August, 2011, from http://www.deewr.gov.au/HigherEducation/Review/Documents/PDF/Higher%20Education%20Review_one%20document_02.pdf.
- Hincks, P. (2008) Geography Issues, Victorian Curriculum and Assessment Authority (VCAA). Melbourne.
- Howe, N., Strauss, W. (2000) *Millennials Rising: The Next Great Generation*, New York, NY, Vintage Books.
- Kennedy, G. E., Judd, T.S., Churchward, A., Gray, Krause, K. (2008) First year students' experiences with technology: Are they really digital natives? *Australasian Journal of Educational Technology*, 24, 108-122.
- Leung, C., C. (2007) Employers warn on VCE maths drop-off. *The Age*. Melbourne.
- Lyons, K. & Davies K. (2011) Working Papers of the Skilled Workforce Development Initiative of the QLD Surveying and Spatial Industry. accessed July 2011, <http://www.spatialinfoservices.com.au/reports.html>

- McAneny, B. (2010) Generation Z and learning, accessed July 2011, <http://www.prelude-team.com/blog/2010/08/24/generation-z-and-learning>
- McCrindle, M. (2010) The New Generations at Work: Attracting, Retaining & Training Generation Y, accessed July, 2011, <http://www.mccrindle.com.au/resources/whitepapers/The-New-Generations-at-Work.pdf>
- McCrindle, M., & Beard, M. (2007) Seriously Cool: Marketing, Communicating & Engaging with the Diverse Generations. accessed July, 2011, http://www.mccrindle.com.au/wp_pdf/SeriouslyCool.pdf
- McCrindle, M. (2011) Why are Generation Z Unique? accessed, July 2011, http://www.generationz.com.au/about_why.html
- McDougall, K., Williamson I., Bellman, C. & Rizos, C. (2006) Challenges Facing Spatial Information and Geomatics Education in the Higher Education Sector. *Combined 5th Trans Tasman Survey Conference & 2nd Queensland Spatial Industry Conference 2006*. Cairns.
- Mueller, J. (2011) Generation Z Characteristics, accessed July 2011, http://www.ehow.com/info_8056211_generation-characteristics.html
- Oblinger, D., Oblinger, J. (2005). Is it Age or IT: First Steps Toward Understanding the Net Generation. In O. Oblinger. D., J. (Ed.), *Educating the Net Generation*: EDUCAUSE.
- Pupedis, G. & Bellman, C. (2009) Recruitment and Retention of Students in the Geospatial Sciences, in *Proceedings of the Surveying & Spatial Sciences Institute Biennial International Conference*, Ostendorf, B., Baldock, P., Bruce, D., Burdett, M. and P. Corcoran (ed.), Surveying & Spatial Sciences Institute, Adelaide, Australia, pp. 829-843 (2009 Surveying & Spatial Sciences Institute Biennial International Conference)
- Raines, C. (2003) Managing Millennials. *Connecting Generations: The Sourcebook*. Manlo Park, California, Crisp Publications.
- Ross, E. (2010) How to connect to Generation Z, accessed July 2011, <http://www.smartcompany.com.au/internet/20100422-how-to-connect-to-generation-z.html>
- Spatial Education Advisory Committee (2007) Spatial Information Industry Workforce Plan, accessed, July 2011, <http://www.crcsi.ecampus.com.au/mod/resource/view.php?id=674>
- Spatial Education Advisory Committee (2008), Career Pathways and Building Skills for the Spatial Information Industry, accessed July, 2011, <http://www.crcsi.ecampus.com.au/mod/resource/view.php?id=1077>
- Skene, J., Cluett, L., Hogan, J. (2007) Engaging Gen Y students at university: what web tools do they have, how do they use them and what do they want? *10th Pacific Rim First Year in Higher Education Conference 2007*. Brisbane.

Mapping Cycling Pathways and Route Selection Using GIS and GPS

Matthew Huntley, Xiaoye Liu, Kevin McDougall, Peter Gibbings

Faculty of Engineering and Surveying
The University of Southern Queensland
Toowoomba, QLD 4350, Australia

Email: matthew.huntley89@gmail.com
xiaoye.liu@usq.edu.au
kevin.mcdougall@usq.edu.au
peter.gibbings@usq.edu.au

ABSTRACT

Increasing demand for crude oil is a major factor affecting the price of fuel for motor vehicles, which also can influence public transport prices. Future transport could prove to be costly so the use of bicycles could become more common for local travel. When riding a bicycle from one place to another, the shortest route is not always the easiest one; terrain has more of an influence on how tired a person gets than the distance travelled. A commuter cycling to work may know that shortest route, however they may not realise that this route requires the outlay of more energy than a longer route.

This study investigated the easiest route from the University of Southern Queensland (USQ) to the central business district (CBD) in Toowoomba in terms of energy used. This was done by collecting GPS data of cycling lanes and pedestrian paths for six different routes. The data was uploaded to GIS software, then processed and analysed to ensure the data was suitable for calculating energy usage. An energy equation was then developed to calculate the energy expenditure for a cyclist riding up and down terrain. Using this equation the energy for the six different routes were calculated to find the route that used the least amount of energy and to also find the route that was the most energy efficient from USQ to the CBD.

KEYWORDS: Cycling, pathways, route selection, energy, GIS and GPS

1 INTRODUCTION

With the price of fuel rising and the cost to operate motorised vehicles becoming more expensive, the use of bicycles will become a more common method of travel. Active transport is now being encouraged to reduce the use of automobiles, save energy and to reduce pollution. Cycling has been around for over century. Over time it has advanced dramatically, today being used not only as a means of transportation but also for recreation. It is used for a range of applications such as competitive cycling, downhill mountain biking and many other sports. A study done on bicycle commuter routes in Guelo, Canada, found that most commuters travelled up to 5.5 kilometres from home to work. Most commuters didn't divert much from the shortest path and were found to use major road routes (Aultman-Hall *et al.*, 1997).

There is extensive planning from local government to design cycling and pedestrian networks in towns and cities. These networks can be made up of cycling and pedestrian pathways or they can be made up of bike lanes. Pathways consist of paths that go through parks and open spaces whereas bike lanes are designated on streets and roads. Councils try to create routes that are

continuous, reasonably direct, functional (serving a variety of destinations), part of a network, safe, and passing through parks and open spaces where possible (IBI Group, 2000). The pathways also need to have smooth surfaces, be well marked and be easily accessible.

Creating a network that incorporates all the above factors, takes considerable time and planning (ANJEC, 2004). A way of encouraging cycling and making it more convenient is to have facilities such as bike racks and lockers, amenities (showers and toilets) at desired locations and along the way and lastly for it to be aesthetically pleasing. To do this, town planners must find the most popular routes and develop these facilities (Leigh *et al.*, 2009). One way in which the local city councils can find the most common routes of the commuters and recreational cyclists is to survey the community (Rissel *et al.*, 2010). This gives the council insight into what paths people like to travel to work or what paths are available. This can help the council to plan or update cycling pathway or road networks. The authorities must also ensure that the networks provide sufficient access to the most important areas such as: schools, businesses, shopping centres, churches, libraries and other community facilities. Studying all the factors together with how people use, need and feel about cycling pathways can help increase the use of cycling.

Previous research on analysis of cycling routes using GIS shows some of the processes used when analysing data (Aultman-Hall *et al.*, 1997). The data was collected using surveys of the community and digitalising all the data into a GIS. Using the surveys filled out by the commuters, a number of possible routes were established. Attributes were then given to the different routes for further analysis. The attributes used included; what intersections had traffic lights, major intersections, gradient, speed limits, bridges & railway crossings.

Shumowsky (2005) discussed the process of creating a map of the pedestrian and bike pathways using GIS software. This map would also be used for planning for future pathways. A key element in the pathway design was the traffic volumes. Keeping away from roads that were too busy was important so that bikers and pedestrians would both feel and be safe when enjoying the routes (Shumowsky, 2005). This is information that needs to be noted in the analysis of routes. Routes that followed creeks and rivers were also a key element in the planning. Using all the information from the GIS system and the public input the best routes around the city were determined.

Gray and Bunker (2005) used GIS as a tool to analyse data and select common routes of commuters. The sources of data used to do this include public transport route and stop data, public transport timetables, street, park, topography, bikeway data, public domain aerial photographs, population data, employment data, and infrastructure and services benchmarks. By analysing all this data, the best routes to Brisbane CBD (central business district), the unsafe bicycle routes and the best locations for new routes could be determined (Gray and Bunker, 2005).

With the active transport networks on a Geographic Information System (GIS), users can view all the different pathways and can also see the inventory and facilities (toilets and parking) of routes (Gray and Bunker, 2005). Viewers can decide whether they want to take a route that has bike paths or whether they want to use bike lanes on a street. Today, most methods of choosing cycling routes are determined on the shortest and safest routes. This however may be a poor way of determining a route as these routes may mean that more energy is expended.

There are many different factors that must be taken into account when forming an equation for energy usage for push bikes. Some of these factors are rolling resistance, wind, rider's weight, bicycle weight, speed, distance, time, revolutions per minute (RPM), slope, and crank size and angle (Abbiss *et al.*, 2009). It is difficult to incorporate all these factors into one formula. Another important aspect of the formula is the output (Atkinson *et al.*, 2003). There are a number of different units that energy can be displayed as. It can be displayed as: watts, kilojoules, calories, and horsepower. Depending on which output is required this determines what the equation will be like. There are three main factors that need to be considered when creating a formula for calculating energy. The first is what output is required, second, what the equation will be based on (time or distance), and lastly what other factors will be taken into account. From previous research most equations calculate the energy on flat terrain rather than slopes and are based on time rather than distance. This makes it very difficult when trying to formulate an equation including slope and distance. Al-Haboubi (1999) discussed the information for working out the caloric energy used when cycling. What Al-Haboubi (1999) does not look at however is applying this formula to slope and distance. His equation only includes speed, technique, RPM and cyclist weight.

This paper presents a way to map cycling pathways when considering the energy usage and route selection from Toowoomba campus of the University of Southern Queensland (USQ) to the CBD of Toowoomba using GIS and GPS. It aims to assist people in identifying the easiest way from USQ to CBD and accordingly create greater incentive to use a bicycle.

2 MATERIALS AND METHODS

2.1 Study area

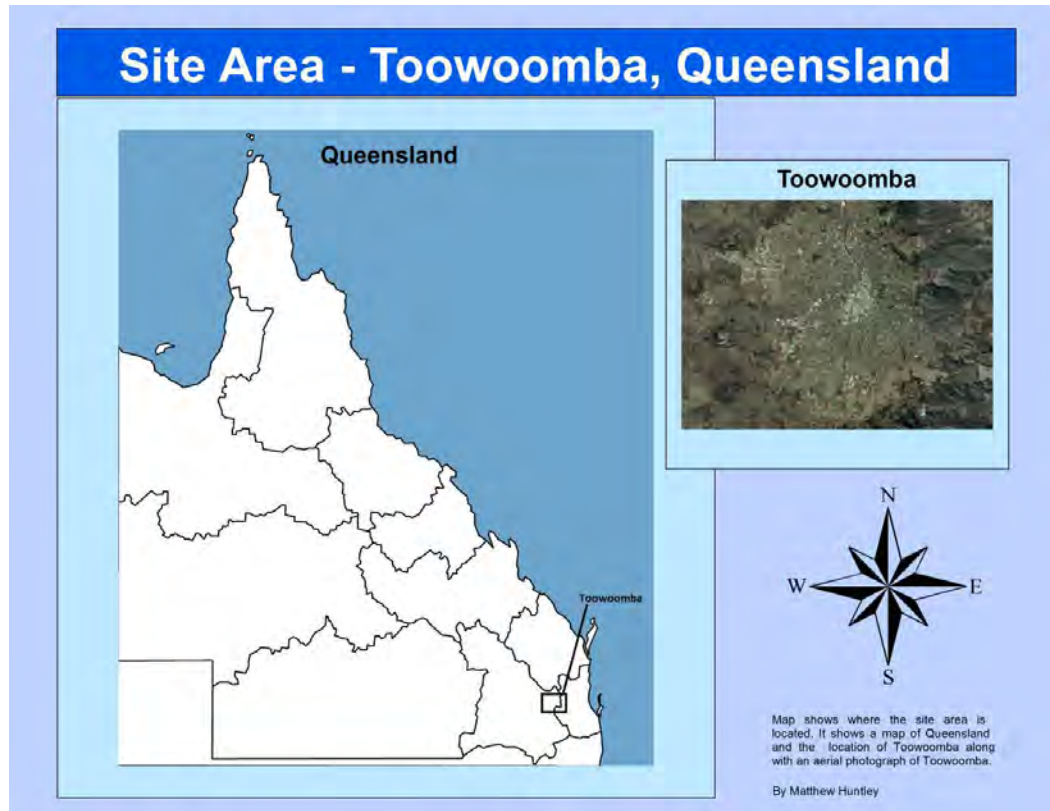


Figure 1: Study area

The study area is in the region of Toowoomba Regional Council, Queensland, Australia, covering the area of the Toowoomba city. The Toowoomba City is the regional centre of the Darling Downs, located approximately 130km out of Brisbane, Queensland, Australia (ANRA, 2009). The city sits on the crest of the Great Dividing Range, around 700 metres above sea level. The majority of the city is west of the divide.

2.2 Data

A 5-metre resolution DEM was generated using aerial photography in the study area. The GPS data on six different routes were collected from the USQ to the office of the Toowoomba Regional Council in the CBD. A frame was made up for a bicycle, with a magnet to be sited on top of the black plate, which the GPS receiver could then be mounted onto (Figure 2). With the method of collection, accuracy, and the equipment decided on, the logging interval was then chosen. A logging interval of five metres was picked for most of the terrain but in some areas with steep slopes and at corners, a logging interval of one second would be used. The optimum time for collecting the field data was worked out. This was done by looking at factors such as: satellite configuration, peak traffic times and weather forecasts. With the ideal times for data collection in terms of satellites these times must then be cross referenced with traffic times. In Toowoomba the peak traffic times are from about 8:00am till around 9:30am and from 2:30pm till 4:30pm. When

taking these statistics into account it was clear that the best time to collect the data would be either from 6:00am till 8:00am or from 9:30am till 12:30am.

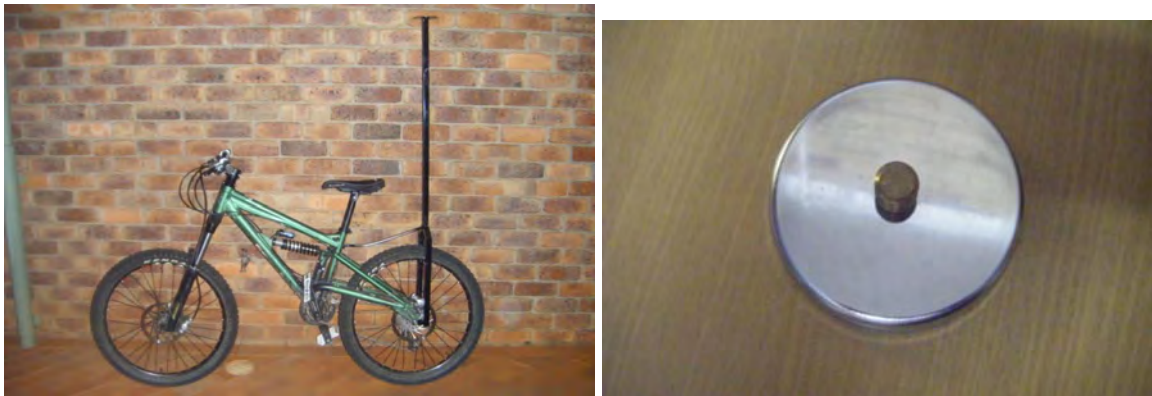


Figure 2: Bike frame and magnet

2.3 Methods

The GPS data was post-processed in Trimble Pathfinder Office software. Once the data was post-processed and the accuracy was checked, the data was then exported as a shape file so it could be used in ArcGIS software. The DEM was used to create a slope map in the study area. The slope was classified as seven classes: 0 – 1%, 1 – 3%, 3 – 5%, 5 – 7%, 7 – 10%, 10 – 15%, and 15 – 30%.

Equation one for the ascent, while equation two is used for the decent, which are modified version of Mueller (2010):

$$Calories = (S_c \times (BW \times 2.2) \times 60) \times \left(\frac{D}{S}\right) + (0.22 \times (HD \times 3.28)) \quad (1)$$

$$Calories = (S_c \times (BW \times 2.2) \times 60) \times \left(\frac{D}{S}\right) - (0.1 \times (HD \times 3.28)) \quad (2)$$

where, S is average speed; S_c is speed coefficient; BW represents the body weight (kg) of the cyclist; D is the distance (km); and HD represents the height difference.

The equation 1 can be divided into two sections. The first section computes the energy use for riding a bike on level terrain. The second section is adding energy to compensate for travelling uphill because it takes more energy to ride up a hill. The equation 2 is similar to the first and it calculates the energy usage of riding downhill, while the second part of this equation subtracts energy to compensate for riding downhill.

Calculations were carried out using the above equations in the ArcMap and a Microsoft Excel spreadsheet. The routes were divided up into a number of small sections to make it easier to calculate. The information gathered using ArcMap was collected by using the slope classification layer and the collected GPS point data. Where a slope started a point would be selected and the point number and its height were noted. Where there was another change in grade, another point was selected and again the number and height was noted. A distance between these two points was then calculated using the inquire function and all this information was stored in the spreadsheet.

When the height difference was negative the terrain was going downhill so the respected formula was used. When the height difference was positive the terrain was uphill and therefore the uphill equation was applied. This was done in turn for each route until the results were found. These results were then tested by applying the above equations for the start and finish point for all the different distances.

3 RESULTS AND DISCUSSION

The total energy used and the length of the six routes are shown in Table 1. The results show that route 4 used the least amount of energy to get from USQ to the CBD while route 5 used the most energy. Route 1 is the shortest route while route 5 is the longest one. These three routes are analysed in details.

Table1 Energy used and length of the six routes

Route	Energy (calories)	Distance (kilometres)
1	144.97	6.031
2	149.99	6.147
3	148.46	6.234
4	144.19	6.151
5	154.67	6.434
6	152.40	6.297

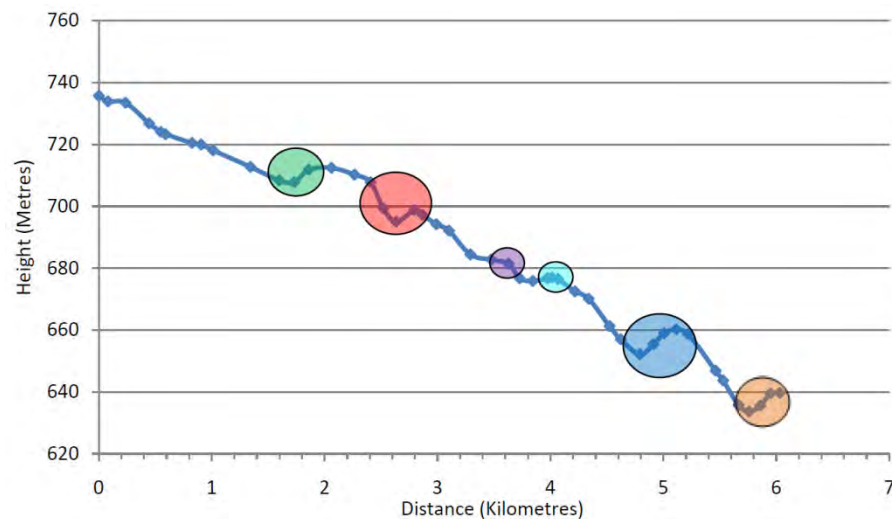


Figure 3: Route1 Profile

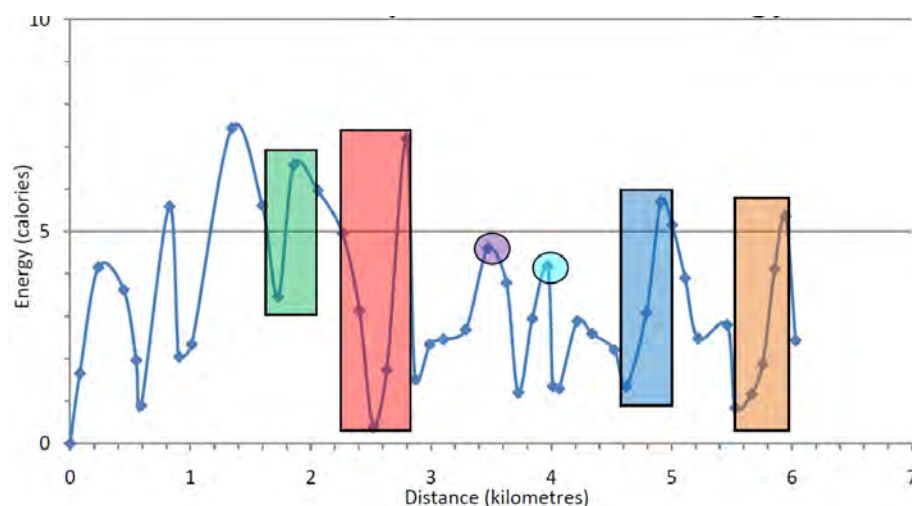


Figure 4: Route 1 Acquired Distance versus Energy

Figures 3 and 4 depict route 1's profile and acquired distance versus energy. Analysis of these figures can determine where most of the energy was used. It can be seen that in the first kilometre that the energy fluctuates as the terrain slopes downhill. These fluctuations are due to the different distances between points. By the second kilometre the energy has built up to a small

peak which can be seen in the green box. The green circle in the profile graph shows that there is a small rise in the terrain and this is what is causing this energy peak. The larger fluctuations before these points are caused by large distances between points which causes a build-up of energy. From two kilometres to three, the energy drops all of a sudden and this is because of the steep slope in the terrain, which can be seen in the Figure 3. These areas of interest are highlighted with the red rectangles and circles in each graph. From this point there is a large rise in energy use and this is because of the uphill section in the terrain which again can be seen in the graph. After the three kilometre mark the energy usage decreases and fluctuates again. This is because there are two small rises or ridges in the terrain which is causing the rises in energy usage. These peaks can be seen in the purple and cyan circles. As the graph reaches the five kilometre mark there is another spike in the energy seen in blue. This is because the terrain has reached a low point and there is a steep slope, which can be seen in the blue circle on the profile graph. The energy graph then shows a slow decline in the energy used from the five kilometre mark to approximately the five and a half kilometre mark. In this section the profile shows that the terrain is downhill before it rises from the five and a half kilometre mark to the finishing point. This small rise is reflected in the energy graph with another spike in the energy used which can be seen in orange.

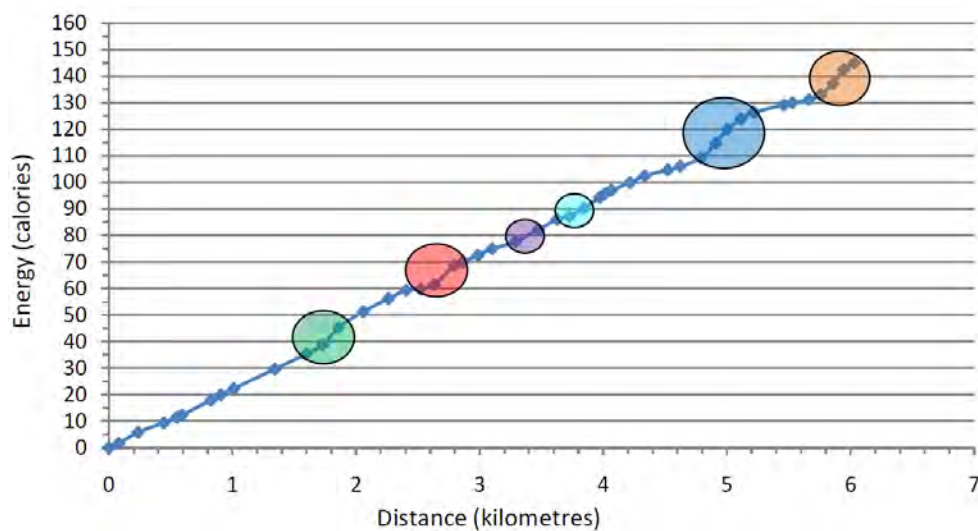


Figure 5: Route 1 Acquired Energy versus Acquired Distance

Figure 5 shows the acquired energy versus acquired distance for route 1, and the effects of the terrain on energy usage can be seen. The coloured circles on this graph correspond to those in the previous graphs for route 1. Where the line gets steeper is where there is a rise in the terrain and where the line flattens out or dips the terrain declines. The red and the blue circles are areas in the route that require large portions of energy to complete and in turn having a major influence in the final outcome.

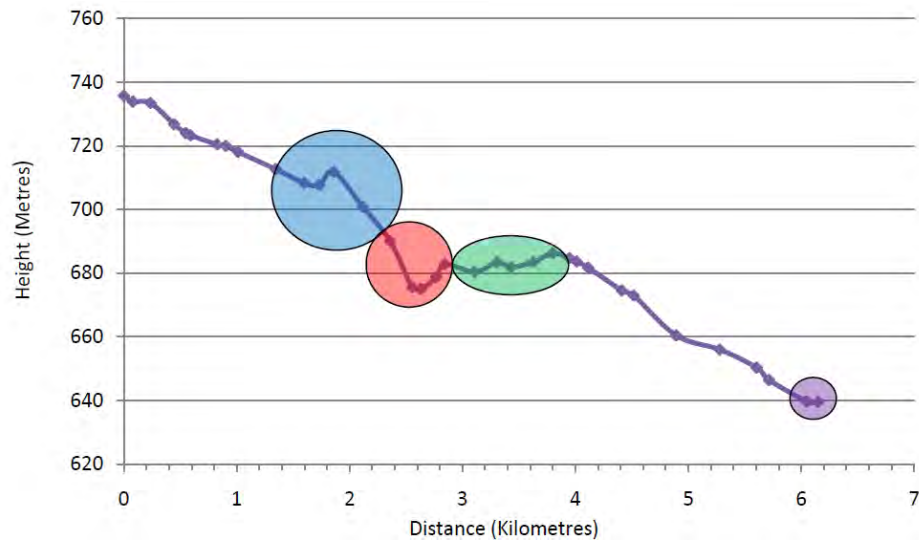


Figure 6: Route 4 Profile

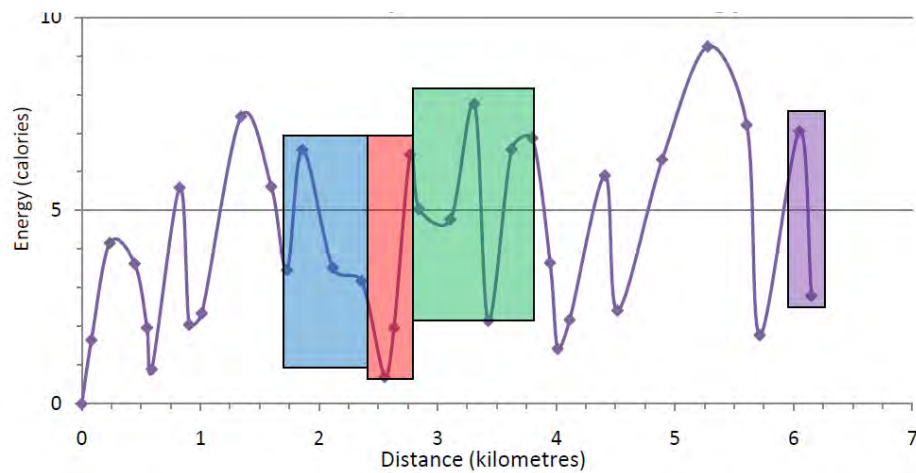


Figure 7: Route 4 Acquired Distance versus Energy

Route 4 was the route that used the least amount of energy from USQ to the CBD. It used a total of 144.19 calories over the route and the total distance was 6.151 km. It can be seen from Figures 6 and 7 there are no large peaks in energy used. From the start point to the first kilometre mark there are fluctuations in the energy graph (Figure 7). This is not due to the terrain but it is due to longer distances used between grades resulting in more energy being used. Continuing along the route from the one kilometre mark the graph continues to fluctuate until it reaches the 2.6km mark. When looking at Figure 6, it can be seen in the blue circle that the terrain rises slightly followed by a long steep drop. When viewing the energy graph the blue rectangle represents the same area mentioned in the blue circle of the terrain graph. The energy rises up and then falls steeply before flattening out briefly. As seen in the red rectangle in Figure 7 the energy drops steeply again and then rises again very steeply. This is reflected in the red circle in Figure 6 where the terrain drops right down to a low and then rises again. From here the green circle in Figure 6 shows a number of small peaks and ridges. The corresponding colour in the energy usage graph shows that there are rises and falls in energy where the terrain is alternating. From 3.75 km mark the terrain declines at a continuous rate, flattening out in some areas. The energy graph however shows fluctuations and again this is mainly due to the different distances used between grades, however areas where the terrain flattens out contributes to these fluctuations. This continues until the 6 km mark before the terrain levels out which can be seen in the purple circle. The energy graph again reflects this in the corresponding colour with a spike in the energy used.

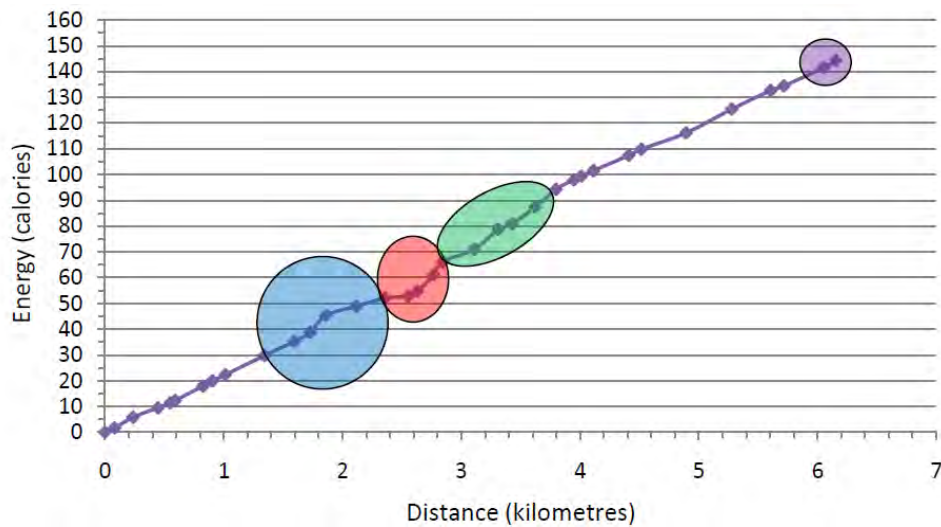


Figure 8: Route 4 Acquired Energy versus Acquired Distance

The Figure 8 shows the corresponding colours of Figures 6 and 7. These coloured circles help to show which areas use large or small amounts of energy. It is evident that the areas in the blue circle (1.8 km mark) and the red circle (2.6 km mark) are sections of the route that influenced the total energy usage. The blue circle section is an area where a considerable amount of energy was saved due to the large downhill section. The red circle also saved some energy however there was a steep incline in the terrain which caused a large amount of energy to be used.

The route 5 is the one that requires the most amount of energy and it is also the longest route. It used a total of 154.67 calories over the total distance of 6.434 km.

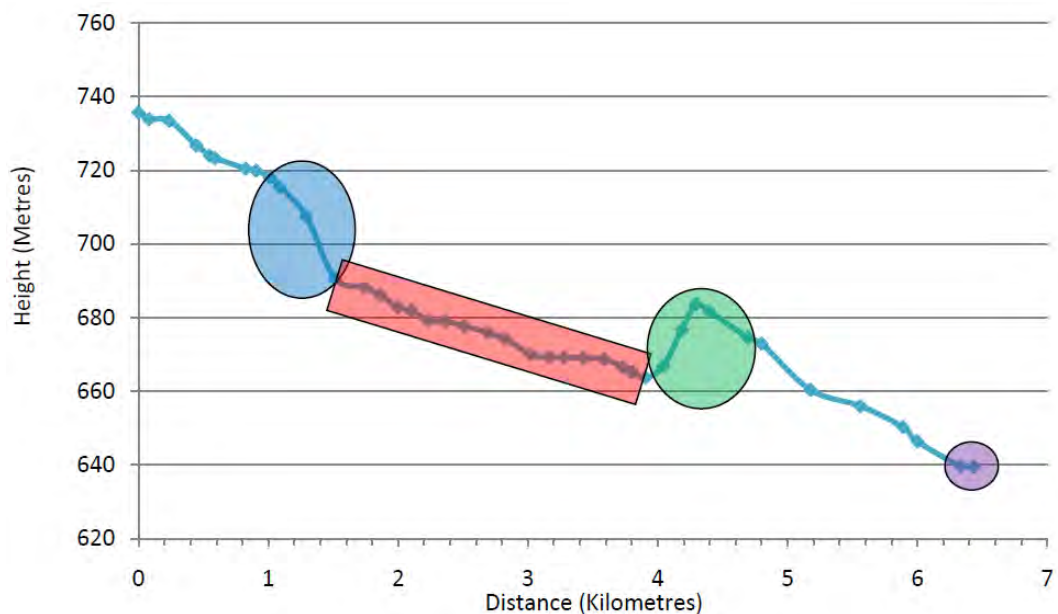


Figure 9: Route 5 Profile

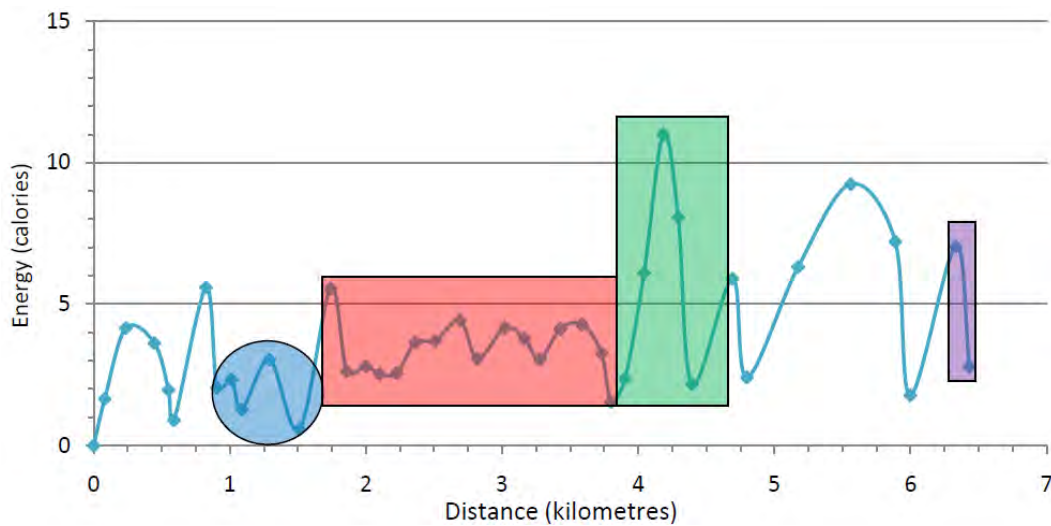


Figure 10: Route 5 Acquired Distance versus Energy

From the start point to the approximately the one kilometre mark there are fluctuations in the energy usage. Like the other routes this is due to the different distances used between grades amounting to spikes in energy. The blue circle seen in Figure 9 from the 1km mark to the 1.5 km mark, shows that there is a steep decline in the terrain. This is reflected in Figure 10 where the energy used is very low around 2-3 calories. The terrain then goes into a continuous decline from 1.5 km mark to the 3.75 km mark and this is represented by the red rectangle in Figure 10. Again Figure 10 reflects the terrain in the red rectangle showing the energy staying quite low and having small fluctuations. Figure 9 shows the green circle from the 3.75 km mark to the 4.8 km mark. This green circle illustrates where the terrain rises steeply. This is reflected in green in Figure 10 where the energy expenditure rises steeply as well. From this point to the 6.25 km mark the terrain begins to decline again at a slightly steeper grade then the previous section. The grade changes fluctuate in areas before flattening out after the 6.25 km mark seen in purple. Figure 10 shows that there are fluctuations in energy usage from the 4.8 km mark to the 6.25 km mark. This is because of the inconsistent distances used as well as areas where the grade changes slightly. The purple rectangle in Figure 10 also represents where the terrain flattens out and there is a spike in the energy.

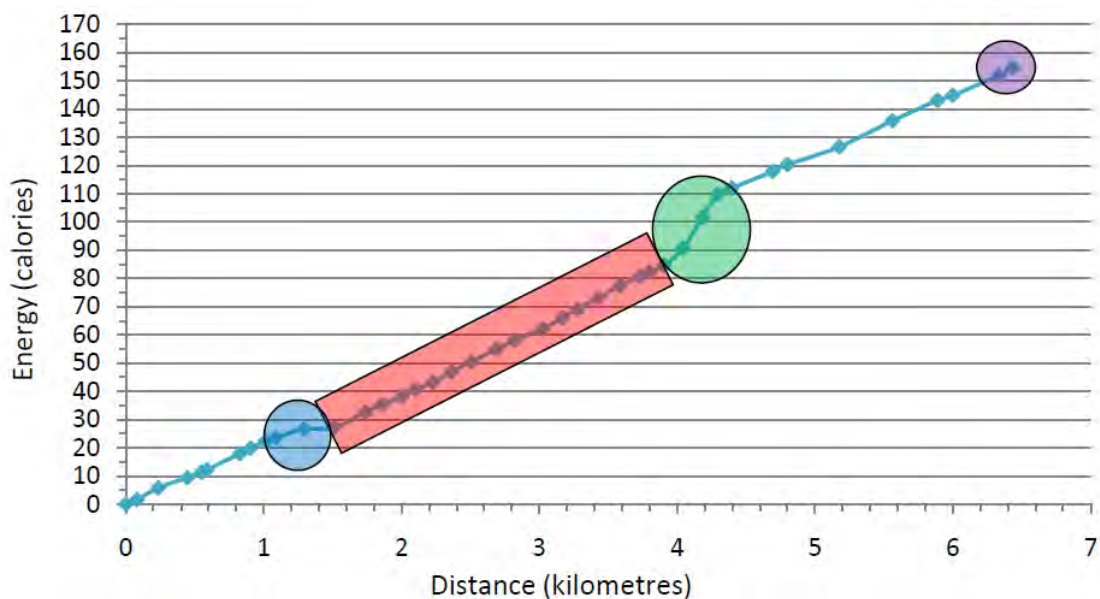


Figure 11: Route 5 Acquired Energy versus Acquired Distance

Figure 11 shows the acquired energy used along the route. It helps display where the energy was used and saved. The blue circle (1.2 km mark) is an area where a large amount of energy was saved because of the decline in the terrain. Although the red rectangle shows an area where the energy is increasing this is also an area where energy has been saved. The grade of this section is not very steep which indicates that the terrain is downhill which it is when looking at Figure 9. The green circle (4.2 km mark) in Figure 11 shows an area where the most energy of the route was used. This was caused by the steep incline in the terrain. Overall the route only had one major uphill section in the terrain, the rest was consistently downhill. The main reason for this route using so much energy is predominantly because it is approximately 250 m longer than most of the other routes.

Route 2, route 3 and route 6 are also broken down and viewed individually as the route 1, route 4 and route 5. After analysing all the routes, areas that have influenced the energy expenditure have been identified. It was also discovered that different distances in the routes could also have contributed to the differences in energy used. By dividing the distance by the calories a ratio can be established that gives energy efficiency. This means that the energy efficiency can be determined for each route to work out which route was the most energy efficient. The Table 2 shows that route 4 is not only the route that used the least amount of energy but it is also the route that is the most energy efficient. It can also be seen that route 3 was the next most energy efficient route. Route 3 used the third lowest amount of energy and is also the fourth shortest route. The results also show that route 1 and 5 were both the third most efficient routes. This is a very interesting result as route 1 is 403 metres shorter than route 5. This proves that the shortest route is not always the easiest route. Route 5 is the longest route so by it being the third most efficient route also reinstates the fact the shortest route is not always the most easiest. Route 6 is the fourth most energy efficient route with it being the route that used the fifth least amount of energy and it was the second longest route. The least energy efficient route was route 2. Route 2 is the second shortest distance and used the fourth lowest energy. This reinstates the fact that the shorter the route doesn't mean it is easier.

Table 2 Route Energy Efficiency

Route	Efficiency	Rating
1	0.04160	3
2	0.04098	5
3	0.04199	2
4	0.04266	1
5	0.04160	3
6	0.04132	4

4 CONCLUSION

This study demonstrated the success of mapping and analysis of the cycling pathways using GIS, GPS and developed energy expenditure formulae. Using GPS data and a DEM generated from aerial photography, six cycling pathways from Toowoomba campus of the University of Southern Queensland to the Toowoomba CBD were analysed using GIS software in terms of energy expenditure and the route length. A slope map was created from the DEM and classified based on slopes. The longest and shortest routes were determined, and the routes that requires the least and the most energy were determined as well. The effects of the terrain on energy usage was also analysed in details for different routes. Terrain areas that have different influence on the energy expenditure were analysed. The energy efficiency, which was defined as the ratio of distance to the calories, was calculated for each route. One route that used the least amount of energy was also identified as the most energy efficient route.

REFERENCES

- Abbiss, C. R., Peiffer, J. J. & Laursen, P. B. (2009). Optimal cadence selection during cycling. *International SportMed Journal*, 10(1), 1-15.
- Al-Haboubi, M. H. (1999). Modelling energy expenditure during cycling. *Ergonomics*, 42(3), 416-427.
- ANJEC. (2004). Pathways for the Garden State: a local government guide to planning walking, bikeable communities. Association of New Jersey Environmental Commissions, <http://www.anjec.org/pdfs/pathways.pdf>.
- ANRA. (2009). Groundwater management unit: Toowoomba City Basalt. Available: <http://www.anra.gov.au/topics/water/overview/qld/gmu-toowoomba-city-basalt.html> Accessed 08 August 2011.
- Atkinson, G., Davison, R., Jeukendrup, A. & Passfield, L. (2003). Science and cycling: current knowledge and future directions for research. *Journal of Sports Sciences*, 21(9), 767-787.
- Aultman-Hall, L., Hall, F. L. & Baetz, B. B. (1997). Analysis of Bicycle Commuter Routes Using Geographic Information Systems: Implications for Bicycle Planning. *Transportation Research Record* 1578(1), 102-110.
- Gray, M. & Bunker, J. M. (2005). Kelvin Grove Urban Village: The use of GIS in active transport planning. *Proceeding of 27th Conference of Australian Institutes of Transport Research*, 7-9 December 2005, Brisbane, Australia.
- IBI Group. (2000). Calgary Pathway and Bikeway Plan Report 2000. http://www.calgary.ca/docgallery/BU/trans_planning/cycling/pathway_and_bikeway_plan/part_1.pdf.
- Leigh, C., Peterson, J. & Chandra, S. (2009). Campus bicycle-parking facility site Selection: exemplifying provision of an Interactive facility map. *Proceeding of Surveying & Spatial Sciences Institute Biennial International Conference*, 28 September – 2 October 2009, Adelaide, Australia.
- Mueller, K. (2010). Cycling calories. Available: www.infinutrition.eu/library/Products%20page%20infinIT.doc. Accessed 25/06/2010.
- Rissel, C., Merom, D., Bauman, A., Carrard, J., Wen, L. M. & New, C. (2010). Current cycling, bicycle path use and willingness to cycle more - findings from a community survey of cycling in Southwest Sydney, Australia. *Journal of Physical Activity and Health*, 7(267-272).
- Shumowsky, R. (2005). Developing a bike/pedestrian plan using ArcInfo and public participation. *Proceeding of 2005 Esri International User Conference*, 25–29 July 2005, San Diego Convention Center, San Diego, California.

Wetland Mapping Using Remote Sensing Imagery and ModelMap

Xuan Zhu

School of Geography & Environmental Science
Monash University
Building 11, Clayton Campus
VIC 3800, Australia
Xuan.Zhu@monash.edu

ABSTRACT

Wetlands are an important part of the natural environment. They have rich biodiversity and provide habitat for animals and plants. Accurate mapping of wetlands is critical for monitoring their changes and assessing the effects of natural and anthropogenic factors on wetland ecosystems. Remote sensing imagery has been widely used for mapping wetlands and their changes. This paper presents a study on the comparison of two classification tree-based techniques – Random Forests (RF) and Stochastic Gradient Boosting (SGB) – for wetland mapping. RF and SGB are essentially statistical techniques based on classification and regression tree analysis. Through RF, classification trees are ‘grown’ using a random selection of the explanatory variables. Tens or hundreds of these trees may be grown for each classification. The final classification decision is based on the mode of the class’s output by individual trees. SGB produces a classification model in the form of decision trees. It uses classification errors to refine the tree development and integrates multiple classification trees into a single best classification. The two techniques are available in ModelMap, a software package for the free R statistical program. The study uses the Jiuzhaigou Nature Reserve in China as a case and assesses the accuracy of wetland mapping of the two methods.

KEYWORDS: wetland mapping; remote sensing; Random Forests; Stochastic Gradient Boosting; classification trees

1 INTRODUCTION

Wetlands are significant areas with rich biodiversity and support both aquatic and terrestrial animal and plant species. Basically, they are areas where water covers the soil. They may be permanently wet or ephemeral, fresh or saline, boggy marshes or filled with ponds (Nielsen & Brock, 2009). Wetlands can come in many different shapes and sizes, and can appear in many different geographic settings. Wetland ecosystems are part of our natural wealth. They deliver a wide range of ecosystem services, such as carbon sequestration (Gorham, 1991), flood mitigation (Olhan et al. 2010), fish and fibre provision, refuges for endangered wildlife (Kennedy & Mayer, 2002), water quality protection (Carter, 1999; Ambasht, 2008) and recreation. The importance of wetland conservation is now well established as a matter of international public policy (Ramsar Convention Bureau, 2004).

Accurate mapping of wetlands is critical for monitoring their changes and assessing the effects of natural and anthropogenic factors on wetland ecosystems. Remote sensing imagery has been widely used for mapping wetlands and their changes. Traditionally, aerial photography is used, which provides a very accurate method for wetland mapping. However, single-date aerial photographs are not informative with regard to wetland dynamics, as wetland water levels can vary widely within and among years (Cowardin & Golet, 1995). Repeatable wetland mapping over time via aerial photography is expensive and infeasible over the long term (Wright and Gallant, 2007). In addition, wetland maps derived from aerial photographs have an extensive time lapse between imagery acquisition and final map production (Baker et al., 2006). Satellite remote sensing offers a

method for automated, reproducible and less expensive wetland mapping. It provides imagery with better spectral, radiometric and temporal resolutions, but decreased spatial resolutions compared with aerial photography. Several studies on large-area wetland mapping demonstrated that overall accuracy is not compromised with coarser satellite images, comparable to that of aerial photographs (Bolstad and Lillesand 1992; Sader et al. 1995; Harvey and Hill, 2001; Baker et al., 2006). Therefore, satellite remote sensing has become a popular means for automated wetland mapping and monitoring.

Large-area satellite remote sensing of wetlands is based on computer-based classification of multispectral image data. Classification techniques generally fall into one of two camps, supervised or unsupervised. Unsupervised classification techniques classify the image into the land cover classes with no input from the user, whereas supervised classification techniques require at least some user interaction (mainly to create training sites or observations to aid classification) (Jensen, 2005). Some wetland mapping studies have successfully utilised supervised rule-based or classification tree-based methods with the reported overall accuracy of over 70% (Sader et al. 1995; Baker et al., 2006). Typical classification tree-based methods include classification tree analysis (CTA), Random Forests (RF) and Stochastic Gradient Boosting (SGB).

CTA, also referred to as classification and regression trees (CART), is a statistical technique used to determine a set of if-then logical (split) conditions that permit accurate prediction or classification of cases via tree-building algorithms (Breiman et al., 1984). A CTA analysis often produces dichotomous trees, in which leaf nodes represent classifications and branch nodes represent conjunctions of features that lead to those classifications. In image classification, it typically operates by recursively parsing the training observations in a form of binary partitioning based on the values of the selected explanatory variables such as spectral responses and ancillary data (e.g. slope, vegetation and soils) (Lawrence and Wright, 2001; Wright and Gallant, 2007). The classification trees formed by CTA can be used as classification rules either by themselves in rule-based classifiers or combined with expert knowledge. CTA is essentially non-parametric and non-linear (Breiman et al., 1984). No assumptions are made regarding the underlying distribution of values of the explanatory variables. Therefore, it can handle numerical data that are highly skewed or multi-modal, as well as categorical variables with either ordinal or non-ordinal structure. CTA uses efficient algorithms to search for all best possible variables for classification from those provided by the analyst, and is a relatively automatic “machine learning” method with relatively little user input. Many studies reported higher accuracies in land cover classification resulted from CTA than other classification methods such as maximum likelihood, minimum distance to means, and parallelepiped (Hansen, 1996; Friedl and Brodley, 1997; Lawrence and Wright, 2001; Pal and Mather, 2003). It has been successfully applied in wetland mapping (Wright and Gallant, 2007). However, CTA does not necessarily produce the optimal classification tree, and the results may be adversely affected by inaccuracies and outliers in the training data (Moisen, 2008). Several ensemble classification methods have been developed to improve CTA. Random Forests (RF) and Stochastic Gradient Boosting (SGB) are two of these newer methods that use CTA as building blocks.

RF uses bagging, or bootstrap aggregating, to form an ensemble of classification trees (Breiman, 2001). Bagging involves developing multiple classification trees by selecting random subsets of the explanatory variables and original training data. RF uses an improved method of bagging. It creates a number of classification trees, each trained on a bootstrapped sample of the original training data, and searches only a random sample of the explanatory variables for a split at each tree node. For classification, each tree in RF casts a unit vote for the most frequent class to the input data. The output of RF is determined by a majority vote of the trees. This method is not sensitive to noise or overtraining, and has achieved an overall accuracy of over 80% in land cover classification, better than CTA (Gislason et al., 2006; Rodríguez-Galiano et al., 2011).

SGB is a hybrid of the bagging and boosting approaches (Friedman, 2002). Boosting starts with a standard classification tree built from the entire sample of the training data. It uses iterative re-training, in which the incorrectly classified training data are given higher weighting, which results in a new classification tree that emphasises the most difficult classification problems in the training data. The process is iterated to generate a set of classification trees, which are then used to determine the correct classification using a plurality rule (Friedman, 2002). Boosting generally reduces both the variance and the bias of the classification, but it can overtrain and is sensitive to

noise (Gislason et al., 2006). SGB combines boosting with bagging to effectively reduce the sensitivity to inaccuracies and outliers (i.e. noise) in the training data. It selects a random subset of the training data to perform the boosting, uses a steepest gradient algorithm to give higher weights on the misclassified training data that are close to their correct classification (rather than the worst classified data), and builds many small classification trees sequentially from pseudo-residuals from the previous tree (Friedman, 2002; Lawrence et al. 2004). At each iteration, a classification tree is built from a random subset of the training data producing an incremental improvement in the classification. Finally, all the small classification trees are stacked together as a weighted sum of terms. The overall accuracy of classification improves progressively with each additional term. SGB has achieved an overall accuracy of more than 90% in land cover classification using high resolution satellite images (such as IKONOS), but only 62% with moderate resolution Landsat ETM+ images (Lawrence et al. 2004). Baker et al. (2006) reported an 86% overall accuracy in wetland mapping using Landsat ETM+ images.

Both RF and SGB are refinements of the conventional CTA. As an improved version of bagging, RF was reported to be comparable to boosting in terms of accuracies, but without the drawbacks of boosting (Breiman, 2001). SGB integrates boosting and bagging approaches and has been shown in most cases to generate significantly higher accuracies than CTA and other boosting methods (Friedman, 2002). The objective of this study is to compare the performance of RF and SGB in wetland mapping using Landsat ETM+ imagery, using the Jiuzhaigou Nature Reserve in China as a case study.

2 METHODS

2.1 Study area

The Jiuzhaigou Nature Reserve is located at the northern foothills of the Duangar Peak in the Min Mountains that forms part of the eastern edge of the Tibetan Plateau in the north of Sichuan Province, China (Figure 1). It covers an area of more than 700 square km. The reserve is renowned for its Karst landforms that create a unique landscape of 17 spectacular waterfalls, 114 crystal blue lakes, 5 tufa terraces, 47 springs, and 11 derived flows that move through the snow-capped mountains and travel for about 50km along the Y-shaped forested valley (Liu et al., 1996). The elevation ranges from 1,800 to 4,800m. The average annual temperature is 7.3°C. Temperature ranges between -20.2°C to 32.6°C. The annual rainfall ranges from 649.6 to 940.6 mm. The climate of Jiuzhaigou shows a clear alternation of dry and rainy seasons. Its spectacular topography and the presence of different parent materials and aspects have produced numerous niches for diverse flora and fauna. It is the home of some 140 bird species and a number of endangered plant and wildlife species, including the giant panda and the Sichuan takin (Liu et al., 1996). The reserve was designated as a World Heritage Site in 1992 and a World Biosphere Reserve in 1997. Agricultural and pastoral practices and tree logging within the reserve have ceased since 2000. The current human activity is mainly tourism.

The Jiuzhaigou Nature Reserve has all palustrine, riverine, and lacustrine wetland systems (Cowardin et al., 1979). Palustrine wetlands include vegetated wetlands as well as small, shallow, open water bodies. The riverine system consists of river channels containing open water or nonpersistent vegetation. Lacustrine wetlands comprise both permanent and intermittent lakes containing nonpersistent vegetation. The goal of the study is not to classify these different types of wetland from remote sensing imagery, but to focus on the separation of wetland from upland at a coarse level. The goal is to identify an accurate and automated technique for mapping wetlands.

Jiuzhaigou Nature Reserve



Figure 1: Location of the Jiuzhaigou Nature Reserve

2.2 Wetland mapping

The study used both RF and SGB for mapping wetlands of the Jiuzhaigou Nature Reserve using the Landsat-TM images acquired in September 2004 (the wet season), and then compared the mapping accuracies of the two techniques. The Landsat-TM images were re-sampled to 40m resolution. Six bands of the TM images (excluding Band 6 - the thermal band) were used as spectral variables. The ancillary data set used in this project is the slope map derived from a 40-m DEM generated from the 1:100,000 topographic map. 876 training sites were developed for five broad types of land cover in the study area, including wetland, woodland, pasture, dry land and glacier/permanent snow. These training sites were selected using the on-site surveys, 2002 land use map of the study area (1:50,000) and QuickBird satellite images (with a spatial resolution of 0.6m) acquired in November 2004. Land cover maps were first generated with RF and SGB using the TM images, slope data and the training sites, from which wetland maps were then derived. The land cover classification was conducted in ModelMap – a software package for the free R statistical program (R Development Core Team, 2008).

ModelMap was developed for building predictive models using RF and SGB (Freeman et al., 2010). It provides an interface between several existing R packages to automate and simplify the process of RF and SGB analysis and modelling using spatial data (in either ESRI Grid or ERDAS Imagine image raster data formats). ModelMap enables model building over large geographical areas using training data with RF or SGB, model validation with an independent test set and cross-

validation, and model output in the form of probability maps. In this study, the following procedure was used for wetland mapping.

- 1) Identifying major types of land cover from the TM images and generating probability maps of each type of land cover

Using six spectral variables (TM Bands 1 – 5 and Band 7) and slope as the explanatory variables, RF and SGB models were built with ModelMap respectively for classifying the TM images into individual land cover maps. A pair of RF and SGB models was built for each of the five types of land cover using the same training data sets. Each model was then used to produce a probability map for a particular type of land cover. Each probability map represents the probability that a particular type of land cover is present at every location.

- 2) Combining the probability maps to derive wetland maps

The probability maps were read into ArcGIS for map production and for identifying wetlands. There are five RF-derived and five SGB-derived probability maps respectively for wetland, woodland, pasture, dry land and glacier/permanent snow. The five RF-derived probability maps were combined via overlay to create the RF-derived land cover map, in which each location was assigned to the type of land cover that has the highest probability at that location. The five SGB-derived probability maps were combined in the same way to produce the SGB-derived land cover map. Each of the two land cover maps was finally reclassified to generate a wetland map.

- 3) Assessing mapping accuracies

The performance of RF and SGB for mapping the probability of presence of each type of land cover was assessed using four accuracy measures: sensitivity, specificity, Kappa and the area under the ROC (Receiver Operating Characteristic) curve (or AUC). The first three are dependent on probability thresholds. Sensitivity measures the proportion of actual positives which are correctly identified (e.g. the percentage of actual wetlands that are correctly identified as occurring at a location). Specificity measures the proportion of true negatives which are correctly identified (e.g. the percentage of non-wetlands that are correctly identified at a location where wetland does not exist). Kappa is a statistic measuring the agreement on classification by taking into account the agreement occurring by chance. It is the proportion of correctly classified locations after accounting for the probability of chance agreement. A detailed review of the three measures can be found in Fielding and Bell (1977).

An ROC curve is a graphical plot of sensitivity, or true positive rate, against false positive rate ($= 1 - \text{specificity}$) (Fawcett, 2006). The best possible classification method would produce a point in the upper left corner or coordinate (0, 1) of the ROC plot, which represents 100% sensitivity (no false negatives) and 100% specificity (no false positives). It represents a perfect classification. A completely random assignment would give a point along a diagonal line where the true positive rate equals to the false positive rate. Points above the diagonal represent good classification results, while points below the line poor classification results. The area under the ROC curve, or AUC, provides a measure of overall performance of a classifier. A good classifier should have an AUC near 1, while a poor classifier has an AUC near 0.5.

Standard error matrix measures, including overall accuracy, producers accuracy and users accuracy, were used to assess and compare the accuracies of the RF-derived and SGB-derived wetland maps. The accuracy assessment data obtained through on-site surveys and interpretation of QuickBird satellite images included 69 random points sampled in wetlands and 731 random points sampled in non-wetland areas.

3 RESULTS

3.1 Probability maps

RF and SGB in ModelMap were used to produce probability maps for each type of land cover. Figure 2 shows the RF-derived probability maps respectively for wetland, woodland, pasture, dry land and glacier/permanent snow (on all the maps, darker the area, lower the probability of presence). The SGB-derived probability maps for the five types of land cover are shown in Figure 3.

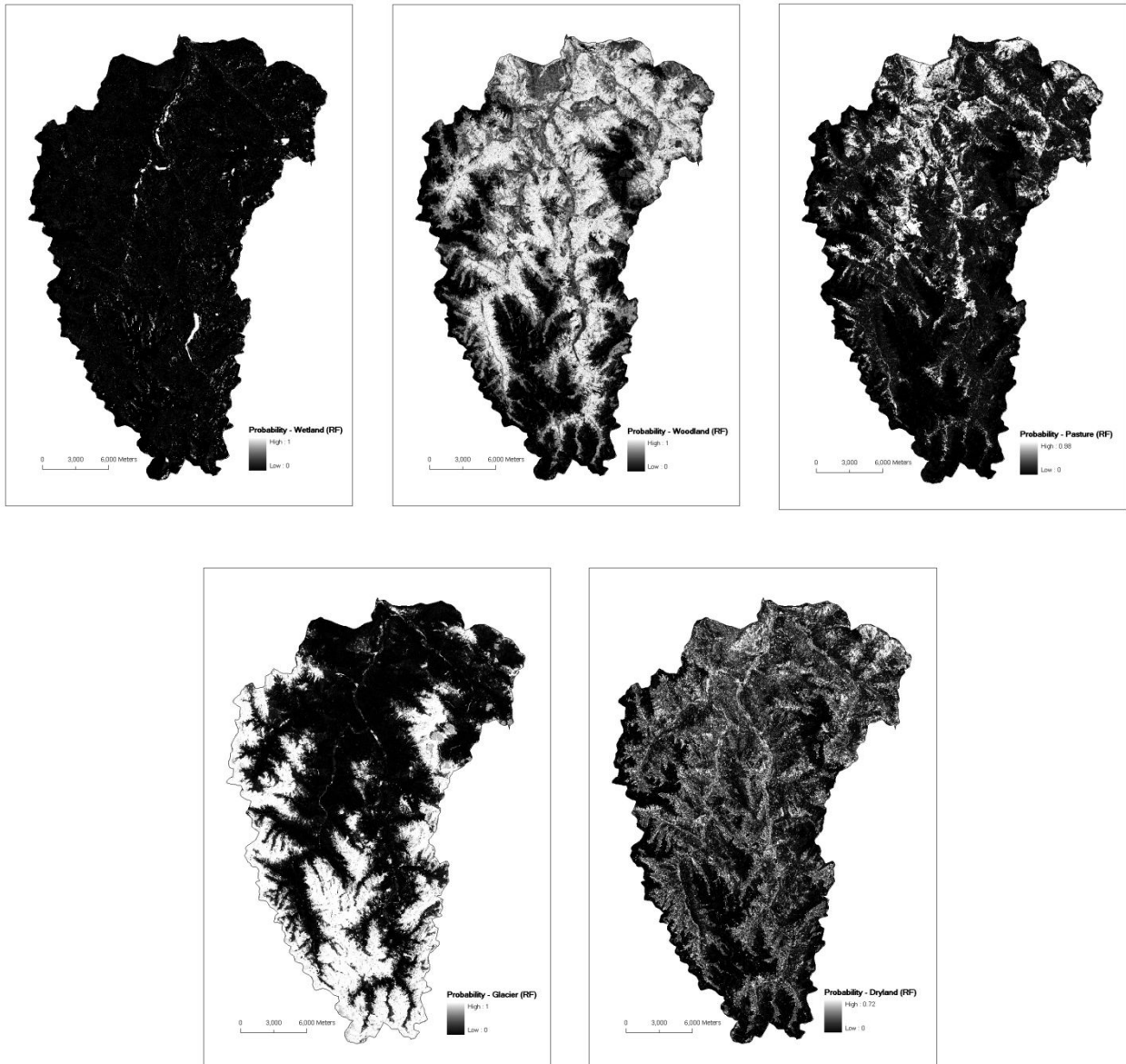


Figure 2: RF-derived probability maps

The RF-derived probability maps look similar to the corresponding ones derived by SGB. According to Table 1, they both performed very well in predicting the presence of wetland, and glacier and permanent snow, moderately well in predicting woodland and pasture, and poorly in predicting dryland. SGB had higher values of Kappa for all the types of land cover except dryland. However, the differences of Kappa between RF and SGB are very small, indicating that the two techniques produced a similar level of agreement on classification. Compared with RF, SGB had a larger proportion of actual positives that were correctly identified and a lower proportion of actual

negatives that were correctly identified for all the types of land cover except glacial/permanent snow. The poor performance on dryland may be because this type of land cover is distributed in some isolated small pockets in the nature reserve and the number of the training points sampled in these areas is quite small. There were substantial confusions between dryland and pasture.

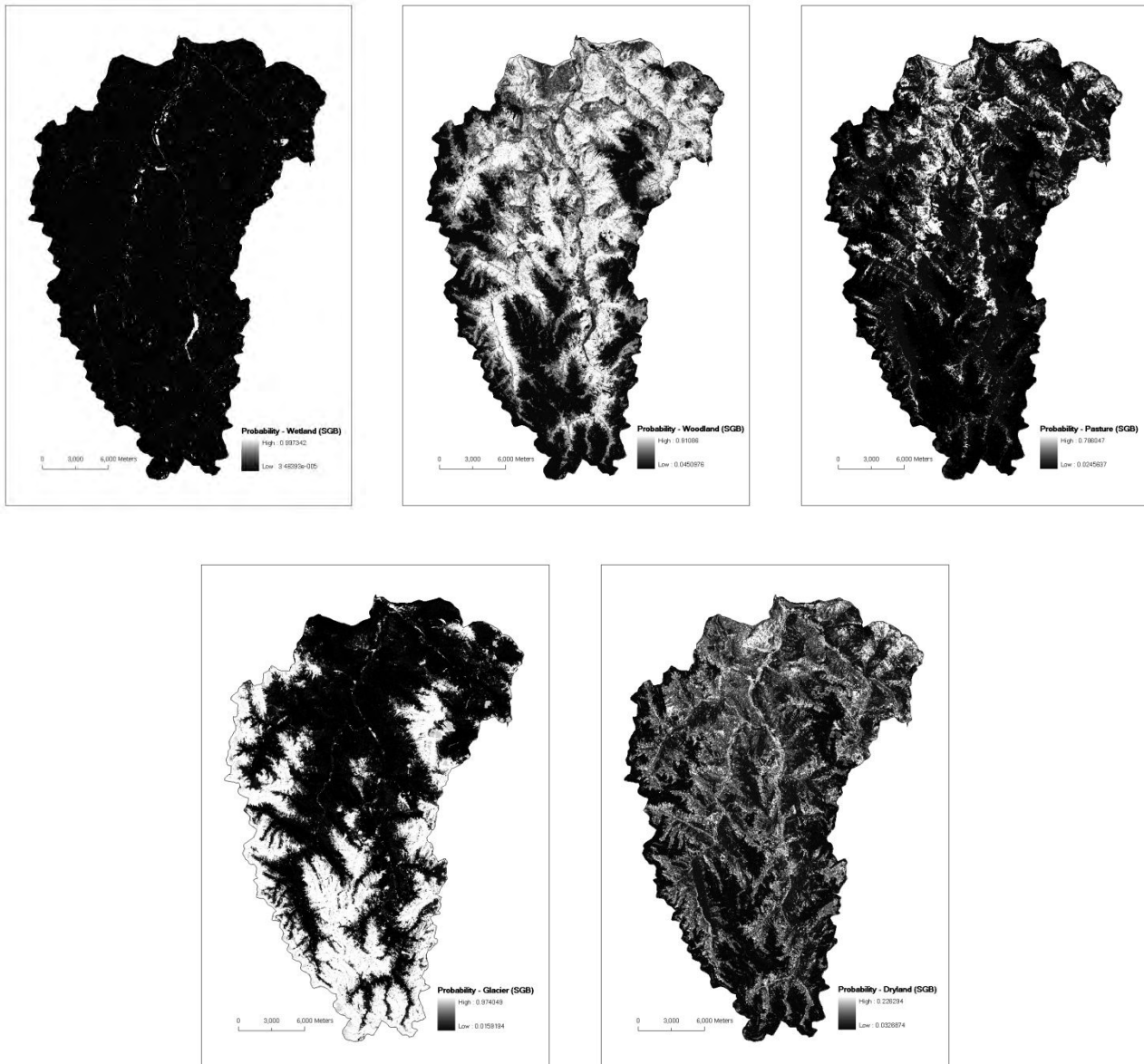


Figure 3: SGB-derived probability maps

Table 1 Accuracy measures for each type of land cover with RF and SGB (the thresholds were chosen to maximise Kappa)

Land Cover	Kappa		Sensitivity		Specificity		AUC	
	RF	SGB	RF	SGB	RF	SGB	RF	SGB
Wetland	0.99	0.99	0.98	0.99	0.99	0.98	0.99	0.99
Woodland	0.65	0.66	0.79	0.90	0.87	0.77	0.91	0.91
Pasture	0.53	0.57	0.62	0.63	0.95	0.95	0.89	0.89
Dryland	0.12	0.11	0.25	0.32	0.92	0.88	0.69	0.70
Glacier/snow	0.84	0.86	0.90	0.87	0.94	0.97	0.97	0.98

3.2 Wetland maps

The RF-derived wetland map was created by combining the six probability maps in Figure 2. The SGB-derived wetland map was generated by combining the six probability maps in Figure 3. The two wetland maps are shown in Figure 4. The general patterns of wetland are similar in the two maps. The SGB-derived wetland map contains many more speckles. As shown in Tables 2 and 3, the two maps had an equal Kappa value of 0.74 (a very strong agreement) and almost identical overall accuracies (96% with SGB and 97% with RF). However, the RF-derived map had a users accuracy of 96% for wetland, much higher than 82% for the map derived by SGB. The SGB-derived map had a slightly higher producers accuracy of 71% for wetland than the RF-derived map, whose producers accuracy is 62%. Therefore, SGB resulted in more commission errors, while RF created more omission errors. Errors in the SGB-derived wetland map mainly resulted from non-wetland areas incorrectly classified as wetland. The majority of errors in the RF-derived wetland map were caused by misclassification of wetland as non-wetland.

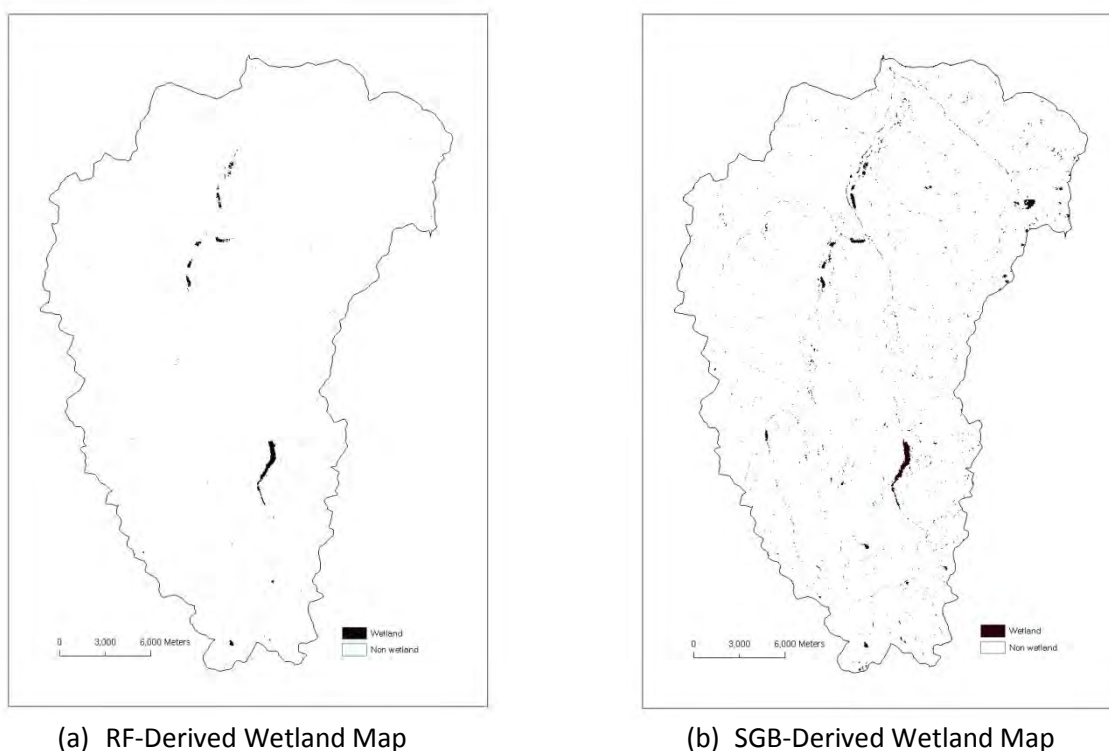


Figure 4: Wetland maps

Table 2 Error matrix for RF

Classified Data	Reference Data			Users Accuracy
	Wetland	Non-Wetland	Row Total	
Wetland	43	2	45	96%
Non-Wetland	26	729	755	97%
Column Total	69	731	800	
Producers Accuracy	62%	99%		
Overall Accuracy = 97%			Kappa = 0.74	

Table 3 Error matrix for SGB

Classified Data	Reference Data			
	Wetland	Non-Wetland	Row Total	Users Accuracy
Wetland	49	11	60	82%
Non-Wetland	20	720	740	97%
Column Total	69	731	800	
Producers Accuracy	71%	98%		

Overall Accuracy = 96%

Kappa = 0.74

4 CONCLUSION

RF and SGB have been reported to achieve improved accuracy of predictions and classifications relative to other classification methods (Breiman, 2001; Friedman, 2002). Both are nonparametric and robust to noise in the data, and can handle large data sets, missing data and outliers. SGB and RF were used independently in land cover and wetland classifications using satellite imagery and achieved satisfactory outcomes (Lawrence et al., 2004; Baker et al., 2006; Gislason et al., 2006).

In this paper, the two classification-tree based techniques were used and assessed for land cover and wetland mapping using moderate resolution satellite images. The results indicated that the two methods are powerful and effective in land cover and wetland classification. In individual land cover mapping, RF and SGB achieved a similar level of agreement on classification on all types of land cover. They both performed well in estimating probabilities of the presence of wetland and glacier/permanent snow, moderately well for pasture and woodland, but poorly in predicting the presence of dryland. The final wetland maps were generated by identifying the type of land cover that has the highest probability of presence at every location in the study area. Both RF and SGB achieved high overall accuracies and class accuracies in the final wetland mapping. However, SGB produced more commission errors than RF.

The study used ModelMap to implement RF and SGB. It demonstrated that ModelMap is simple to use and efficient in handling large image data sets. However, the RF and SGB methods implemented in this package are standard ones, which do not work well for unbalanced samples (Breiman et al., 2004; Liu et al., 2008). The poor performance of the RF and SGB in ModelMap on dryland is due to this problem as dryland represented a very small group in the training sample compared with other types of land cover. Re-sampling might be needed to increase the representation of dryland in the sample. In addition, the RF is less likely to overfit data, but the SGB may overfit data (Freeman et al., 2010). Overfitting may be the major cause of the higher number of commission errors generated by the SGB in its derived wetland map.

In general, the accuracies of land cover and wetland mapping with RF and SGB using moderate resolution satellite images were high. As demonstrated in this study, their performance is largely dependent on the classes of land cover to be identified and the quality and representativeness of training and accuracy assessment data. There is no doubt that their performance also relies on the explanatory variables used. In this study, Landsat TM spectral data and slope were used as the explanatory variables in modelling. It is anticipated that including more ancillary data, such as vegetation and soils, would further improve the accuracy of wetland mapping as demonstrated in Wright and Gallant (2007).

REFERENCES

- Ambasht R. (2008). Wetland ecology: An overview. Proceedings of the National Academy of Sciences India Section B- Biological Sciences, 78, 3 – 12.
- Baker C., Lawrence R., Montagne C. Patten D. (2006). Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models. Wetlands, 26, 465-474.

- Bolstad P.V., Lillesand T.M. (1992). Rule-based classification models: flexible integration of satellite imagery and thematic spatial data. *Photogrammetric Engineering and Remote Sensing*, 58, 965-971.
- Breiman L. (2001). Random forests. *Machine Learning*, 45, 5-32.
- Breiman L., Chen C., Liaw A. (2004). Using random forest to learn imbalanced data. *Journal of Machine Learning Research*, no. 666, Department of Statistics, University of California, Berkeley, CA.
- Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984). *Classification and Regression Trees*. Wadsworth & Brooks/Cole Advanced Books & Software.
- Carter V. (1999). Technical aspects of wetlands — Wetland hydrology, water quality, and associated functions. U.S. Geological Survey Water-Supply Paper, 2425.
- Cowardin L.M., Carter V. Golet F.C., Laroe E.T. (1979). *Classification of Wetlands and Deepwater Habitats of the United States*. U.S. Fish and Wildlife Service.
- Cowardin L.M., Golet F.C. (1995). U.S. Fish and Wildlife Service 1979 wetland classification: A review. *Vegetation*, 118, 134–15.
- Fawcett T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861–874.
- Fielding A.H. and Bell J.F. (1977). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38-49.
- Freeman E.A., Frescino T.S., Moisen G.G. (2010). *ModelMap: An R Package for Modeling and Map production using Random Forest and Stochastic Gradient Boosting*. USDA Forest Service, Rocky Mountain Research Station, Ogden, UT, USA. URL <http://CRAN.R-project.org/>.
- Friedl M.A., Brodley C.E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61, 399–409.
- Friedman J.H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis*, 38, 367-378.
- Gislason P.O., Benediktsson J.A., Sveinsson J.R. (2006). Random Forests for land cover classification. *Pattern Recognition Letters*, 27, 294–300.
- Gorham E. (1991). Northern peatlands: Role in the carbon cycle and probable responses to climatic warming. *Ecological Applications*, 1, 182–195.
- Hansen M., Dubayah R., Defries R. (1996). Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, 17, 1075–1081.
- Harvey K.R. and Hill G.J.E. (2001). Vegetation mapping of a tropical freshwater swamp in the Northern Territory, Australia: a comparison of aerial photography, Landsat TM and SPOT satellite imagery. *International Journal of Remote Sensing*, 22, 2911-2925.
- Jensen J. (2005). *Introductory Digital Image Processing: A Remote Sensing Perspective*, 3rd Edition. Pearson Education.
- Kennedy G., Mayer T. (2002). Natural and constructed wetlands in Canada: An overview. *Water Quality Research Journal of Canada*, 37, 295 - 325.

- Lawrence R.L., Bunn A., Powell S., Zambon M. (2004). Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis, *Remote Sensing of Environment*, 90, 331–336.
- Lawrence R.L., Labus M.A. (2003). Early detection of Douglas-fir beetle infestation with sub-canopy resolution hyperspectral imagery. *Western Journal of Applied Forestry*, 18, 202–206.
- Lawrence R.L., Wright A. (2001). Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric Engineering and Remote Sensing*, 67, 1137–1142.
- Liu S.Q., Tang X.C., Tang B.X., Yang, Y. (1996). *The World Natural Heritage: Ecological Environment and Protection in Jiuzhaigou*. Chengdu Science and Technology Press, Chengdu.
- Liu X.Y., Wu J. and Zhou Z.H. (2008). Exploratory Undersampling for Class-Imbalance Learning. *Systems, Man, and Cybernetics, Part B: Cybernetics*, IEEE Transactions, 39, 539-550.
- Moisen G.G. (2008). Classification and regression trees. In S.E. Jorgensen and B.D. Fath, editors, *Encyclopedia of Ecology*, Vol. 1, 582-588. Elsevier.
- Nielsen D., Brock M. (2009). Modified water regime and salinity as a consequence of climate change: prospects for wetlands of Southern Australia. *Climatic Change*, 95, 523 – 533.
- Olhan E., Gün S., Ataseven Y., Arisoy, H. (2010). Effects of agricultural activities in Seyfe Wetland. *Scientific Research and Essays*. 5, 9–14.
- Pal M., Mathers P. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86, 554–565.
- Ramsar Convention Bureau (2004). The list of wetlands of international importance: 04 July 2011. URL <http://www Ramsar.org/pdf/sitelist.pdf>
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- Rodríguez-Galiano V.F., Abarca-Hernández F., Ghimire B., Chica-Olmo M., Atkinson P.M., Jeganathan C. (2011). Incorporating spatial variability measures in land-cover classification using Random Forest. *Procedia Environmental Sciences*, 3, 44–49.
- Sader S.A., Ahl D., Liou W.S. (1995). Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine. *Remote Sensing of Environment*, 53, 133–144.
- Stehman S.V., Wickham J.D., Smith J.H., Yang L. (2003). Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: Statistical methodology and regional results. *Remote Sensing of Environment*, 86, 500–516.
- Wickham J.D., Stehman S.V., Smith J.H., Yang L. (2004). Thematic accuracy of the 1992 National Land-Cover Data for the western United States. *Remote Sensing of Environment*, 91, 452–468.
- Wright C., Gallant A. (2007). Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sensing of Environment*, 107, 582–605.

Characterising Heterogeneous Vegetated Surfaces Using Multiangular Satellite Data

Geoff McCamley

School of Mathematics and Geospatial Science, RMIT University, Melbourne Australia
S3142528@student.edu.au

Ian Grant

Bureau of Meteorology, Melbourne Australia
i.grant@bom.gov.au

Simon Jones

School of Mathematics and Geospatial Science, RMIT University, Melbourne Australia
Simon.jones@rmit.edu.au

Chris Bellman

School of Mathematics and Geospatial Science, RMIT University, Melbourne Australia
Chris.bellman@rmit.edu.au

ABSTRACT

Bidirectional Reflectance Distribution Functions (BRDF) represent variations in surface reflectance resulting from changes in a satellite's view and solar illumination angles. BRDF representations have been widely used to assist in the characterisation of vegetation. However BRDF effects are often noisy, difficult to interpret and are the spatial integral of all individual features present in a pixel.

This paper describes the results of an approach to understanding how BRDF effects can be used to characterise vegetation. The implementation of the Ross Thick Li Sparse BRDF model using MODIS is a stable, mature data product with a 10 year history and is a ready data source. Using this dataset, a geometric optical model is proposed that seeks to interpret changes in the Normalised Difference Vegetation Index (NDVI) with respect to changes in the viewing angle. The model described in this paper expresses the BRDF effects in terms of NDVI, vegetation density and a vegetation height-to-width ratio. This alternative parameter set seeks to offer a more directly interpretable characterisation of vegetated surfaces. The model proposed here has been applied to MODIS pixels along a transect from Melbourne to Darwin (Australia) and considered at various temporal periods. This model reveals additional information not existing in reflectance data. For example, using the height-to-width ratio, areas of vegetation are revealed that do not have an accompanying significant increase in NDVI.

KEYWORDS: BRDF, Vegetation, NDVI

1 INTRODUCTION

Bidirectional Reflectance Distribution Functions (BRDF) represent variations in surface reflectance resulting from changes in a satellite's view angle and sun elevation angle. BRDF representations have been widely used to assist in the characterisation of vegetation (Lovell and Graetz, 2001) (Hill et al., 2008) (Grant, 2000). Variations in BRDF effects have been observed in past studies and metrics have been developed to characterise vegetation using this information, e.g. Structural Scattering Index (SSI) (Gao et al., 2003), Anisotropic Factor (ANIF) (Sandmeier et al., 1998). Past studies have also shown BRDF effects to be noisy and difficult to interpret (Gao et al., 2003).

Sensors capable of sufficient angular sampling and revisit frequencies necessary for determining BRDF effects tend to have moderate to low spatial resolution. BRDF effects cannot be directly derived from Landsat data due to the small variation in illumination and viewing angles and infrequent revisits of the sensor (Li et al., 2010). Other sensors with similar or higher spatial resolutions than the 30m provided by Landsat, will also typically have small variations in angular sampling and infrequent revisits limiting their ability to derive BRDF effects. As a consequence, BRDF effects derived from satellite data will be a spatial integral of the different surface features present in a pixel. Therefore to study BRDF effects of specific features, surfaces often need to be found that are homogeneous at a given spatial scale, but these surfaces are often of limited interest, for example dried salt lakes.

This paper describes an alternative approach to interpreting BRDF effects. A simple model is developed in which a pixel is considered to be a composite of two elements: a soil layer having a low NDVI response and a vegetation component with a higher NDVI response. NDVI is derived from MODIS reflectance bands as $(R_2 - R_1) / (R_2 + R_1)$. The soil layer may be considered a flat surface and the vegetation components vertical prism shaped protrusions. The vegetation prisms may be described in dimensions that are a proportion of a pixel's dimensions. Considering the geometry of the surfaces, an expression can be derived for changes in NDVI with view angle. Large view angles will bring the vertical surfaces of vegetation (with a higher NDVI response) into view and obscure an area of soil (with a lower NDVI response). Treating changes in NDVI across a range of viewing angles derived from the MODIS BRDF representation as 'observations', numerical values for the alternate parameter set within the model may be derived that provide a best fit (lowest root mean square error (RMSE)) to MODIS NDVI across the same range of viewing angles. This approach seeks to replace the wavelength specific isotropic, volumetric and geometric BRDF parameters within the MODIS model with another set, namely NDVI, density and the height-to-width ratio of the vegetation. Formulation of this model is expressed in equations (2) – (5) of Appendix 1 which also contains a more detailed explanation of the model's development.

The dependence of NDVI on view angle has three parameters:

- the density of the vegetation cover (D),
- the height-to-width ratio of the vegetation (H), and
- NDVI difference between the vegetation components and the soil layer.

Application of the model requires determination of the NDVI profiles for soil and vegetation components present within the pixel. In this model NDVI profiles for soil and vegetation may be thought of as two end-members in a spectral mixture analysis where both the reflectance and angular properties are considered. In well defined, homogeneous small areas the NDVI profiles for vegetation and soil layers may be determined by direct observation or observation of similar surfaces components in another location or at another epoch. However, application of this model to spatially extensive areas (for example, a Melbourne to Darwin transect), where vegetated surfaces are variable, non-homogeneous and not well defined, the end-members must be estimated. The simplest approach that imposes a minimum of assumption is to use $NDVI_{veg} = 1$ and $NDVI_{soil} = 0$ with no associated angular variability. Use of these values further simplifies equations (2) – (5) in Appendix 1. Substituting these values into equations (2) – (5) results in the following expression:

$$\text{NDVI}(\theta) = \text{NDVI}(0^\circ) (1 + H \tan(\theta)) \quad (1)$$

Where: $\text{NDVI}(0^\circ)$ is the NDVI as observed from a 0° view angle, 0° illumination and 0° relative azimuth, which may be evaluated from MCD43B1 product as the normalised difference of the NIR and Red weights of the isotropic kernel, H is the Height-to-Width Ratio of the vegetation, θ is the viewing angle.

In equation (1), consideration of the difference in the NDVI response between the vegetation components and the soil layer is lost from the model described in Appendix 1, i.e. equation (1) assumes a maximum NDVI response difference between vegetation components and the soil layer. Furthermore, $\text{NDVI}(0^\circ)$ becomes the direct measure for the fraction of the pixel covered by vegetation and a separate value representing the density of the vegetation (i.e. D in the model) is no longer derived. In the manner as described above, a value for H in equation (1) can be determined from a best fit (lowest RMSE) to NDVI computed across a range of viewing angles using the MODIS BRDF representation.

2 METHOD

This study uses an established BRDF representation contained within existing data sets. This is a pragmatic approach to data collection and processing, and means that any results have the potential for broad and immediate application. NASA's Moderate Resolution Image Spectrometer (MODIS) sensor provides a BRDF representation in the MCD43 product. Specifically, the MODIS MCD43B1 product provides BRDF parameters of the Ross-Thick-Li-Sparse BRDF model at 500 m resolution in eight-day intervals. The MODIS BRDF representation models the atmospherically corrected surface reflectance as a function of the sun and view direction providing weights for a linear combination of isotropic, volumetric and geometric kernel functions. The weights are derived separately for each of the seven observed visible to mid-infrared bands (Lucht et al., 2000). The MODIS BRDF representation has been shown to be accurate (Schaaf et al., 2002) and derived reflectance less susceptible to noise than individual BRDF parameters (Lucht and Lewis, 2000).

The MCD43 product is distributed by the Land Processes Distributed Active Archive Centre (LP DAAC, <https://lpdaac.usgs.gov/>) as 10 degree tiles in a sinusoidal projection. For the work described here, the tiles have been mosaiced and re-projected to a Geographical grid covering the Australian continent, using the MODIS Re-projection Tool (MRT) software distributed by the LP DAAC and applied by the Australian Government's Commonwealth Scientific and Industrial Research Organisation (CSIRO) (Paget and King, 2008)

Using the MCD43B1 dataset, a transect of 6,180 pixels has been extracted between Melbourne (Lat $37^\circ 49'$ S, Long $144^\circ 58'$ E) and Darwin (Lat $12^\circ 27'$ S, Long $130^\circ 51'$ E). For each pixel along the transect, the model as described by equation (5), has been applied enabling derivation of a height-to-width ratio (H).



Figure 1: Transect from Melbourne to Darwin, being a distance of approximately 3,100 km. The white line is path of the transect and the red square indicates the location of an enlargement shown in Figure 5.

The quality assurance flags associated with the MCD43B1 product indicated poorer BRDF inversion results during summer epochs for pixels along the transect nearing Darwin. This may reasonably be assumed the consequence of cloud cover during the wet season that reduced the number of observation available when performing the BRDF inversion. No adjustment or removal of observed pixels has been made in response to lower quality BRDF inversions.

3 RESULTS

For pixels along the transect, the model as described by equation (1) has been applied and a height-to-width ratio derived. Epochs have been selected that represent seasonal variations. The model generally provided a good fit, with low RMSE and correlation coefficients generally exceeding 0.9 to NDVI when derived across the same range of view angles using the MODIS BRDF representation.

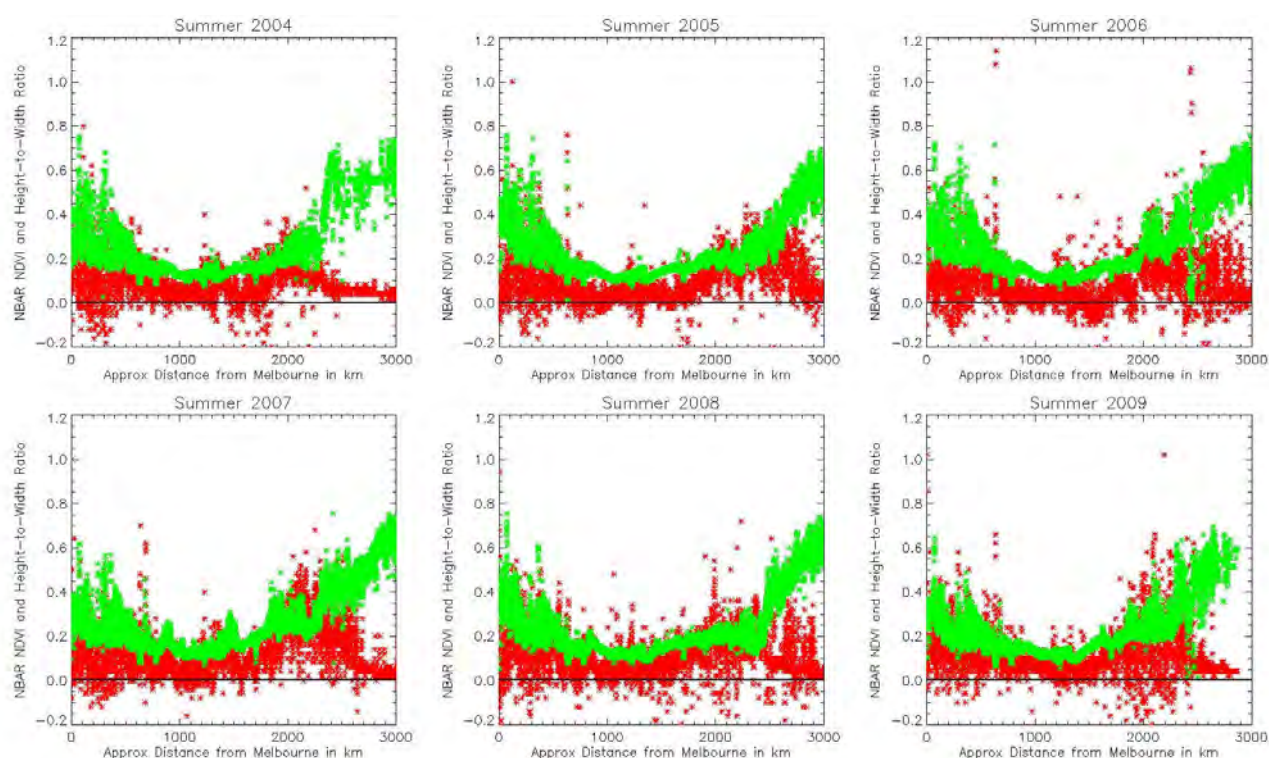


Figure 2: NBAR NDVI (green) and derived Height-to-Width ratio (red) on the y-axis plotted against distance along the transect from Melbourne on the x-axis. Epochs are for the last week of January each year from 2004 to 2009, in summer. Only pixels where the correlation is greater than 0.9 between the model and NDVI derived from MODIS BRDF representation have been included in the plots.

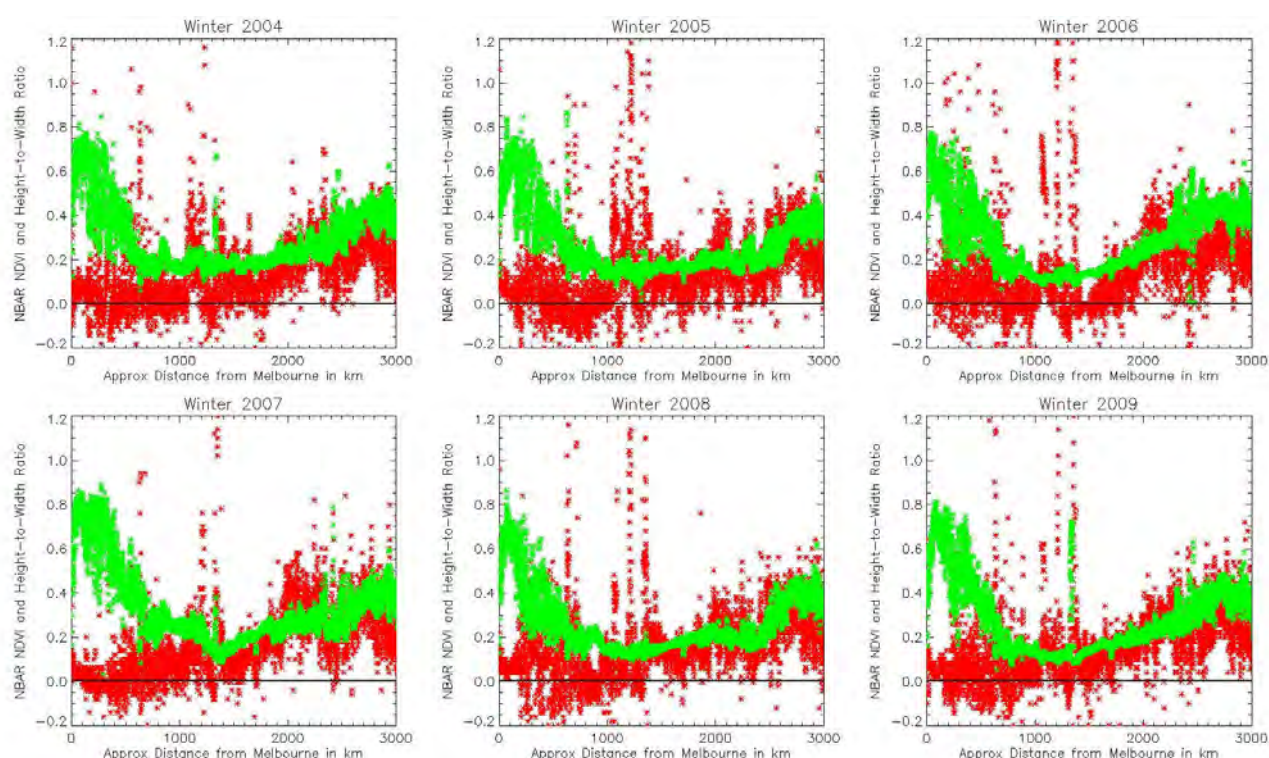


Figure 3: NBAR NDVI (green) and derived Height-to-Width ratio (red) on the y-axis plotted against distance (x-axis) along the transect from Melbourne. Epochs are for the first week of July each year from 2004 to 2009, in winter. Only pixels where the correlation is greater than 0.9 between the model and NDVI derived from MODIS BRDF representation have been included in plots.

Nadir BRDF Adjusted Reflectance (NBAR) NDVI derived from the MODIS MCD43 product has also been determined for pixels along the transect. Within Figures 2 and 3, NBAR NDVI exhibits the expected profile along the transect: falling NBAR NDVI moving north-west from Melbourne, low NBAR NDVI across the centre of Australia and the Simpson Desert and rising again in tropical areas approaching Darwin. NBAR NDVI was higher in winter than during summer nearer to Melbourne. The opposite pattern is observed as the transect approaches Darwin: higher NBAR NDVI during the summer wet season and lower NBAR NDVI during the winter dry season.

3.1 Discussion

The results in Figures 2 and 3 display a number of general relationships, although the data, in particular the height-to-width ratio, is noisy at the pixel scale. The derived height-to-width ratio has a positive relationship with NBAR NDVI for values less than 0.4; that is as NBAR NDVI increases the height-to-width ratio also increases. For NBAR NDVI greater than 0.4, there is a negative relationship with the height-to-width ratio: as NBAR NDVI increases beyond 0.4, the height-to-width ratio decreases. The negative relationship between the height-to-width ratio and NBAR NDVI is particularly evident in the winter epochs nearer to Melbourne and in the summer epochs nearer to Darwin (refer Figures 2 and 3). This relationship can also be seen by plotting the derived height-to-width ratio against NBAR NDVI as shown in Figure 4.

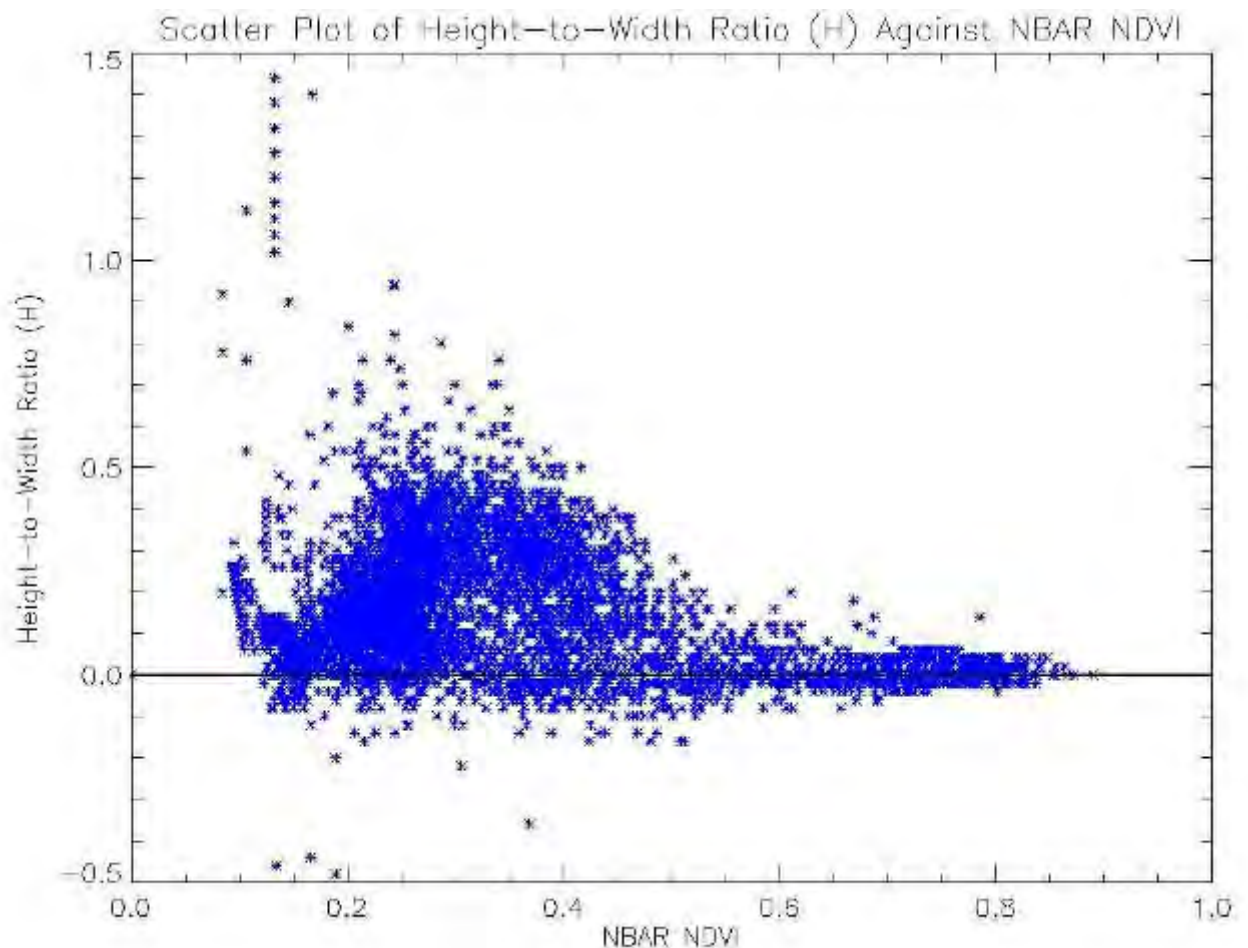


Figure 4: Derived height-to-width ratio (y-axis) plotted against NBAR NDVI (x-axis) for pixels along the Melbourne – Darwin transect. The selected epoch being the first week of July 2007, in winter.

An interpretation of the negative relationship is that as the vegetation becomes denser and the canopy approaches closure, the soil layer becomes less visible at progressively larger viewing angles. The mutual obscuration of vegetation components is not accounted for by the model and

thus for vegetation densities where this begins to occur (NBAR NDVI > 0.4) the resulting height-to-width ratio approaches zero and the results lose validity. Therefore application of this model is limited to surfaces with an NBAR NDVI response less than 0.4.

Perhaps more significant are features at specific locations along the transect that appear to have the converse of the general relationship between the height-to-width ratio and NBAR NDVI. For winter epochs, at approximately 600km, 1,100km, 1,200km and 1,300km along the transect from Melbourne there are four large positive spikes in the height-to-width ratio. These appear without significant corresponding changes in NBAR NDVI. Spatially these align to the Willandra lakes in Mungo National Park, Tooncatchyin Creek, Cooper Creek and Warburton Creek which feed into Lake Eyre. A similar effect, but less well defined occurs around the Davenport range in the Northern Territory (approx 2,000km from Melbourne).



Figure 5: Location of Tooncatchyin Creek, Cooper Creek and Warburton Creek along the Melbourne – Darwin transect. The red line indicates the path of the transect.

These spikes in the height-to-width ratio are more prominent in winter. During summer epochs the increases in the height-to-width ratio at these locations are significantly less pronounced, but it is still possible to detect as a general rise in the height-to-width ratio. The spatial extent and degree of definition of these spikes in the height-to-width ratio varies between years, being more extensive in winter 2005 and less extensive in winter 2009 when there appears a corresponding significant increase in NBAR NDVI in a neighbouring location around Warburton creek.

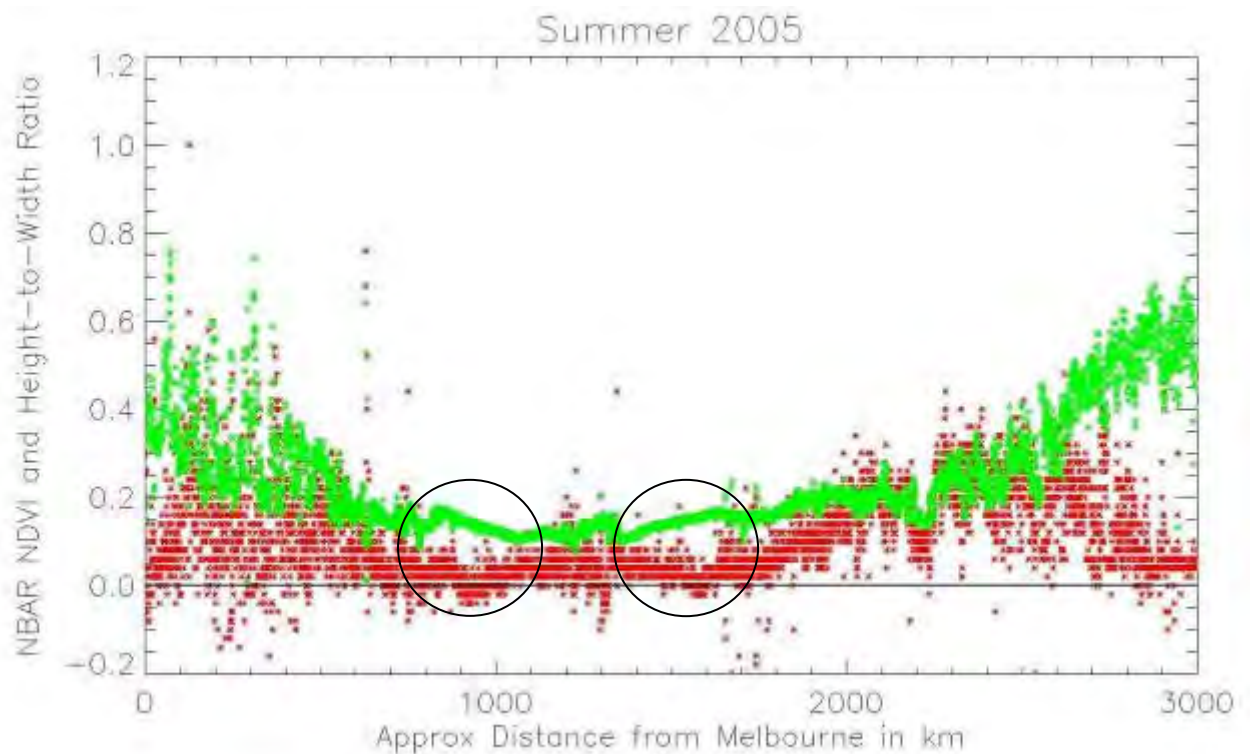


Figure 6: NBAR NDVI (green) and derived Height-to-Width ratio (red) on the y-axis plotted against distance from Melbourne on the x-axis. Figure 6 is an enlargement of the plot for the same epoch shown in Figure 2.

In contrast to the winter spikes in the height-to-width ratio, the height-to-width ratio is consistently low during summer in areas 800 – 1,000km and 1,400 – 1,700km along the transect from Melbourne. These locations are indicated by the black circles in Figure 6. They correspond spatially to an area south-west of the Strzelecki Desert and the Simpson Desert through which the transect passes. These areas are arid and may generally be expected to have very low vegetation cover. Additionally the variance of the height-to-width ratio at these locations is comparatively low compared to the variability in the height-to-width ratio shown at other locations along the transect. This effect can be seen at the same locations at other epochs, but is less pronounced than that shown for the epoch in Figure 6. The slight rise in the height-to-width ratio in between the areas of the Strzelecki and Simpson deserts corresponds to the Tooncatchyin, Cooper and Warburton Creeks features discussed above, however appearing in Figure 6 during summer when the variation height-to-width ratio is less pronounced.

It is difficult to fully understand or explain the underlying cause of height-to-width ratio variations at these locations without specific surface cover knowledge at these particular epochs (for example rainfall and flood events). However where the height-to-width ratio spikes, the corresponding surface feature is associated with the presence of a water course or lake. Where the height-to-width ratio is consistently low in comparison to NBAR NDVI, the corresponding surface feature is related to an absence of water courses or lakes. This suggests that these particular variations in the height-to-width ratio may be associated with the presence of vegetation in an otherwise arid landscape, although with an insufficient coverage to influence a significant change in NBAR NDVI. The height-to-width ratio may be sufficiently sensitive and able to amplify the presence of low density vegetation from the background soil that is only apparent when viewed obliquely. Consistent with this, when there is effectively no vegetation cover present (for example sand dunes in the Simpson Desert), NDVI is invariant to viewing angle and the derived height-to-width ratio will be close to zero, indicating that the NBAR NDVI response is from bare soil alone. The height-to-width ratio may therefore be used in arid regions to determine if the NDVI response is from bare soil or if low density vegetation cover is also present and contributing to the NDVI response.

4 CONCLUSION

This paper describes a model for interpreting BRDF effects associated with NDVI and has demonstrated the process with application to a transect from Melbourne to Darwin, Australia. The model is theoretically based on geometric optical properties and empirically provides a good fit to NDVI changes with viewing angle as derived from the MODIS BRDF representation.

The model is able to identify features not appearing in reflectance data alone, in particular, the height-to-width ratio identifies low vegetation cover from background soil when no significant NDVI response is observed. This may allow separating the NDVI responses between background soil and vegetation.

Further analysis is needed to better understand surface conditions at particular epochs. Alternatively, understanding of the approach might be assisted by application to other better understood study surfaces. Furthermore, the relationship between NDVI and the height-to-width ratio needs quantitative definition.

5 APPENDIX 1

5.1 Geometric Optical Model Formulation

In order to quantifiably interpret the variations in NDVI with respect to changes in viewing angles a geometric optical model is proposed. A pixel may be considered to be a mix of soil components and vegetation components. If a pixel is defined as having a dimensionless area of 1 (that is 1×1), a random distribution of vegetation prisms may be considered on the surface having dimensions $d \times d \times (H \times d)$ where; 'd' is a dimensionless value between 0 and 1 representing the length and width of the vegetation prism in the horizontal plane and 'H' is a dimensionless height-to-width ratio of the vegetation. A height-to-width-ratio of 1 represents cubes of vegetation, values greater than 1 being taller elongated vegetation such as pine trees and values less than 1 being low and flat vegetation. This enables vegetation cover within the pixel to be defined as a fraction of the pixel's dimensions. The use of prism shaped objects to represent vegetation has the benefit of computational simplicity and the shape of protrusions should not be important with prisms having been used in past BRDF modelling (Roujean et al., 1992). Larger viewing angles will bring the vertical surface of the vegetation into view and obscure an area of soil.

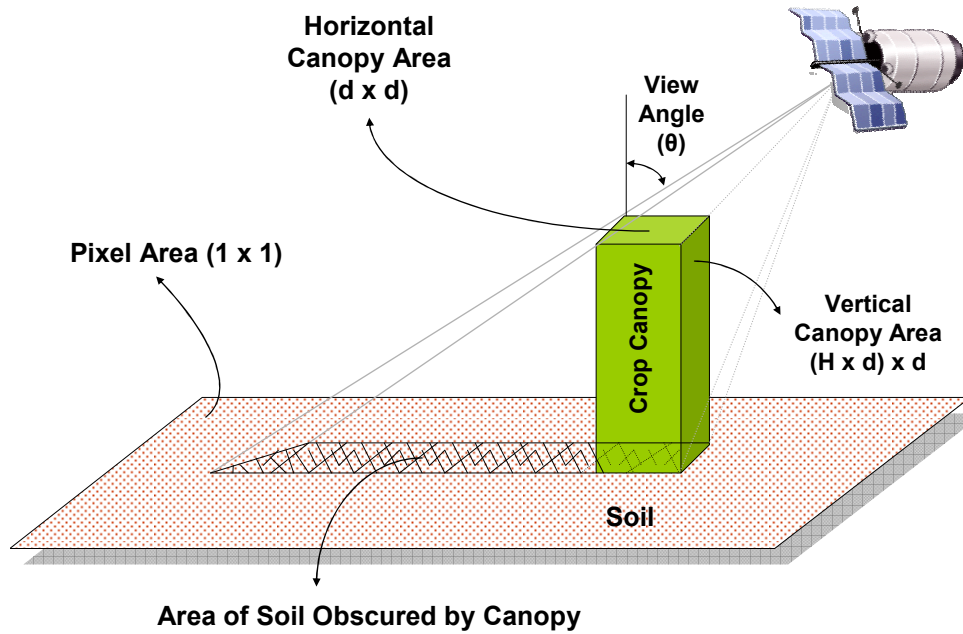


Figure 7: A single 'prism' of vegetation is depicted on a flat surface of soil with an area of soil obscured via the vegetation prism based on the viewing geometry from the sensor.

Based on the geometry of these surfaces, an expression for changes in NDVI, as a function of viewing angle can be described as the sum of the horizontal and vertical surface components:

$$\text{NDVI}_{\text{Total}}(\theta) = \text{NDVI}_{\text{horizontal}}(\theta) + \text{NDVI}_{\text{vertical}}(\theta) \quad (2)$$

$$\text{NDVI}_{\text{horizontal}}(\theta) = D \text{NDVI}_{\text{Veg.}}(\theta) + (1-D) \text{NDVI}_{\text{Soil}}(\theta) \quad (3)$$

$$\text{NDVI}_{\text{vertical}}(\theta) = H D \tan(\theta) [\text{NDVI}_{\text{Veg.}}(\theta) - \text{NDVI}_{\text{Soil}}(\theta)] \quad (4)$$

where: H is the height-to-width ratio of vegetation which is constrained to an arbitrary range from - 2 to + 2, $D = d^2$ and represents the density of vegetation cover with a range $0 \leq D \leq 1$, with value 0 being for all soil and no vegetation and 1 being for complete vegetation coverage and no soil; and θ is the viewing angle.

Equation (3) is the linear combination of soil and canopy vegetation index on the horizontal plane for the case when only the horizontal vegetation surfaces are present ($H = 0$). Equation (4) is the additional contribution that the vertical plane of the canopy makes to the observed vegetation index as a function of the viewing angle including the additional soil obscuration relative to the $H = 0$ case. Total vegetation index response (equation (2)), as a function of viewing angle is therefore the sum of equations (3) and (4). The model assumes that the vegetation prisms do not overlap from the satellite viewpoint and neglects shadows. If viewed at zenith ($\theta = 0$) or the vegetation has no vertical profile ($H = 0$), then equation (4) becomes zero. Increases in NDVI as a consequence of higher viewing angles will be greatest when:

- the vegetation index difference between canopy and soil is greatest, and/or
- the height-to-width ratio of the vegetation is large (that is larger values of H), and/or
- the density of vegetation is higher (that is larger values of D).

The model is independent of the number of vegetation prisms present within a pixel. For a given density (D) and height-to-width ratio (H), a single large prism of vegetation or a large number of small prisms (as would realistically be expected) will make visible the same surface areas of soil and vegetation at the same viewing angle. The volume that the vegetation occupies as a fraction of the pixel's dimensions in 3-D space will be a function of the number of individual vegetation prisms present in a pixel, density (D) and the height-to-width ratio (H). However, consideration of the number of individual vegetation prisms can be ignored, as the resulting vegetation volume is not considered within this model.

Values for the two model parameters H and D may be determined by numerical methods that provide a best fit (lowest root mean squared error (RMSE)) between the model and vegetation index values derived from the MODIS BRDF parameters across a range of viewing angles 0° to 57°; this being at the upper limit of MODIS's angular sampling (Wanner et al., 1997). In applying this model, values for vegetation indices derived from the MODIS BRDF representation across a range of viewing angles effectively become 'observations'. The model's approach seeks to replace the wavelength-specific isometric, volumetric and geometric parameters within the MODIS BRDF model with the alternate parameter set of: $NDVI_{veg}$ $NDVI_{soil}$ vegetation density (D) and a height-to-width ratio (H) which offer more direct interpretation of vegetation structure.

The proposed model may be considered a re-expression of the BRDF effects into an alternate, more readily interpretable set of parameters. The model may also be thought of as a mixing model, where the end members are soil and vegetation and the determination of the component quantities are based upon both spectral and angular effects.

In applying this model, a representative characteristic profile for $V_{soil}(\theta)$ and $V_{veg}(\theta)$ must first be determined. This acknowledges that soil and vegetation end-members also have angular profiles that vary with viewing angle. This approach assumes that there are two, and only two component elements present in each pixel. The end-member (that is $NDVI_{veg}(\theta)$ and $NDVI_{soil}(\theta)$) may be obtained from observations (satellite or in situ), although in practise this may be difficult in all but a very select number of situations. An example where satellite observations may allow determination of the end-members is for sufficiently large areas of single species cropped fields where observations made during periods of bare soil prior to planting and of the mature crop canopy prior to harvest may be used for determination of the end-members in the model. The characteristic profiles for soil and vegetation should be uniquely defined for each pixel. A simple expression that provided a good fit for the NDVI profile of most surfaces was found in the form of:

$$NDVI_{soil}(\theta) = A_{soil} + B_{soil} \theta + C_{soil} \tan(\theta) \quad (5)$$

$$NDVI_{veg}(\theta) = A_{veg} + B_{veg} \theta + C_{veg} \tan(\theta) \quad (6)$$

Values for 'A', 'B' and 'C' can be derived that provide a best fit (minimum RMSE) for representative samples of soil and crop derived from the MCD43 product across a range of viewing angles. The end-member angular signatures expressed by equations (5) and (6) can be substituted into equations (3) and (4). Expressions in other forms (for instance a polynomial in θ) may equally provide good characteristic profiles for soil and crop canopy as there is no interpretation implied from the parameters or in the structure of equations (5) and (6). In application of this model as described in this paper, equations (5) and (6) are set equal to 0 and 1 respectively.

Application of this model is based on changes in viewing angle only. Values for the solar zenith angle and relative azimuth must be determined and fixed when generating 'observations' using MODIS's BRDF representation which the model seeks to interpret. In application of this model a 0° solar zenith angle and a 0° relative azimuth have been used. Using 0° for the solar zenith angle will make consideration of the relative azimuth selected irrelevant. A 0° solar zenith angle and 0° viewing angle will represent the 'hot spot', which is the angularly narrow reflectance enhancement around the antisolar direction. The range of viewing angles from 0° to 57° used for deriving the model parameters represents consistent viewing away from the hot spot without the need to consider viewing angles that will cross the hot spot. This enables the model to remain relatively simple in its formation. The MODIS BRDF representation is reciprocal between solar and viewing angle, therefore whilst a 0° solar zenith angle may not have been observed by MODIS in

deriving the BRDF representation, a 0° viewing angle may have been observed and will yield the same BRDF result. Past studies have shown that linear BRDF models can be stably extrapolated to zero zenith angles.

REFERENCES

- GAO, F., SCHAAF, C. B., STRAHLER, A. H., JIN, Y. & LI, X. 2003. Detecting vegetation structure using kernel based BRDF model. *Remote Sensing of Environment*, 86, 198 - 205.
- GRANT, I. F. 2000. Investigation of the Variability of the Directional Reflectance of Australian Land Cover Types. *Remote Sensing Reviews*, 19, 243 - 258.
- HILL, J. M., AVERILL, C., JIAO, Z., SCHAAF, C. B. & ARMSTON, J. D. 2008. Relationship of MISR RPV parameters and MODIS BRDF shape indicators to surface vegetation patterns in an Australian tropical savanna. *Canadian Journal of Remote Sensing*, 34, S247 - S267.
- LI, F., JUPP, D., REDDY, S., LYMBURNER, L., MUELLER, N., TAN, P. & ISLAM, A. 2010. An Evaluation of the Use of Atmospheric and BRDF Correction to Standardise Landsat Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
- LOVELL, J. L. & GRAETZ, R. D. 2001. Analysis of POLDER-ADEOS data for the Australian continent: the relationship between BRDF and vegetation structure. *Journal of Remote Sensing*, 23, 2767 - 2796.
- LUCHT, W. & LEWIS, P. 2000. Theoretical noise sensitivity of BRDF and albedo retrieval from the EOS-MODIS and MISR sensors with respect to angular sampling. *International Journal of Remote Sensing*, 21, 81-98.
- LUCHT, W., SCHAAF, C. B. & STRAHLER, A. H. 2000. Algorithm for the Retrieval of Albedo from Space Using Semiempirical BRDF Models. *IEEE Transactions on Geoscience and Remote Sensing*, 38.
- PAGET, M. & KING, E. 2008. *MODIS Land Products for Australia CSIRO Marine and Atmospheric Research Internal Report No. 004, CSIRO Marine and Atmospheric Research Canberra, Australia. ISBN 978192142342 (pdf)* [Online]. Available: <https://rs.nci.org.au/lpdaac/> [Accessed].
- ROUJEAN, J.-L., LEROY, M. & DESCHAMPS, P.-Y. 1992. A Bidirectional Reflectance Model of the Earth's Surface for the Correction of Remote Sensing Data. *Journal of Geophysical Research*, 97, 20,455 - 20,468.
- SANDMEIER, S., MULLER, C., HOSGOOD, B. & ANDREOLI, G. 1998. Physical Mechanisms in Hyperspectral BRDF Data of Grass and Watercress. *Remote Sensing of Environment*, 66, 222 - 233.
- SCHAAF, C. B., GAO, F., STRAHLER, A. H., LUCHT, W., LI, X., TSANG, T., STRUGNELL, N. C., ZANG, X., JIN, Y., MULLER, J.-P., LEWIS, P., BARNSLEY, M., HOBSON, P., DISNEY, M., ROBERTS, G., DUNDERDALE, M., DOLL, C., D'ENTREMONT, R. P., HU, B., LIANG, S., PRIVETTE, J. L. & ROY, D. 2002. First operational BRDF, albedo nadir reflectance product from MODIS. *Remote Sensing of Environment*, 135 - 148.
- TIAN, Y., ROMANOV, P., YU, Y., XU, H. & TARPLEY, D. 2010. Analysis of Vegetation Index NDVI Anisotropy to Improve the Accuracy of the GS-R Green Vegetation Fraction Product.
- WANNER, W., STRAHLER, A. H., HU, B., LEWIS, P., MULLER, J.-P., X, L., BARKER SCHAAF, C. L. & BARNSLEY, M. 1997. Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR data: Theory and algorithm. *Journal of Geophysical Research*, 102, 17,143 - 17,161.

Mapping Coastal Carbon Sinks in South Australia

Matt Miles, Darcy Peters and Peter Fairweather

SA Department of Environment and Natural Resources
GPO Box 1047, Adelaide. South Australia 5001
matthew.miles@sa.gov.au

Abstract

The aim of this project was to assess the application of existing habitat mapping to estimate carbon sequestration rates in coastal and marine environments in South Australia. Following on from a global summary of coastal carbon quantification science (Laffoley and Grimsditch 2009), this project sought out Southern Australian literature and local knowledge to investigate coastal carbon sequestration in maps. By telling stories about the dynamics of carbon in the coastal region, lessons can be learned about what parts of ecosystems we can manage better or re-establish in order to minimise atmospheric carbon levels.

The project identified the main processes that drive carbon sequestration rates in three broad ecosystems: saltmarshes, mangroves and seagrass meadows. These coastal and benthic marine habitats were matched to field verified mapping from aerial photography compiled over the last 15 years. From available Australian literature, carbon sequestration rates were assigned, mapped and summarised inside and outside of proposed Marine Protected Areas.

The study has estimated a carbon sequestration contribution for these ecosystems, such that comparisons with terrestrial systems can be approached while also providing a window into policy opportunities for conservation and climate change outcomes. While marine carbon is currently excluded from IPCC carbon accounting methods, this study helps our understanding of the issues involved and contributes to a monitoring and research agenda in this emerging field.

KEYWORDS: coast, marine, carbon sequestration, climate change, habitat mapping

1 Introduction

The contribution of "Blue Carbon" (i.e. carbon in aquatic ecosystems) to climate change, is being discussed in literature and media. The World Bank has recently released a report titled "Mitigating Climate Change through Restoration and Management of Coastal Wetlands and Near-shore Marine Ecosystems," which finds that drainage and degradation of coastal wetlands emits significant amounts of carbon dioxide directly into the atmosphere and leads to decreased carbon sequestration (Crooks, et al 2011). The report emphasised the need for: protecting coastal wetlands; creating incentives for avoiding/reversing their degradation; and including the protection of these ecosystems in carbon emission reduction strategies and climate negotiations. Investment in these activities can link to other climate strategies and carbon financing mechanisms, provided that protocols on accounting, verification and reporting of net carbon uptake can be agreed.

This paper reports on a project undertaken in South Australia (SA), to utilise coastal and marine benthic habitat mapping to quantify carbon flows and elucidate factors influencing sequestration potential. This was done by aligning available sequestration rates with mapping of four ecosystem types; saltmarshes, mangroves, seagrass and kelp.

By developing these systems to measure carbon, we can articulate principles for carbon emission mitigation strategies and contribute knowledge and understanding of information gaps to conservation policy and research.

2 SALTMARSH

There are 21,676 ha of tidally influenced saltmarshes in Department of Environment and Natural Resources (DENR) mapping along the coasts of South Australia.

Saltmarshes are intertidal ecosystems dominated by low growing herbs, shrubs and grasses tolerant of high salinity and poorly aerated soils (CRC-CZEWM, 2004). In these ecosystems, it is not vegetation growth that drives carbon sequestration, rather it is the accretion of carbon rich soil sediments via tidal transport (Howe et al, 2009). In the South Australian context, one range of rates was available from an assessment made of coastal systems north of Adelaide (Cook, 2009). The value selected (see Table1) is at the high end of the range for mid to low marshes (0.64 to 2. tC/ha/yr2) in that assessment.

Intertidal marshes on delta flats can receive carbon from river sediment (including flood events) as well as daily tidal inundation depositing sediment from sea water. Saltmarshes open to the sea will have sequestration rates depending on amount of suspended particles in the water column and the speed and depth of inundation (including storm events). Sediment accretion rates in SA saltmarshes are not well known but techniques exist such as sediment elevation tables and spreading of chalk dust to mark stratum levels.

An important feature of saltmarshes regarding net sequestration is that they release only negligible amounts of methane, a powerful greenhouse gas; therefore, the carbon storage benefits of tidal salt marshes are not reduced by methane production as can be the case for freshwater wetlands (Trulio et al, 2007).

3 MANGROVES

There are 15,190 ha of tidally influenced mangroves in DENRs mapping along the coasts of SA.

Mangrove communities in SA are populated by one species - *Avicennia marina* (mangrovetree.org, 2011). Known to be among the southernmost occurrences of mangroves on the globe, they are generally found seaward of saltmarshes and often interspersed with seagrass beds and mudflats.

Carbon sinks to consider in these systems are associated with burial of Carbon in sediments, and net growth of forest biomass. Burial in the soil profile represents long term sequestration, whereas increase in biomass through expansion or (re)planting is relevant only in the shorter term (e.g. tens of years) (Laffoley & Grimsditch, 2009).

Howe et al (2009) makes the observation that landward migration of mangrove systems due to sea level rise is likely to displace saltmarsh communities that are often constrained by infrastructure such as coastal roads and levees. With barriers restricting access to areas for the saltmarshes to colonise, this change in vegetation cover may represent a net loss of sequestration capacity as mangrove sequestration rates are 66% of the rate for saltmarshes.

The value of 139 g/m²/yr is used in this study based on a number of studies across the world using various methods (Laffoley & Grimsditch, 2009). More research is needed to ascertain the confidence this value has in representing SA mangrove areas.

4 SEAGRASS

There are 685,744 ha of seagrass meadows in DENRs mapping off the coasts of SA

A report for the Adelaide Coastal Waters Study (Moore and Westphalen 2007) describes seagrass ecosystems as amongst the most productive plant systems on the planet on a per unit basis. Again with a paucity of local literature to draw on, this study sought to describe sequestration potential in these systems as well as quantify impacts from seagrass losses in South Australian waters.

Seagrass species such as *Posidonia australis* produce “matte” formations that are estimated to store between 30 and 50% of the plants biomass below ground. This underground root system is likely to store carbon for “decades to centuries if the environmental conditions are maintained and seagrass plants continue to live above the stored carbon” (Moore and Westphalen 2007).

Sequestration rates of 68-120 g/m²/yr are described by Moore and Westphalen based on a lower limit of 10% net primary production sequestered into long term storage (in matte formations) and an upper limit of 18%. These rates broadly match with the average rates for seagrasses of 83 g/m²/yr given in Laffoley & Grimsditch (2009) based on Mediterranean species including *Posidonia oceanica*. This project used rates from the local study focussed on local species, to generate upper and lower estimates in favour of the Laffoley & Grimsditch rate.

5 KELP

A number of kelp species occur in SA, particularly in the cooler waters of the South East coast, such as the large kelp (*Macrocystis angustifolia*) and the bull kelp (*Durvillaea potatorum*). Sheltered waters and inlets of more northern waters also support a number of subtropical species where temperatures are high enough for them to survive (PIRSA 2003). While some mapping of reef habitats has occurred, it is insufficient in extent and lacks recording of the vegetated state to assist quantitative assessment of overall sequestration potential.

Laffoley & Grimsditch (2009) describes *Macrocystus* net primary production of 670 – 1300 g/m²/yr or 6.7 – 13 t/ha/yr. This indicates high potential for carbon storage through restoration and protection of kelp forests – and is included to provide a comparison with the other communities described. However, as mentioned above, area estimates and therefore sequestration totals are not included in this project.

6 METHOD

The main source of sequestration rates comes from a publication on “The management of natural coastal carbon sinks” by Laffoley & Grimsditch (2009). This identifies 5 coastal habitat types: Tidal Salt Marshes, Mangroves, Seagrass Meadows, Kelp Forests and Coral Reefs. Four of these occur in and around the South Australian coast (all but Coral Reefs). The four habitat types (hereafter referred to as **carbon classes**) were assigned the sequestration rates in Table 1. These were initially based on Laffoley & Grimsditch (2009) with a range of publications used to validate or modify if a more local study was available (e.g. Cook, 2009 and Moore and Westphalen, 2007).

carbon class	tC/ha/yr	tCO2-e/ha/yr
saltmarsh	2.1	7.7
mangrove	1.39	5.1
seagrass	0.68 - 1.2	2.5-4.4
kelp	6.7 – 13	24.5-47.7

Table 1: shows annual sequestration rates per hectare for each carbon class in metric tonnes of carbon and carbon dioxide equivalents.

Kelp forest mapping is currently being trialled with the use of remote sensing but was not comprehensive enough to be included in the quantitative mapping parts of the study. The three relevant carbon classes were matched to existing habitat mapping data stored in the corporate databases of DENR. These datasets have been collected at various scales over the past decade and included:

- Benthic Habitats at 1:100,000 scale (2008-2010)
- Estuary Habitats at 1:30,000 (2009) and
- Saltmarsh/Mangrove Habitats 1:40,000 to 1:100,000 (2005-2010).

Based on coverage and accuracy, the three layers of habitat mapping were prioritised to build up the best resolution of mapping to represent the carbon classes. Within the mapped layers, attributes recording a percentage ‘coverage’ or ‘density’ assist in describing whether a full sequestration rate be used or a pro rata one. E.g. where a density is recorded as “50-70%”, then

the mid point 60% of rate is applied, or where coverage is recorded as “patchy”, then 50% is used. In addition, saltmarsh habitats occurring above tidal influence were removed as they have much reduced capacity for sequestration based on soil accretion as noted above.

7 RESULTS

The project produced areas of carbon classes and annual sequestration rates. These were presented in a statewide summary as well as broken down for each proposed Marine Protected Area.

Table 2 shows that using existing mapping across the whole of SA, the carbon classes described potentially sequester around half a million tonnes of Carbon per year. This equates to over 2 million tonnes of carbon dioxide equivalent units. Seagrass beds account for 95% of that.

The two map outputs presented here show upper estimate rates for the Upper Spencer Gulf Marine Park Area (Figure 1) and the upper estimate rates for the central coast and gulfs region of SA (Figure 2). These show where pro rata rates are assigned to areas of lesser ‘coverage’ or ‘density’ and where carbon class mapping is absent.

Carbon Class	Area (ha)	% of total Carbon Classes	Sequestration Range			
			lower (tC/yr)	upper (tC/yr)	lower (tCO ₂ -e/yr)	upper (tCO ₂ -e/yr)
Mangrove	15190	2	18,809	18,809	69,030	69,030
Saltmarsh	21676	3	40,295	40,295	147,882	147,882
Seagrass	685744	95	296,443	522,313	1,087,944	1,916,889
Total	722610	100	355,547	581,417	1,304,856	2,133,800

Table 2: total sequestration rates, in tonnes of carbon (tC) and tonnes of carbon dioxide equivalent units (tCO₂-e) per year per carbon class for whole of SA.

8 LIMITATIONS / RESEARCH OPPORTUNITIES

It should be noted that overall, the accuracy and resolution of sequestration estimates used in this study are at the lower end of the scale. They are useful for comparing different habitats and areas, suitable for educational and public relations purposes. To improve this accuracy, a number of issues need investigating; one main area here is concerned with the detail of carbon flow/process models, and how well they are represented by the attributes of the habitat mapping. For example, how well do percentages of cover describe a reduction in sequestration capacity? What Net Primary Production (NPP) rates are appropriate? How can we account for microbial and animal biomass associated (such as calcareous epiphytes) within these ecosystems?

Another gap is in understanding local saltmarsh and mangrove habitats with respect to sequestration rates, e.g. measuring soil accretion rates within networks of monitoring sites would provide locally tested measures.

A third area is more of a technical challenge in how to maintain repeatable ways of mapping carbon classes in coastal environments, in order to monitor not only broad loss of habitats but changes in density or condition.

During the development of this project, a number of other areas related to carbon sequestration were discussed. While too nebulous and under-studied to include even in this initial quantification, they highlight areas for future research in understanding the dynamics of carbon in the SA coastal environment and potential opportunities to assist in climate change mitigation. Loosely relating to the field of biogeochemistry, it was noted that this is an area of study currently with few participants.

Very briefly, these topics included coastal and benthic areas such as un-vegetated sea beds, mud flats, reefs, shell fish beds (e.g. oysters, razorfish, scallops, mussels, etc.), carbonates in subtidal sandy and limestone formations.

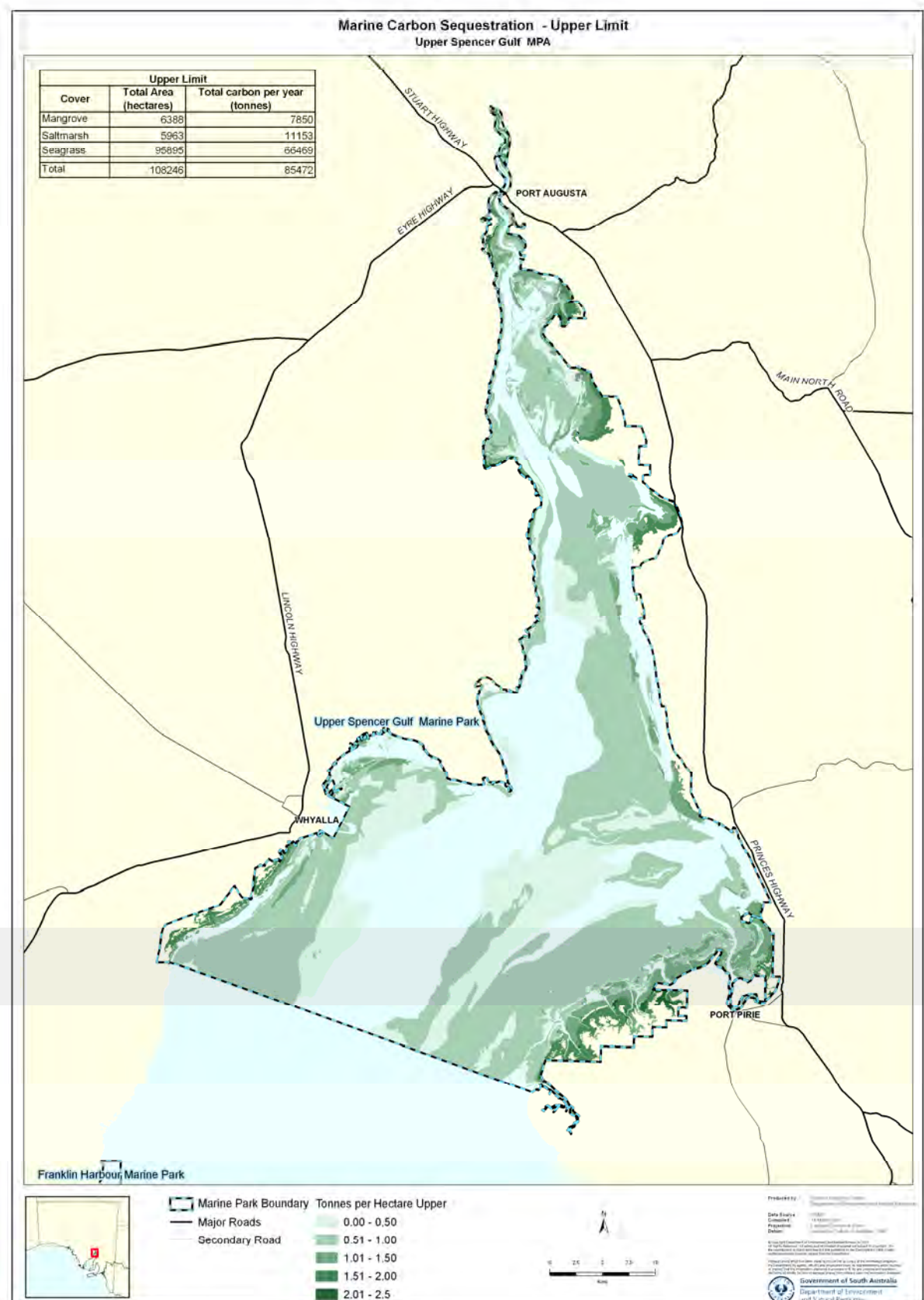


Figure 1: sequestration rates of Upper Spencer Gulf



9 CARBON ACCOUNTING

Given the limitations and opportunities described above, it is clear that we are some way off standardised measures for carbon accounting in coastal environments. Moves are underway however to develop such tools around the world. For example a new methodology for calculating mangrove carbon storage, has now been adopted under the Clean Development Mechanism (CDM) of the UN Framework Convention on Climate Change (UNFCCC) Kyoto Protocol. This method describes carbon stocks and changes to degraded mangrove systems in an accountable way to attract investment for large scale regeneration and planting (IUCN, 2011).

It is envisaged that further coastal systems will have similar tools evaluated and developed over coming years for even wider use, as responses to the challenges and opportunities presented here are mounted.

This project showed that there are varying levels of understanding in the measurement techniques for above and below ground carbon in coastal ecosystems (Cook 2009). This is no doubt mirrored across the range of terrestrial ecosystems providing further challenges for comparative measures between conservation, remedial and direct action activities relating to carbon sequestration.

A further consideration in comparing total impacts of activities across ecosystems is the risk associated with longevity of sequestration. E.g. the value of terrestrial biosequestration effort can be severely reduced by fire events, and coastal systems contain certain risks associated with sea level rise reducing or compromising potential gains.

It is important to note that greenhouse gas benefits are not a sole reason for comparing the value of ecosystem protection strategies. Rather, they add to biodiversity or ecosystem service benefits that are currently encapsulated in strategies such as NatureLinks and No Species Loss. The treatment of “biodiverse carbon” is an emerging field currently focusing on terrestrial ecosystems. A key challenge is to develop biodiversity metrics to work in concert with carbon metrics (Connor and Patterson, 2011).

10 CONCLUSION

This study shows that coastal and benthic environments should be considered when developing carbon sequestration policies and accounting mechanisms. Numerous areas of research remain in order to better understand the carbon dynamics in these environments, however it is clear that degradation or loss of such systems is not only an issue for conservation objectives and ecosystem services but also for carbon emission mitigation and sequestration opportunities. In all, they are an integral part of the system, so we should seek to understand them.

It should be stressed that coastal systems such as sea grass meadows should not as a result of this work be oversold as an answer to climate change, rather that carbon sequestration services in these ecosystems is a further reason to preserve their natural processes. Natural systems are working as hard as they can, any degradation to a carbon sink will require even greater reductions in emitting activity to achieve the goal of net reductions to alleviate effects of climate change.

In this study, published sequestration rates have been combined with existing habitat mapping, to quantify annual Carbon sequestration. Within all Marine Park Areas, estimates have been produced of between 200,000 and 300,000 tonnes of Carbon sequestered per year, this increases to around 500,000 tonnes when including mapped areas for the whole of SA. Converted to carbon dioxide equivalent units, this state-wide upper estimate equates to over 2 million tonnes CO₂-e.

Maps show how the estimates vary across the Marine Protected Areas in relation to density of vegetation as stored in habitat mapping databases. There are obviously assumptions built into this analysis particularly around how well mapped areas match the models of measured sequestration rates from the literature. However as a proof of concept, the study has shown potential to assist carbon accounting processes, visualisation and communication.

It can be argued that coastal carbon sinks remove more carbon from the atmosphere on a per hectare basis than freshwater wetlands due to the lack of methane generated in these saline

environments offsetting the carbon capture benefits. Such claims must be considered carefully within any carbon accounting system as selective inclusion or exclusion can influence concepts of net impacts markedly.

This study has provided a broad brush approach with scope for refinement while suggesting numerous areas for research opportunities and potential policy response. As we consider the introduction of carbon markets and the pursuit of conservation activity along our coasts in the challenges against climate change, such methods and tools will need refinement, adoption and support. This project helps to identify some of the principles that may contribute to such methods.

References

Cook FS, 2009. "Review of salina carbon inventory information, including biomass" Consultant report, Delta Environmental Consulting, Adelaide.

CRC-CZEWM, 2004. "Where River Meets Sea: Exploring Australia's estuaries" Cooperative Research Centre for Coastal Zone, Estuary and Waterway Management. CSIRO Publishing

Crooks, S., Herr D., Tamelander J. , Laffoley D, and J. Vandever. 2011. "Mitigating Climate Change through Restoration and Management of Coastal Wetlands and Near-shore Marine Ecosystems: Challenges and Opportunities." Environment Department Paper 121, World Bank, Washington, DC. http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&theSitePK=523679&entityID=000333038_2010407024117&searchMenuPK=64187283&theSitePK=523679

Howe A. J. Rodríguez J.F. and P.M. Saco. "Surface evolution and carbon sequestration in disturbed and undisturbed wetland soils of the Hunter estuary, southeast Australia", Estuarine, Coastal and Shelf Science., vol. 84, pp.75-86, 2009.

IUCN, 2011. <http://iucn.org/?7595/Mangroves-to-receive-huge-boost-from-new-carbon-credit-rules> viewed 19 July 2011.

Laffoley, D.d'A. & Grimsditch, G. (eds). 2009. "The management of natural coastal carbon sinks." IUCN, Gland, Switzerland. 53 pp

MangroveWatch Australia, 2011. Mangroves of South Australia. http://www.mangrovetwatch.org.au/index.php?option=com_content&view=category&layout=blog&id=85&Itemid=300202 viewed 15 March 2011

Moore, T. and Westphalen, G. 2007. *Australian seagrass meadows as potential carbon sinks. A report for the Environment Protection Authority*, Adelaide

O'Connor, P.J. and Patterson, S. (2011). *Establishing a National Biodiversity Co-benefits Standard for investments in the Carbon Market*. O'Connor NRM PTY LTD.

PIRSA. 2003. Ecological Assessment of the South Australian Beach-cast Seagrass and Marine Algae Fishery. Assessment Report Prepared for Environment Australia, against the Guidelines for the Ecologically Sustainable Management Of Fisheries. For the purposes of Part 13(A) of the Environment Protection and Biodiversity Conservation Act 1999. November 2003. <http://www.environment.gov.au/coasts/fisheries/sa/seagrass/pubs/seagrass-submission.pdf>

Trulio, L. Callaway, J. and S. Crooks. 2007. *White Paper on Carbon Sequestration and Tidal Salt Marsh Restoration*, Dec. 20, 2007. http://www.southbayrestoration.org/pdf_files/Carbon%20Sequestration%20Dec%2020%202007.pdf viewed 15 March 2010

Preparing for the Invasion – Retreat from the coast

Mick Strack

School of Surveying
University of Otago
310 Castle St, Dunedin
mick.strack@otago.ac.nz

ABSTRACT

The effects of Climate Change, such as increasingly severe storm events and sea level rise, are contributing to an increasing risk of coastal erosion. The demand for and price of coastal land continues to increase in spite of the hazard. Engineering solutions (hard defences) have been well tested in the past and have proved to be, at best, temporary and very expensive measures. Other beach protection measures may include beach profile re-grading and sand replenishment. Planning responses include setting aside public coastal margin strips, rolling easements and creating development set-back margins, to provide a spatial buffer between the sea and vulnerable infrastructure. A policy of managed retreat requires existing housing and infrastructure to be moved away from the shore. The preservation of the public and natural character of the coastal environment and landscape is a matter of national importance. The demands of sustainability require us to enhance the natural functioning of the coastal environment, to maintain natural landscapes, and to provide for public access.

However, coastal property owners have an expectation that they have a right to protect their property, that their local authorities will protect their land, and that consent will be granted for continued coastal development. Local authorities have to be creative and innovative in balancing competing demands on the management of the coastal environment. This paper describes the issues and recommends some examples of good practice.

KEYWORDS: coasts, erosion, sustainability, property rights, planning, managed retreat.

1 INTRODUCTION

While the details and effects of Climate Change are still contested, it is irrefutable that the risks of sea level rise and the probability of increased storm surge events are real. Different coasts will be affected in quite different ways. Some coastal land is rising due to post glacial uplift, and land accretes by reliction (Boateng 2010). In other places the coastal land is very vulnerable to erosion, either because it is particularly low lying or the geology¹ is such that the land is easily washed away. There has been much survey work completed to identify low lying land using LiDAR surveys from which vulnerability maps can be derived (Boateng 2010). This work is often concentrated on urban centres where local authorities require early warnings about vulnerable property and infrastructure. In developed and urbanised nations it is likely that national capacity exists to protect such low-lying urban land – given the long time frame available to build defences against incremental sea level rise. The value of protected property and infrastructure is likely to make the cost-benefit ratio positive. The more immediate issue of adapting to sea-level rise emerges where there is less dense development and where the infrastructure assets are less strategically vital, but where private property is at risk. For example, there are many beach-front houses and local access roads (rather than arterial roads) that are currently at risk from erosion. In these situations there are several conflicting interests to consider: the proprietors who feel justified in protecting their property or demanding that the local authority undertakes protective works; the local authority who may feel that hard protection structures are both environmentally detrimental

¹ For example, soft mudstone cliffs, or poorly stabilised dunes.

and financially prohibitive; and the public who expect the beach to retain its natural character and to provide access and recreation opportunities.

It is at the interface between the sea and the land, and between public and private, with which this paper is concerned. Also, whose interests can be defended in a judicial setting, how local authorities balance the competing demands of their communities, and how legislation and policy intervention can provide a fair outcome for all.

2 COASTAL EROSION

The assurances and protections that attach to a fee simple title in New Zealand suggest that title to land is forever, but this “ignores the reality that the underlying land may itself be impermanent” (Turbott 2006;44). It is apparently a logical response to the threat of coastal erosion to seek to protect property from loss, and huge efforts have been exerted to design and build coastal defences. Previous coastal management paradigms have seen this as a challenge: to defend against the natural forces of erosion; to assert mastery over nature and retain property as a core value. That paradigm is shifting in recent years towards an acceptance of nature and a desire to work with her forces.

Erosion is not an evil ... “Indeed erosion has several natural and societal benefits: it liberates sediment for the coastal system that leads to deposition elsewhere, thus maintaining beaches, barriers and dunes; it is a mechanism by which the coastal topography adjusts to minimise wave energy levels at the coast; it provides materials upon which coastal ecosystems depend and it creates the scenic cliffed coastal landscapes that are so valued by society for their aesthetic appeal as well as their geological interest.” (Cooper & McKenna 2007;296 citations omitted). Similarly, O’Riordan et al (2008;154) suggest “a whole new ecological, economic and societal arrangement would emerge on the new coastline” when new wetlands and conservation areas are created during the process of coastal inundation.

Owners of coastal property need to be aware that the nature of their natural ambulatory boundary requires them to relinquish any claims to land lost by a landward movement of the sea. The question about what action such owners may take to defend their land from any loss to the sea and what governments can do to balance conflicting interests is at issue here.

3 OCCUPYING THE COASTAL MARGINS

3.1 Urban centres

Most of our cities have grown up on harbours. Settlers to New Zealand arrived from the sea - their first point of contact was the land adjacent to harbours. Furthermore, as the country proved to be so suitable for agriculture and the production of primary produce, most of the trade, both throughout New Zealand and overseas, was delivered through these harbours. Most harbours and urban waterfronts are now highly modified with reclamations and sea walls protecting essential infrastructure and strategic assets (oil terminals, storage areas, service industries, and more recently, public social spaces). Ongoing investment in this area and maintenance of walls can relatively easily include incremental vertical additions to keep pace with any foreseeable sea level rises. The decision about whether to defend against sea level rise is fundamentally about costs and benefits rather than about environmental concerns or integrity of coastal processes. “In the case of high density developments (major cities) the value of defended infrastructure easily outweighs the costs of defence.” (Cooper & McKenna 2007;294) but “the costs would escalate well beyond the benefits in many low-lying areas, even where there was some settlement in the form of villages or small towns” (O’Riordan et al :147).

3.2 Coastal lifestyle land

Beachside baches are the archetypal representation of New Zealand summers, with the typical bach being a simply built structure used for essential shelter, while most of the living was done out on the beach or at sea. Greater population demands and more affluent lifestyles today are changing that image. There is ever-growing demand for more beach-front property to be

developed, and new luxury housing is taking over from the simple bach. This more recent style of coastal development² suggests that “we have ‘given up’ on preservation and protection of natural character of much of our coast” (Brookes 2001;8).³ Baches were often able to be uplifted or even dragged and relocated away from the encroaching sea – retreat was a reasonable option. The newer housing, with concrete slab floors, cannot be uplifted; protection or demolition are the only options when the sea encroaches.

It has been recorded that coastal property values are not reflecting the hazard threat (Turbott 2006)⁴ nor the temporary nature of that land. American research (cited in Dahm 2002;24) has found that coastal property owners find “the risk [of coastal erosion] acceptable because they ‘want to be there’, the amenities of an oceanfront location (e.g. view, easy access to the beach, water recreation, peace and quiet) appearing to meet deeply felt emotional needs of the people who owned property there.” Property prices rise on the back of strong demand, and that demand often ignores the threat of loss due to erosion. Coastal land still attracts a price premium over non-coastal land, yet it may be eroded away within 100 years, or require considerable additional investment (if allowed) in building coastal defences. Councils and planners are under pressure to consent to more development of the coastal land and current guidance from the Resource Management Act 1991 (RMA) and rules in the district plans do little to restrict future development.⁵ Furthermore, as development proceeds, demands for protection become more insistent. While there is no case for the legal responsibility for councils to protect private assets, a case for protection can develop from the purposes of the RMA which seeks the sustainable management of natural and physical resources to provide for the social, economic and cultural well being of people and communities. “The principles of maintaining and enhancing natural character are hard to apply to an environment where there are well established physical resources present” (Mason para 81).

For beachfront land, the cost benefit analysis for coastal defences is not so clear cut as it is for urban land. Here, the costs are invariably public while the benefits are private. Natural landscapes and character and public access are strongly held values, but they are not easily quantified in dollar terms. Many local authorities, therefore, have implemented policies to abandon hard coastal defences and implement a policy of managed retreat – allowing land to be lost to the sea and requiring property owners to retreat from the coast.

4 NEW ZEALAND POLICY

The New Zealand Coastal Policy Statement 2010 was prepared “to state policies in order to achieve the purpose of the [Resource Management] Act in relation to the coastal environment of New Zealand.” (DoC 2010;5). It is based on the implementation of sustainability goals, including applying a precautionary approach.⁶ The policy statement supports the preservation of the natural character of the coastal environment including natural processes (Policy 13), and allows for the management of subdivision and use (Policy 25), while enhancing the public utility of the coast as accessible and public open space (Policies 18 & 19).

Coastal hazard risks should be managed by “locating new development away from areas prone to such risks; considering responses, including managed retreat, for existing development in this situation; and protecting or restoring natural defences to coastal hazards” (DoC 2010;10 Objective 5). Coastal hazard risks are to be assessed having regard to “the physical drivers and processes that cause coastal change including sea level rise” and to “the effects of climate change

² Recent development typically involves high value holiday homes dominating coastal landscapes, built to maximize view and proximity to the beach, and there are several examples of the creation of artificial waterways and canal-centred development on totally manmade shorelines.

³ Brookes continues (2001;9), “A cohesive and informed community concerned about the health of their local coastal system becomes replaced (if it ever existed) by individuals concerned with the protection of private property and the effective removal of now hazardous coastal processes from their proximity.”

⁴ Even when hazard warning notices are entered on the title document, values continue to increase (Turbott 2006;22)

⁵ See for example the outcome of an appeal to the Environment Court (*ORC v DCC* [2010] NZEnvC 120) that overturned a council decision to deny consent for a dwelling house on coastal land that was occasionally inundated by the sea, and would likely be more so in the future. The grounds for the decision were that proprietors should be able to take responsibility for their own decisions about building in hazard zones. “There comes a point where a consent authority should not be paternalistic but leave people to be responsible for themselves, provided they do not place the moral hazard of things going wrong on other people” (para2).

⁶ Policy 3: “Adopt a precautionary approach towards proposed activities whose effects on the coastal environment are uncertain, unknown or little understood, but potentially significantly adverse” (DoC 2010;12). It is apparent that hard coastal defense structures must fall into this category of activity.

on ... storm frequency, intensity and surges; and coastal sediment dynamics" (DoC 2010;23 Policy 24(a) & (h)). Local authorities shall "avoid redevelopment, or change in land use, that would increase the risk of adverse effects from coastal hazards ... and encourage managed retreat by relocation or removal of existing structures or their abandonment in extreme circumstances (DoC 2010;24 Policy 25).

The Ministry for the Environment is active in planning for climate change effects on coastal margins. A report to the Ministry (Bell et al. 2001;viii) warns that "coastal development and global warming are on an eventual collision course" and that "managed retreat and adaptation are the only reasonable long-term options." Regional councils have also prepared documentation (Turbott 2008) to assist the planning for coastal erosion and particularly the issues arising from implementing managed retreat – that is, abandoning property or relocating structures threatened by coastal processes. But in terms of putting any effective limits on coastal development, the proposed New Zealand Coastal Policy Statement is far from helpful.⁷

5 LITIGATION

5.1 *Falkner v Gisborne District Council 1995*

Not surprisingly given the significant financial stake they hold in their land, coastal property owners usually expect either to be able to defend their land themselves, or ideally to demand the government defend it for them. The right to protect ones land from the sea was argued in the case *Falkner v Gisborne District Council 1995*.⁸ Since the 1920s, the near foreshore of Wainui Beach has been occupied by housing. The local authority undertook various beach protection works, including sheet pile groynes, rock walls, and post and rail walls, in an attempt to retain and support the beach fore-dunes as a line of defence for the inland properties and houses. Consequent upon occasional severe storm damage to the defensive works, a high level of pressure was exerted on the local authority to continue to enhance these defences to protect private property. By the 1990s the Department of Conservation and the Gisborne District Council reassessed the feasibility of providing engineered solutions to holding back the sea, and proposed to allow the sea naturally to take its toll – to implement a policy of managed retreat.

Adjoining landowners, on the other hand, expected that the protection work was a continuing local government responsibility. The expectations that property should not be lost without just compensation led residents to object to the new policy by lodging a claim to the Planning Tribunal and subsequently an appeal to the High Court.

Evidence was presented about the likely increased threat of coastal erosion due to sea level rise, and also about the desirability of maintaining a fight against the forces of nature. The question of whose responsibility it was to protect property was raised at a philosophical level but also at a practical level regarding ownership of land where coastal protection works were built, and the location of Mean High Water Springs (MHWS).⁹

The judicial argument focused around three main issues, namely the common law duty on the Crown to protect the land from invasion, the common law right of individual owners to protect their own property, and whether the statutory intervention of the Resource Management Act 1991 has abrogated or modified those duties and rights (*Falkner* at p466).

In examining the arguments for the rights and duty to protect property, the court quoted Halsbury (at p471): "It is the royal duty of the Crown to preserve the realm from the inroads of the sea by appropriate defences; and every subject has a corresponding right." However, the court observed that the private right was not absolute: "... no doubt the ordinary rights of property must give way to that which is done for the protection and safety of the public" (at 471).¹⁰

The court, in response to the claim that every landowner has a right to protect his or her land from the inroads of the sea, stated that such an approach "manifests a narrow 19th century

⁷ Peart (2009;237) describes the document as "a damp squib."

⁸ *Falkner v Gisborne District Council* [1995] NZRMA 462

⁹ Under the Resource Management Act 1991 the coastal marine area is defined by Mean High Water Springs (MHWS). The fact that the property boundaries were defined as Mean High Water Mark appeared to be ignored in these arguments. Expert evidence of surveyors was required to define the actual location of the MHWS line on the ground, and it was not a straightforward decision. It is no surprise, however, that the tide regularly reached the foot of the hard defensive structure; as it always will on an eroding beach, and the MHWS line becomes defined by the structure rather than as a naturally determined line.

¹⁰ Quoting Coulson & Forbes.

preoccupation with proprietary rights, out of keeping with the more holistic policy concerns of sustainability and environmentalism popular today” (at p475). The court decided that the policy of managed retreat was a justifiable response to coastal erosion and that this process was not an evil that takes away property but an expected result of natural processes.

5.2 *Mason v Bay of Plenty Regional Council 2007*

This case deals with erosion on Waihi Beach. The Regional Council had consented to a plan of defensive structures to protect beachfront properties,¹¹ with an appellant arguing against the proposed structure and for the retention of the natural beach. Building consent had been granted for beachfront properties for many years, and the Council had maintained hard defence structures to protect those dwellings. The council's proposal was not to abandon the defence works but to provide various 'soft' and 'hard' engineered structures that sought to avoid or mitigate adverse effects on the coastal environment.

The court observed: “Plainly enough, this case gives rise to strongly opposed viewpoints – the concern on the one hand of the relevant beachfront property owners that their properties of considerable value and pleasantness of outlook towards the sea should continue to be protected against coastal hazards erosion; and, on the other hand, the interest of the first appellants and others in looking to enhance the amenity of the beach and providing for all-tide public access along all parts of it” (*Mason v BoPRC*, para 59). The council had balanced the desire to restore a natural beach with the desire to protect built infrastructure¹² and arrived at the solution they did. It was not for the court to impose an alternative solution. In this case the redesigned sea walls were consented to; it was appropriate to fund the protection of property, and public interests and values were adequately allowed for.

6 PLANNING RESPONSES

The engineering solutions that have often been the first response to coastal erosion have been shown to be non-sustainable; practically; economically; environmentally; and socially. Changing the ways we develop near the coast may provide more enduring solutions.

There are, of course, many methods and planning tools available to achieve the appropriate balance between private property rights, public rights, and environmental protection. They may require new policies, new legislation, and a new desire to act proactively rather than just deal with the issue when the dramatic news pictures emerge of buildings collapsing into the sea.

6.1 *Buy and lease back*

The beauty of a fee simple title is that it is essentially perpetual and protected by the state, although it is still dependent on a grant from the Crown. The state rarely uses the compulsory acquisition opportunity implicit in the Doctrine of Tenure – that the state has ultimate ownership, authority and control of all land. However, it is at least possible that the state could acquire coastal land (compulsorily or otherwise) then offer it back to the original owner or to the open market for a short term lease (dependent on predictions of erosion loss). The benefit of this approach is that it would “enable the social and economic future of the settlement to remain intact for as long as possible, yet enable the areas at risk slowly to be removed over a period of time.” (O’Riordan, et al. 2008;154). A full economic analysis of value and the market conditions could indicate where costs may fall.

To illustrate this; consider a coastal property at a desirable beach front location valued at say, \$1m. The cost of financing that property either by a bank mortgage or just by taking the opportunity cost of that money would be about \$60,000 pa. If the State was to purchase that property and then lease it back it could expect a return of \$60,000 pa, \$600,000 for 10 years plus adjustments for the price of money. In other words, in about 15 years the State may have funded the purchase. Then depending on whether the land needed to be abandoned in 10 years or 100

¹¹ The council recognised that “in the past, seawalls were the usual response to coastal erosion” but asserted that they arrived at the best solution (redesigned walls) with a “full cost benefit analysis” (*Mason v BoPRC* para 54).

¹² The RMA’s requirement to promote the sustainable management of natural **and physical** resources (emphasis added).

years, the state would earn a return on its land, the leasehold owner would be well notified that the term is not perpetual and that it was a matter of natural forces when the land must be abandoned and turned over to the sea. In the meantime the lessees would retain full occupation, possession and use rights, and furthermore, retain the full market potential of the leasehold interest – they could buy, sell and transmit their interest and even secure a mortgage over the leasehold value. A variation of this arrangement may include that the buy-back is just for the land, an arrangement which would enable the leaseholder to retain full interest in the dwelling, to the extent that it could be removed and relocated at any time prior to the land lease expiry, or the land being engulfed by the sea.

From the proprietor's point of view, the alternative is that in due course, the land will be lost to the sea without any compensation, and the final investment return will be negative. As owner, the Crown can be sure that the land will be cleared of structures before it succumbs to the encroachment of the sea and can allow for the natural coastal processes. It can enhance public access to the coast, and can manage the ecological systems by planting appropriate coastal vegetation to restore natural character and enhance natural defences.

This scenario is predicated not on any legal requirement on the state to compensate for land lost to the sea but on political expediency, and for the opportunity to be proactive in implementing a preferred approach to managed retreat.¹³ At some point the government needs to draw a line about who is getting assistance and who is not. It is unlikely that managed retreat will gain any headway if left to resistant private land owners, especially those who have considerable economic, professional and political clout. However, this solution would be controversial to implement. Central government would be reluctant to intervene in local matters and, as long as local government provides an alternative to the strict implementation of managed retreat,¹⁴ there would be little pressure on government to become directly involved with coastal investment and management.

6.2 Rolling Easements

Rolling easements are “a collection of arrangements under which human activities are required to yield the right of way to naturally migrating shores” (Titus 1998;1313). They have been used to good effect in many USA coastal regions. Rolling easements are strips of land that are defined by and move with the coastal boundary.¹⁵ All land remains in the hands of the owners of the upland parcel (except as it is encroached upon by the sea) but there is a public right of access over that land. They continue to allow productive use of the land until the sea eventually takes the land, which may be decades away.

The property bears no cost of the easement until the sea is ready to take it, except to the extent of granting public access rights over the land. “Rolling easements do not render property economically useless; they merely warn the owner that someday, environmental conditions may render the property useless, and that if this occurs, the state will not allow the owner to protect her investment at the expense of the public. By the time the sea threatens the property, owners will have had decades and perhaps centuries to factor this expectation into their plans; and into the price they paid for their property” (Titus 1989;1389).

Because the protection of public values is a legitimate planning tool¹⁶ the setting aside of such an easement is not a taking of property rights and there should be no expectation that compensation is due for the loss of any rights. But rolling easements could be purchased at affordable prices because the lost property right would have little value. “Developers who deny that the sea will rise would view the policy as costing them nothing” (Titus 1989;1285) while developers who foresee sea level rise would be planning for retreat anyway. “Rolling easement policies ... foster political consensus by forcing developers to concede that sea level rise is likely before they can argue that the regulation will affect property values” (Titus 1989;1331).

Rolling easements do not require particular lines to be drawn on plans or on the ground. They would not depend on defining a MHWL, but can be measured from time to time as an offset

¹³ It also has strong social justice implications. It is problematic when some property owners are compensated for loss when others are not. A current issue with compensation packages being offered to some owners of earthquake damaged houses and not others.

¹⁴ As illustrated by the *Mason v BoPRC* case

¹⁵ That may be the line of MHWL but may just be the toe of a cliff, bank or dune.

¹⁶ Similar to the ability to take an esplanade reserve upon the subdivision of riparian or littoral land under the RMA.

from the toe of a dune or cliff. They would be identified by a note on the survey plan and the title document as a publicly reserved easement in much the same way as a marginal strip is shown after riparian land is alienated from the Crown (subject to the CLRA 1990).

Special legislation would need to be passed to establish the process, the rights, the decision making authority and the detail, and it would take an active stance from government. Given the high public benefit, this would be a responsible response, and one that is not without precedent.¹⁷ The shore line boundary could move, the owner would get plenty of warning of loss and the need to retreat, the public would retain access along the coast, and the beach's natural character could be retained.

6.3 Coastal reserves

The Resource Management Act 1991 forms the statutory basis for the setting aside of most coastal reserves. The setting aside of a 20m wide esplanade reserve along the mark of MHWS is an expected condition of any subdivision consent of land (subject to some exceptions; for example, it is not required if the lots being created are over 4ha) (ss229-237). Esplanade reserves are vested in the local authority, and are established to protect conservation values and enable public access and recreational use. It is the consent to subdivide that triggers the setting aside of the reserve, so it is unlikely that a well integrated and connected coastal reserve will be established by this means, both because developers may create lots larger than 4ha, or not subdivide at all. Furthermore, the esplanade reserves are fixed where they were created, so may continue to be eroded away and cease to provide for their purposes.

An alternative is for the Council to allow an esplanade strip which has significant benefits for the coastal environment. An esplanade strip is defined by the MHWS as and where it lies and however it moves – it is ambulatory. Ownership stays with the land owner, but a public easement in favour of the territorial authority is created over the strip. This will mainly be created upon the subdivision of coastal land, but may also be created by negotiation with a landowner. There are more opportunities available here to use this as a tool to ensure continued public access and to regulate development on the land.

Nevertheless, the RMA process provides little scope to use these reserves proactively to protect the coastal environment, maintain public access, and regulate construction on land under threat of erosion. If they could be created as a condition of any resource consent (not just a subdivision consent) then they could provide a more satisfactory result.

6.4 Development setback lines

Development setback lines can be created in district plans to control and regulate building with a coastal hazard zone. They can be imposed over bare land, which will affect the subdivision potential of coastal land, or over previously developed land where they can affect building locations. A development setback is used “to manage the location of new dwellings under the Building Act, to ensure the houses are safely located and the need for seawalls is avoided. It also informs property owners of the potential risk posed by coastal erosion” (Dahm & Gibberd 2009;2). All development opportunities (defined by the possibility of gaining a land use consent on coastal land) should be restricted to the extent that a generous setback line, measured from the dune or cliff toe (rather than the title boundary), ensures coastal land is free from any development infrastructure or building. The land may remain as open space, or in pastoral or arable use. Public authorities may then be in a position to encourage the restoration of a more natural coastal ecosystem. In this way the land value will remain relatively low as it does not have the inflating effect of future opportunity cost, and there may be an opportunity for public purchase to protect the public and environmental benefit. Any dwelling within the setback zone can be denied further resource consents for expansion of dwelling or other structures like protective walls. This puts owners on notice that the life of their structure is limited and they may have to retreat or suffer from future loss.

Setback lines should take into account a time frame of at least 50 years and up to 100 years in terms of sea level rise expectations, although most commentators recognise that this is

¹⁷ The taking of marginal strips for example.

somewhat speculative. A precautionary approach should be applied to any predictions of future sea level rise. Setback lines can be imposed over coastal land as a normal resource management rule in the district plan, and these could have wide application in many coastal regions now. They allow for the clear space required to accommodate the shifting coastal boundary without the demands for protective structures, but they do not accommodate for public access along the beaches above the line of MHWS.

7 CONCLUSION

“[I]f society wants to retain its natural shorelines, then governments will have to induce property owners to yield their land to the sea” (Titus 1998;1308). However, from a political and pragmatic point of view, it will be difficult to require property owners to relinquish their rights. As Coutts (1989;314) has observed, “the result of planning is likely to be ‘democratic’ in that it imposes the limits that are acceptable to society at large rather than those that are necessary for wise management of the resources.” While property owners hold considerable political power and continue to assert their rights rather than defer to their responsibilities to the wider public (for coastal access) and to the environment (for the protection of natural character of the coast), planning solutions will be contested. Planning approaches in the nature of setting aside land, easements or reserves have the potential to accommodate an acceptable balance of interests, but New Zealand planning practice is not known for its proactive intervention in property rights. Land can be taken without compensation as a condition of a subdivision consent application (for example, as esplanade reserve), but it is unlikely, under current resource management laws, that more interests or encumbrances on property would be taken or imposed by governments, no matter how much they may be in the public interest or how they maintain or restore aspects of the environment. Solutions that infringe on property rights to that extent would require legislative backing and that may be viewed as too hard.¹⁸ However, the sea continues to advance on the land and property retreat is inevitable. We can give up in the future or plan for it now.

REFERENCES

- Boateng, I. 2010. *Spatial Planning in Coastal Regions: Facing the Impact of Climate Change*. FIG Comm 8 Working Group 8.4. Copenhagen.
- Brookes, H. c.2001. Chances Beckoning: Integration in New Zealand’s Coastal Management. retrieved 20/11/09. www.rmla.org.nz/publications/h_brookes.doc
- Cooper, J.A.G. & J. McKenna. 2008. Social justice in coastal erosion management: The temporal and spatial dimensions. *Geoforum*, 39 (2008) 294-306.
- Coutts, B.J. 1989. Mean High Water as a Cadastral Boundary. *Ocean & Shoreline Management*, 12 (1989) 309-330.
- Corrina Dahm, Eco Nomos. 2002. *Beach User Values and Perceptions of Coastal Erosion*. Environment Waikato. Technical Report 2003/03. 68. Retrieved 20/11/09 <http://www.ew.govt.nz/Publications/Technical-Reports/Beach-user-values-and-perceptions-of-coastal-erosion/>
- Dahm, J. & B. Gibberd. 2009. *Coromandel Beaches – Coastal Hazards: Review of Primary Development Setback at Selected Beaches*. Thames Coromandel District Council. Retrieved 20/11/09. web.tcdc.govt.nz/mvxtprod/cache/0012CBA1.001.pdf
- DEFRA. Department for Environment Food and Rural Affairs. 2009. Coastal erosion assistance package Impact Assessment. Retrieved 20/11/09 <http://www.defra.gov.uk/corporate/consult/coastal-change/>

¹⁸ Note for example, a 2005 government proposal to set aside a 5m wide public access easement over private land along all waterways which was angrily protested by rural landowners and the government backed down.

- Department of Conservation DoC. 2008. *Proposed New Zealand Coastal Policy Statement 2008*. Policy Group, Department of Conservation, Wellington. Retrieved 26/06/10 from: www.doc.govt.nz
- Department of Conservation. DoC. 2010. *New Zealand Coastal Policy Statement 2010*. Department of Conservation, Wellington.
- Gisborne District Council. GDC. 2009. Application by GDC Rivers & Land Drainage Department. 2009. For Coastal Permit and Land Use Consent. Retrieved 20/11/09
<http://www.gisborne.govt.nz/assets/Files/Planning-Policy/Wainui-Beach-works/Stage3-ResourceConsentApplicationAEE.pdf>.
- Kalo, J.J. 2005. North Carolina Oceanfront Property and Public Waters and Beaches: The rights of littoral owners in the Twenty-First Century. *North Carolina Law Review* [2005] 83;1427-1505.
- Milligan, J. & T. O'Riordan. 2007. Governance for Sustainable Coastal Futures. *Coastal Management*, 35:4 499-509 July 2007.
- O'Riordan, T., S.A. Nicholson-Cole, & J. Milligan. 2008. Designing Sustainable Coastal Futures. *21st Century Society*, 3:2, 145-157 June 2008.
- Peart, R. 2009. *Castles in the Sand*. Craig Potton Publishing. Nelson.
- Stallworthy, M. 2006. Sustainability, Coastal erosion and Climate Change: An Environmental Justice Analysis. *Journal of Environmental Law* (2006) 18:3, 357-373
- Titus, J.G. 1998. Rising Seas, Coastal Erosion, and the Takings Clause: How to save wetlands and beaches without hurting property owners. *Maryland Law Review*, 57:4 (1998) 1279-1399.
- Turbott, C. 2006. *Managed Retreat from Coastal Hazards: Options for Implementation*. Environment Waikato. Technical Report 2006/48: 89. Retrieved 20/11/09 from: <http://www.ew.govt.nz/publications/Technical-Reports/Managed-Retreat-from-Coastal-Hazards-Options-for-Implementation/>

Case law

- Mason v Bay of Plenty Regional Council*. Environment Court. A98/07.
- Falkner v Gisborne District Council* [1995] NZRMA 462
- Otago Regional Council v Dunedin City Council* [2010] NZEnvC 120. [2010] NZRMA 263

CORSnet-NSW: Improving Positioning Infrastructure for New South Wales

Volker Janssen, Joel Haasdyk, Simon McElroy and Doug Kinlyside

Land and Property Information, NSW Department of Finance & Services

346 Panorama Avenue, Bathurst NSW 2795, Australia

Volker.Janssen@lpi.nsw.gov.au, Joel.Haasdyk@lpi.nsw.gov.au

Simon.McElroy@lpi.nsw.gov.au, Doug.Kinlyside@lpi.nsw.gov.au

ABSTRACT

CORSnet-NSW is a rapidly expanding network of Global Navigation Satellite System (GNSS) Continuously Operating Reference Stations (CORS) providing fundamental positioning infrastructure for New South Wales that is accurate, reliable and easy to use. It is built, operated and managed by Land and Property Information (LPI), a division of the NSW Department of Finance & Services. Currently (September 2011) consisting of 68 reference stations, the network will grow to more than 120 CORS within the next two years. CORSnet-NSW supports a wide range of GNSS applications in areas such as surveying, agriculture, mining and construction, and provides a solid platform for research and innovation involving satellite positioning technology. It undergoes a continuous program of maintenance to ensure the best possible positioning infrastructure is available to NSW. This paper presents the current status of the CORSnet-NSW rollout and discusses several issues critical to providing (and achieving) reliable GNSS positioning of homogeneous and high accuracy across the State. These issues include the direct connection to the national datum via GDA94(2010) coordinates, the importance of site transformations and absolute GNSS antenna models, considerations regarding GNSS-based height determination using CORS, and the automated monitoring of CORSnet-NSW sites for quality control. Efforts to support regional geodesy in the Asia-Pacific region and future plans are also outlined.

KEYWORDS: CORSnet-NSW, CORS, positioning infrastructure, datum, Network RTK, APCV.

1 INTRODUCTION

Global Navigation Satellite System (GNSS) Continuously Operating Reference Station (CORS) networks are being introduced across Australia and internationally to provide improved access to positioning infrastructure for a wide range of GNSS applications in areas such as surveying, mapping and asset management, precision agriculture, engineering and construction, airborne imaging and sensors, and utilities management. Benefits include datum definition, rationalisation of infrastructure, establishment of multi-user systems, positioning services that are similar across and between networks, consistent and reliable connectivity to the national datum, and the ability to provide a degree of legal traceability for satellite-based positioning. Real Time Kinematic (RTK) GNSS in particular, once initialised, provides high-precision coordinates and allows 'real-world digitising' with the ability to significantly enhance productivity. It has been shown that CORS networks are well-suited to support improving cadastral infrastructure with RTK GNSS techniques (Janssen et al., 2011a).

CORSnet-NSW is a rapidly growing network of GNSS CORS providing fundamental positioning infrastructure for New South Wales that is accurate, reliable and easy to use (Janssen et al., 2010; LPI, 2011a). This network aims to support the spatial community and provide stimulus for innovative spatial applications and research using satellite positioning technology. It is built, owned and operated by Land and Property Information (LPI), a division of the NSW Department of

Finance & Services. CORSnet-NSW undergoes a continuous program of maintenance to ensure principal positioning infrastructure is available across all of NSW, while maintaining national and international standards and best practice (e.g. ICSM, 2006, 2007; Lands, 2006) to accommodate established and developing positioning, navigation and timing applications. Acknowledging that other reference station providers, both private and government, may need to establish, operate and co-exist in the State, CORSnet-NSW forms the backbone of datum realisation for all spatial applications in NSW, ensuring seamless, consistent and accurate positioning across the State. LPI accepts the inclusion of all other suitable reference stations in its network, including in areas already serviced by CORSnet-NSW, to ensure redundancy and continuation of services (LPI, 2011b).

LPI's first CORS was installed in 1992 in Bathurst to support internal survey and aerial photography operations (Kinlyside and Yan, 2005). In 2004, a network of seven CORS was installed in the Sydney metropolitan area and made available to the public one year later under the name SydNET (Roberts et al., 2007). A renewed effort of expansion to extend the coverage of CORS throughout NSW commenced in 2009 and corresponded with the rebranding of the network as CORSnet-NSW (Janssen et al., 2010).

This paper presents the current status of CORSnet-NSW and discusses several issues critical to providing (and achieving) reliable GNSS positioning of homogeneous and high accuracy across the State. These issues include the direct connection to the national datum via GDA94(2010) coordinates, the importance of site transformations and absolute GNSS antenna models, considerations regarding GNSS-based height determination, and the automated monitoring of CORSnet-NSW sites for quality control. Efforts to support regional geodesy and future plans are also outlined.

2 CURRENT NETWORK STATUS AND ROLLOUT

The network currently (September 2011) consists of 68 CORS tracking multiple satellite constellations, mainly located in the highly populated coastal region and the eastern part of the State. Efforts are underway to expand CORSnet-NSW to over 120 stations within the next two years. Figure 1 illustrates the coverage of CORSnet-NSW, showing stations that are operational (indicated by small triangles) as well as some planned stations (indicated by small circles). A 150 km radius around active stations is shown to illustrate sub-metre Differential GPS (DGPS) coverage, while a 50 km radius indicates the maximum coverage area for single-base RTK operation at the 2-cm level (horizontally). Network RTK (NRTK) coverage at the 2-cm level (horizontally) is shown as a pink polygon extending from the Sydney metropolitan area towards the north and south, in areas that have sufficient station density to support this technique.

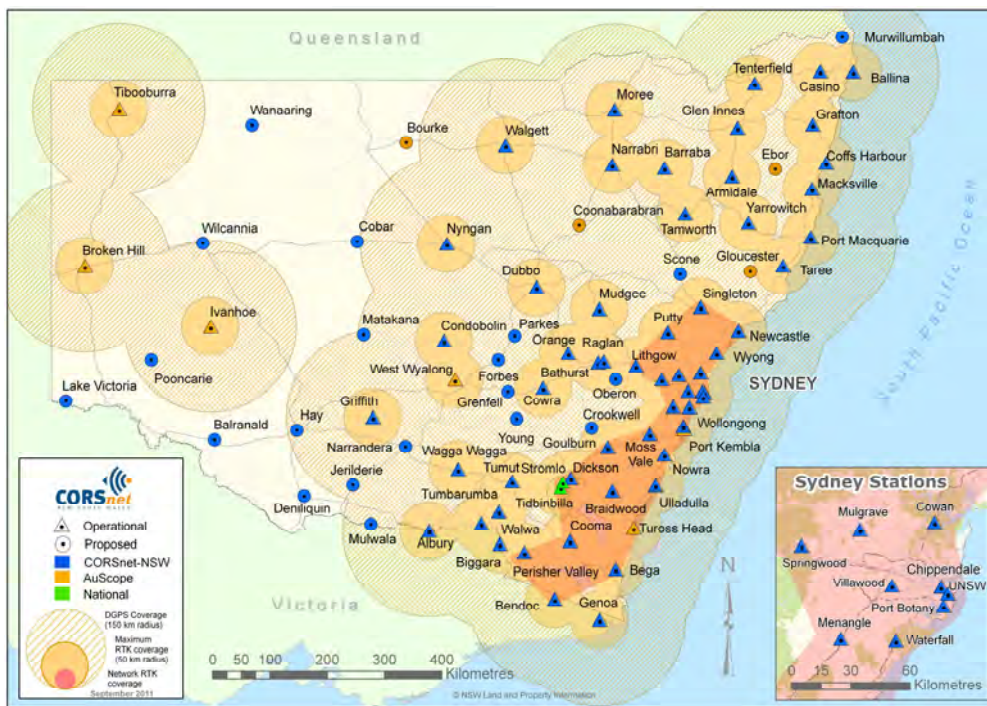


Figure 1: Current coverage of CORSnet-NSW (September 2011).

Currently 83% of the area of NSW is covered by the DGPS service, while single-base RTK is available to 35% of NSW. It should be noted that the State covers a very large area of about 802,000 km², most of which is sparsely populated. As a result, the single-base RTK service is not expected to reach 100% state-wide coverage since dense CORS coverage cannot be justified in all areas. Currently the DGPS service is available to 99.5% of the population, while 94% of the population is covered by single-base RTK.

CORSnet-NSW is operated and managed by a multi-disciplinary specialist team consisting of a technical group (eight staff) and a customer support group (three staff). The network is fully funded by LPI who has committed 7.25 million Australian dollars in capital investment for the Survey Infrastructure Improvement Program of which CORSnet-NSW is a major part. Over the last two years, LPI has invested one million Australian dollars in software, being the second institution in the world to install Trimble Navigation's VRS³Net CORS management software. Several CORSnet-NSW stations have been built to geodetic specifications with support from federal sources, allowing their participation in the scientific, national AuScope CORS network (Janssen, 2009a).

A large number of CORSnet-NSW stations are hosted by local councils, and in the near future several sites will be hosted by private industry. LPI collaborates with the ACT Planning and Land Authority to provide CORS services across the Australian Capital Territory. LPI also collaborates with the Victorian Department of Sustainability and Environment in order to ensure consistent positioning services in the border region between the two states. Currently 80% of CORSnet-NSW stations are hosted by our partners, and this percentage is expected to rise.

All CORSnet-NSW reference stations are equipped with the most recent dual or triple constellation GNSS hardware, purposely mixing GNSS equipment from different manufacturers. In order to provide a legally traceable survey monument that allows the GNSS antenna to be oriented to True North without the need to introduce an antenna height, the new CORSnet-NSW Adjustable Antenna Mount (CAAM) was developed by LPI and a patent submission has been accepted.

Recently, guidelines providing recommendations for the technical design, installation, operation and maintenance of GNSS CORS used in CORSnet-NSW have been published with the objective of ensuring interoperable GNSS CORS across New South Wales (LPI, 2011b). It is anticipated that these guidelines may be adopted by other CORS network operators and may become a national standard.

Station density will be much greater in the east of the State than the west due to application requirements and potential user benefits. As LPI progresses with the rollout of CORS, more users

will have services available to them and the level of service may also improve from its current levels. Due to the lower station density in the State's west, NRTK services will initially not be available state-wide. However, the ongoing modernisation of GPS and the delayed but imminent full operational capability of GLONASS, paired with the deployment of additional GNSS, are expected to support 2-cm level NRTK with larger inter-CORS spacing in the future. Feng and Li (2008) have shown that the use of triple-frequency signals offered by next-generation GNSS will allow maximum distances between CORS to be doubled from currently about 70-90 km to about 140-180 km while maintaining positioning accuracy at the 2-cm level.

Wang et al. (2010) presented a comprehensive study of NRTK user performance, outlining the risks involved in pushing inter-CORS distances beyond 70-90 km for dual-frequency operation. It was also shown that the GNSS rover's Coordinate Quality (CQ) indicators are generally overly optimistic, especially in larger NRTK cells, confirming earlier findings from Edwards et al. (2010) and agreeing well with a recent study carried out in NSW (Janssen et al., 2011b; Janssen and Haasdyk, 2011b).

3 CORSNET-NSW PRODUCTS

Available CORSnet-NSW products include 2-cm level single-base RTK and NRTK services, real-time DGPS, and the provision of RINEX data for post-processing applications. RINEX data for virtual reference stations are expected to be available at the end of 2011. Users can choose from a range of subscriptions available through the major GPS/GNSS equipment suppliers and LPI. Subscription fees contribute to covering the operating and maintenance costs in order to ensure a sustainable and permanent CORS network for NSW well into the future.

Real-time data are provided via Radio Technical Commission for Maritime Services (RTCM) data streams (Heo et al., 2009) at 1-second intervals via the Internet, accessed by users in the field via wireless cellular networks (Yan et al., 2009). NRTK data are provided according to both the Virtual Reference Station (VRS) approach and the Master-Auxiliary Concept (MAC). Compared to single-base RTK, the NRTK solution enables the distance dependent errors (i.e. ionospheric and tropospheric delays and orbit errors) to be modelled more reliably across the network (Janssen, 2009b). It also allows the correction data provided to a user to be optimised based on their (changing) location within the network, thus effectively eliminating the degradation of RTK positioning accuracy with increasing distance from a single base station (Janssen et al., 2011b; Janssen and Haasdyk, 2011b). Should a CORS go down for any reason, an automatic switch is made to utilise an alternative reference station for the NRTK solution, without the need for the user to manually switch to another RTK reference station. NRTK operation thereby improves real-time service availability and reliability.

4 PROVIDING AND ACHIEVING HOMOGENEOUS GNSS POSITIONING ACROSS NSW

Several issues are critical to providing (and achieving) reliable GNSS positioning of homogeneous and high accuracy across the State. This section discusses the direct connection to the national datum via GDA94(2010) coordinates, the importance of site transformations, GNSS-based height transfer considerations, the use of absolute GNSS antenna models, and the automated monitoring of CORSnet-NSW sites for quality control.

4.1 GDA94(2010) vs. GDA94(1997)

The Geocentric Datum of Australia (GDA94) is the basis for geodetic infrastructure in Australia (ICSM, 2006). For a review of coordinate systems, datums and associated transformations in the Australian context the reader is referred to Janssen (2009c). The introduction of CORS or 'active' control marks has revolutionised positioning for spatial professionals. However, these CORS must work in tandem with traditional 'passive' marks in the ground. Issues arise for high-accuracy applications simply because the new control marks are far more accurate than the old. Today we find ourselves working in this challenging transition period

between old and new, which will continue until the introduction of a new national datum sometime in the next decade.

For a reliable NRTK or virtual RINEX solution to be possible, reference station coordinates must have a homogenous accuracy of better than 15 mm (Ramm and Hale, 2004) because multiple CORS are used to model the distance dependent errors across the network. In order to provide CORS users with these new services and to ensure that they are accurate, reliable and easy to use, it was essential to introduce a new set of highly accurate and consistent 3-dimensional coordinates for use in CORSnet-NSW. It was also important to ensure that these new coordinates are consistent with those from other states, particularly Victoria and its state-wide GPSnet CORS network (Hale et al., 2007), so it is possible to share CORS infrastructure and provide seamless positioning across the nation. Victoria completed a state-wide geodetic re-adjustment of its ground control network a few years ago, which provides excellent comparisons with GPSnet-derived positions across the entire state, typically at about a centimetre in metropolitan areas and at the sub-decimetre level in rural areas.

In NSW, a complete re-adjustment of the geodetic ground control network will not take place until a new national datum has been introduced. This is due to the large amount of marks involved (SCIMS currently contains 237,000 survey marks) and the enormous work effort that would be required. Another reason is that users dealing with large datasets (e.g. those handling GIS datasets) should not be forced to transform these multiple times.

Instead, a new ad-hoc realisation of the national datum, GDA94(2010), was introduced by LPI for CORSnet-NSW (Janssen and McElroy, 2010). In order to avoid confusion, the original definition of GDA94 is now referred to as GDA94(1997) in NSW. The concept of different and regular realisations of the same datum is very familiar to geodesists, scientists and surveyors working with International Terrestrial Reference Frame (ITRF) coordinates. This global datum is refined, or realised, every three or so years in order to improve its accuracy, based on increasing amounts of data, and improved modelling and processing techniques (Altamimi et al., 2011).

Legal acceptance of position is an important consideration for some GNSS users and also managers of CORS networks (Hale et al., 2007). Since spatial professionals must work within the constraints of current NSW legislation which requires them to connect to local ground control, all CORSnet-NSW sites are coordinated with both GDA94(1997) and GDA94(2010) coordinates.

GDA94(1997) is the original realisation of the current national datum which was adjusted in 1997. It is sometimes also termed 'local' GDA94 because it is based on local connections between control marks. Many of these coordinates were derived from pre-1980 terrestrial observations. Shortcomings in the initial datum definition (that were not obvious at the time) and the process of propagating coordinates through many layers of measurements and adjustments over the years have caused significant distortions in GDA94(1997), rendering it unsuitable for CORS network operation. Across NSW, known distortions reach up to 0.3 metres in the horizontal component. GDA94(1997) coordinates for each CORSnet-NSW site are determined by LPI through a GNSS-based local tie survey in order to provide connections to the existing ground control network. These coordinates are made available via the Survey Control Information Management System (SCIMS) database (LPI, 2011c).

GDA94(2010) is a later realisation of the national datum. The year in brackets was chosen somewhat arbitrarily to set it apart from the earlier realisation and indicates the date of its introduction. It is sometimes also termed 'global' GDA94, although 'regional' or 'national' would be more correct. GDA94(2010) provides a direct connection to the Australian Fiducial Network (AFN) and its successor, the Australian Regional GPS Network (ARGN), exclusively via GNSS observations. The directness of this GNSS-based connection removes many biases and facilitates virtually distortion-free (a millimetre or two every ten kilometres) spatial control across the network, thereby providing at least an order of magnitude improvement in the positioning framework (Ramm and Hale, 2004). This realisation should also ensure the viability of CORSnet-NSW well into the future, as it is designed to be highly compatible with the planned new national datum.

GDA94(2010) coordinates for all CORSnet-NSW sites are obtained via Regulation 13 certification. These so-called 'Reg 13' certificates are issued by Geoscience Australia, a facility accredited by the National Association of Testing Authorities (NATA). Geoscience Australia determines 3-dimensional site coordinates, which are stated on these certificates, based on one complete week of GNSS data and highly traceable, standardised, scientific processing. Certificates

are valid for five years and provide a Recognised Value Standard for positioning infrastructure with respect to the national datum. Through this facility the site coordinates are linked to a standard of measurement in accordance with the National Measurement Regulations 1999 and the National Measurement Act 1960. Consequently, it assists users in establishing some legal traceability of GNSS positions when CORS data are used. GDA94(2010) coordinates also provide interoperability between existing CORS networks and Geoscience Australia's online GNSS processing service, AUSPOS (GA, 2011). These GDA94(2010) coordinates are available from the CORSnet-NSW website only (LPI, 2011a).

4.2 Importance of Site Transformations

The GDA94(2010) realisation is essential to provide real-time users with reliable, horizontal positioning at the 2-cm or better level. This, of course, means that CORSnet-NSW users obtain positions referenced to GDA94(2010). While this is suitable for applications where users are interested only in absolute accuracy and repeatability (e.g. precision agriculture), spatial professionals are generally required to connect to the existing local survey control network due to legislative requirements or to be compatible with spatial data already referenced to local control. In order to obtain output that is consistent with local ground control marks, it is therefore essential to perform a site transformation (also known as site calibration or localisation) for every real-time survey where existing survey control is located nearby.

The site transformation is performed by observing several established ground control marks surrounding the survey area and calculating a locally valid transformation between the CORSnet-NSW reference frame, GDA94(2010), and the local ground control network, GDA94(1997). This is typically done via a menu tool incorporated in the GNSS rover software. Once the site transformation is performed and found acceptable, it is automatically applied and real-time GNSS positioning then refers to the existing local control network. The use of site transformations is already established good practice to reduce the extent of distortions in GDA94(1997). However, in NSW it is now essential to account for the larger differences in coordinates between the two realisations of GDA94.

In an ideal world, real-time GNSS positioning should be directly compatible with coordinates specified on local survey ground control marks. Therefore a consistent, state-wide geodetic infrastructure based on GDA94(2010) coordinates, or something similar, is the ideal solution. The planned introduction of a new national datum for Australia, based in large part on GNSS observations, is expected to solve this problem. Theoretically, this will remove the need for site transformations.

4.3 GNSS-based Height Transfer Considerations for CORS Users

In regards to vertical coordinates, GNSS-based height transfer is possible by converting ellipsoidal heights (h) determined by GNSS to orthometric heights (H) that refer to the Australian Height Datum (AHD71) (Roelse et al., 1971). This is achieved by applying the geoid undulation (N), also known as geoid-ellipsoid separation, geoid height or N value (e.g. Featherstone and Kuhn, 2006; Janssen, 2009c):

$$H = h - N \quad (1)$$

In practice, geoid undulation information plays two crucial roles (Rizos, 1997): On the one hand, N values are necessary to convert (non-GNSS) geodetic control information (i.e. orthometric heights) into a mathematically equivalent reference system to which GNSS results refer (i.e. ellipsoidal heights). On the other hand, N values are required to obtain orthometric heights (i.e. physical meaning) from GNSS-derived ellipsoidal heights (i.e. geometrical meaning). The growing use of CORS networks for GNSS-based height determination has substantially increased the importance of accurate, absolute N values.

In the traditional base-rover field scenario, the published, local AHD71 height of a temporary GNSS reference station set up on a local ground control mark is converted to an ellipsoidal height using equation (1). The ellipsoidal height of the rover is then determined via RTK or post-

processing techniques and converted back to AHD71 using the same equation. The entire process is based on the *calculated* ellipsoidal height of the reference station. Most of the error in the absolute N values cancels since the conversion is applied from AHD71 to ellipsoidal height and back again (Figure 2). The absolute N values involved may have large errors (e) but by applying the height conversion twice (forward and backward), the AHD71 height of the rover is only contaminated by the small difference in relative N value errors (ignoring any GNSS observational errors).

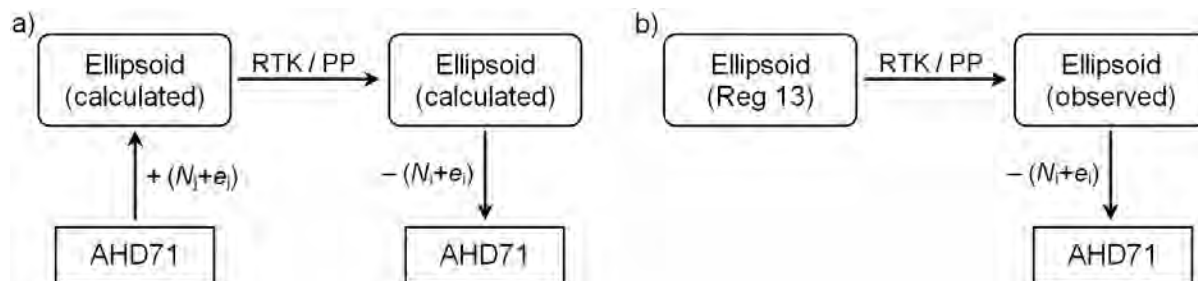


Figure 2: GNSS height transfer methodology using RTK or post processing (PP) in the traditional base-rover scenario (a) and using CORS (b).

In the CORS scenario, the height conversion is only applied once (at the rover end). It is based on an *observed* ellipsoidal height at the CORS, which is generally determined via Regulation 13 certification in Australia (see section 4.1). As the height conversion is only applied once (from ellipsoidal height to AHD71), any error (e) in the absolute N value will fully propagate into the AHD71 height of the rover. Consequently, the absolute accuracy of N values is now more important than ever for AHD71 height determination using GNSS techniques. Fortunately, the recently released AUSGeoid09 geoid model (Brown et al., 2011; Featherstone et al., 2011) provides N values with unprecedented absolute accuracy across NSW (Janssen and Watson, 2010).

4.4 Absolute GNSS Antenna Modelling

Since multiple CORS, often with different antenna types, are used to model the distance dependent errors across the network, appropriate GNSS antenna modelling is crucial for CORS network operators (Janssen and Haasdyk, 2011a). Consequently, it is also an important issue for users of CORS data.

GNSS observations are measured to the antenna phase centre (APC). The APC is not only offset from the actual survey mark but also undergoes variations depending on the azimuth and elevation of the GNSS satellites and the signal frequency (Figure 3). These antenna phase centre variations (APCV) are deviations from the mean APC and can cause additional errors of up to 20 mm in the measurement to a single GNSS satellite. The all-important antenna reference point (ARP), generally located at the bottom of the antenna, is the reference point for measuring instrument heights and for antenna models. Users can confirm its exact location by consulting readily available documentation from the International GNSS Service (IGS, 2011a) to ensure the rover's antenna height to the ARP is correctly measured in the field. The published coordinates of a CORS typically refer to the survey mark which is identical to the ARP if no antenna height is present.

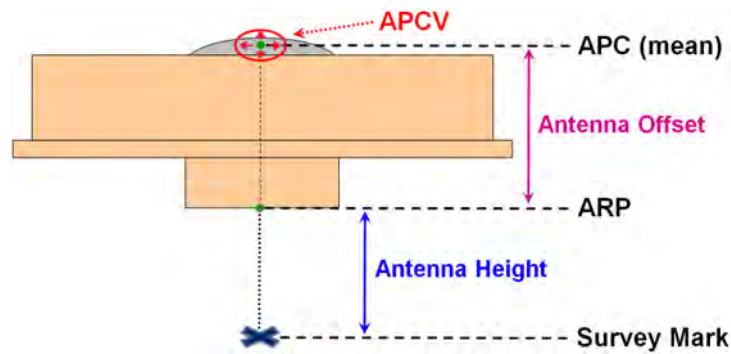


Figure 3: Antenna reference point (ARP), antenna phase centre (APC) and its variation (APCV).

In order to correctly account for the antenna offset as well as any phase centre variations, GNSS antenna types have been calibrated by a number of organisations around the world to generate models. These antenna models provide North, East, Up offsets between the ARP and the mean APC as well as variations dependent on the azimuth and elevation of the received satellite signal, for each frequency. While the North and East offsets for modern GNSS antennas are generally less than a millimetre, they can reach 7 mm for older antennas. The Up offset is very much dependent on the size and design of the antenna and can exceed 200 mm, thus introducing a large error into the height component of the positioning result if not considered.

Not surprisingly, the size, material and design of the antenna have a large effect on the magnitude and distribution of the antenna's phase centre variations. Since GNSS antenna manufacturers have become very good at designing and building symmetric antennas, the azimuth-dependent component is less of a concern than the elevation-dependent component of the variations. However, there are significant differences between antenna types. While the average rover antenna shows a much smaller magnitude of variations than the more expensive CORS choke ring antenna (especially for low elevations), its variations are far less symmetric and show larger differences in the pattern between frequencies (Figure 4).

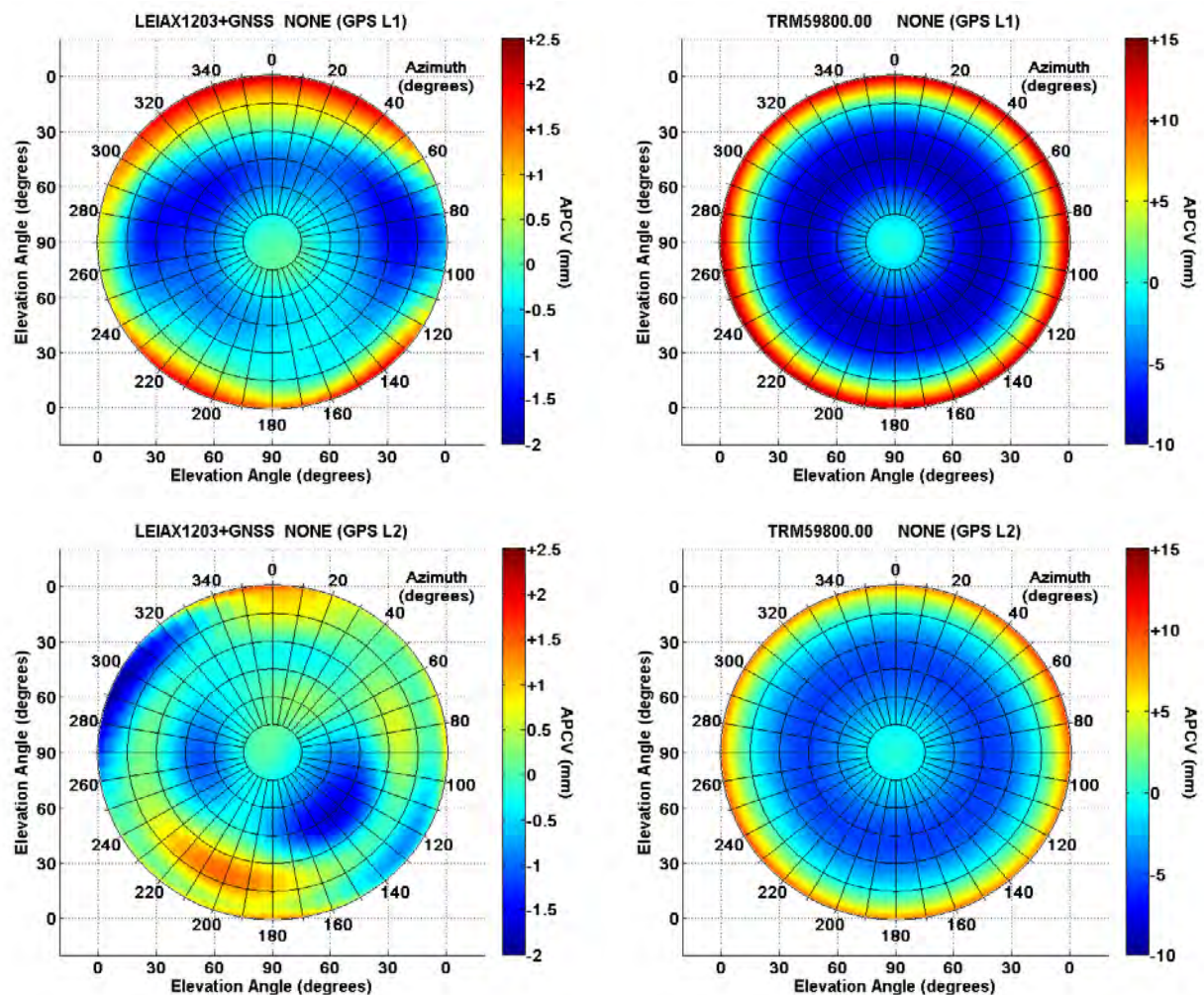


Figure 4: Absolute APCV as a function of satellite azimuth, elevation and frequency for the LEIAX1203+GNSS rover antenna (left) and the TRM59800.00 choke ring antenna (right), both without radome.

In the past, ‘relative’ APCV models were used based on one specific antenna type with assumed zero APCV as a reference antenna (the Dorne Margolin choke ring antenna from Allen Osborne Associates known as AOAD/M_T). However, these have been replaced by ‘absolute’ APCV models because the products of the International GNSS Service (IGS), such as rapid and precise orbits used by CORS operators, are now based on the more rigorous absolute calibrations. While the use of relative APCV models provides correct results if (and only if) no IGS products are used, the combination of relative and absolute APCV models in one project will lead to significant errors, especially in the vertical component.

Absolute GNSS antenna calibrations are performed by several organisations, ideally with a robot rotating and tilting the antenna in an anechoic (i.e. echoless) chamber. Once approved by the IGS, the absolute APCV model parameters, often determined by combining the values obtained from calibrating several antennas of the same type, are made freely available to the spatial community (IGS, 2011b). The published IGS antenna models have recently experienced significant improvement, due to the inclusion of more antenna calibrations and GNSS frequencies, coinciding with the release of ITRF2008 (Altamimi et al., 2011). These models use the ANTEX file format (Rothacher and Schmid, 2010) and follow the international naming convention for GNSS equipment (IGS, 2011c).

While more than one APCV parameter set can exist for a particular antenna type, only one set (generally the best available) is included in the IGS list which is updated regularly to include new antennas. Using the parameters approved by the IGS therefore allows consistency and avoids confusion in regards to which APCV parameter set is the most appropriate. CORSnet-NSW users

are strongly advised to use the absolute antenna models provided by the IGS for both post-processing and real-time operations.

For real-time operations, CORSnet-NSW transmits data specifying all CORS antennas as a 'null antenna', i.e. an antenna with zero antenna offsets and zero APCV. This is achieved by using the absolute APCV corrections obtained from the IGS to reduce the observations to the ARP. Any CORS antenna heights present are automatically considered by the rover through the transmitted RTK/NRTK messages. Therefore, the user does not need to take into account which antenna is used at the CORS site(s) because modelling is taken care of behind the scenes. This considerably simplifies the user's fieldwork because no CORS antenna models have to be uploaded into the rover. The user only has to ensure that the rover equipment applies the appropriate *absolute* IGS APCV model of the rover antenna used in the field.

For post-processing, null antennas are not utilised. Following the RINEX standard, data files from CORS sites or a virtual reference station continue to have observations measured to the APC and will indicate which antenna type has been used. The user should ensure that *absolute* IGS APCV models for both the CORS and the rover are imported and selected in the data processing software. It should be noted that the absolute APCV parameter settings only need to be imported once into the post-processing software and updates are only necessary when a new antenna type is added or in the rare event that the parameters approved by the IGS are updated.

CORSnet-NSW users should apply the following 'golden rules' in regards to GNSS antenna modelling (Janssen and Haasdyk, 2011a):

- Absolute antenna models are GNSS best practice and should be used for real-time and post-processing applications.
- Relative and absolute antenna models should never be mixed.
- IGS absolute antenna models are available from IGS (2011b) or GNSS equipment suppliers.
- CORSnet-NSW uses IGS absolute antenna models, IGS products, and the null antenna principle. Therefore, all CORS 'look the same' from a user perspective. In order to avoid confusion and provide consistency, users are strongly advised to use IGS absolute antenna models.
- Setting the elevation mask at the rover to 10-15° not only reduces atmospheric and multipath errors but also APCV effects.
- All antenna heights should be measured vertically to the ARP, in millimetres and inches, and converted between the two as a check.
- For high-accuracy static surveys, the rover antenna should be oriented to True North.

4.5 Automated Monitoring of CORSnet-NSW

Quality control and integrity monitoring of CORS infrastructure is becoming increasingly important for legal traceability of data and measurements as well as for long-term stability studies of station coordinates. CORSnet-NSW is monitored by determining high-precision daily coordinate solutions using the Bernese 5.0 software (Dach et al., 2007) in an automated process (Haasdyk et al., 2010). Station coordinates are obtained in ITRF and transformed into GDA94. The ongoing analysis of these coordinates can reveal (1) site specific velocities of the network at higher densities than those provided by the global IGS network, allowing comparisons with existing tectonic plate models, (2) medium density sampling of the local distortions present in GDA94(1997) and the distortions in ellipsoidal heights derived from AHD71 in conjunction with AUSGeoid09, and (3) trends in site coordinates which can reveal local ground deformation.

For each CORSnet-NSW site, the resulting coordinate time series showing the difference of the observed station coordinates from the official coordinates is made available on the CORSnet-NSW website (Figure 5). Results show that station coordinates are calculated with millimetre-level precision, while velocities are obtained with 2-4 mm/yr precision and agree with the expected tectonic motion across NSW.

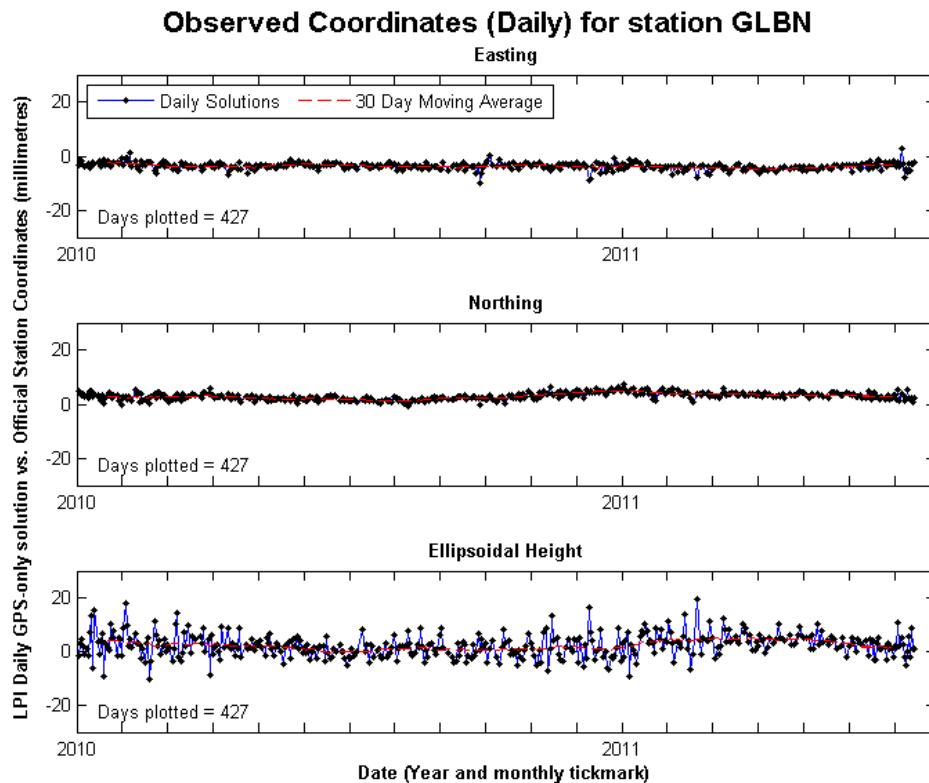


Figure 5: Observed position vs. official position of Goulburn CORS.

5 CONTRIBUTION TO APREF

All modern geodetic datums use reference systems closely aligned with ITRF. The latest realisation of ITRF (ITRF2008) has a precision of a few millimetres (Altamimi et al., 2011), forming a robust basis for any regional or national geodetic datum. The Asia-Pacific Reference Frame (APREF) initiative aims to improve the geodetic infrastructure in the Asia-Pacific region by creating and maintaining a modern regional geodetic framework closely linked to ITRF and based on continuous GNSS data (Dawson and Hu, 2010). Benefits include the development of geodetic datums and CORS networks to support regional development, monitor geophysical hazards and sea level change as well as to synchronise geodetic activities across the region (Stanaway and Roberts, 2010). All CORSnet-NSW sites are contributing data to APREF, thus not only strengthening the regional framework but also allowing the independent monitoring of all CORSnet-NSW sites.

In order to support APREF, all 4-character CORSnet-NSW site IDs have been checked (and changed in a few cases) to avoid conflict with existing national and international CORS sites. In addition, each CORSnet-NSW site has been assigned a 9-digit APREF DOMES number by Geoscience Australia. The DOMES (Directory of MERIT Sites) number was introduced in the early 1980s during the Monitoring of Earth Rotation and Intercomparison of the Techniques (MERIT) campaign, which investigated the relationship between several new space-geodetic techniques, to unambiguously identify each mark involved in this campaign (Wilkins and Mueller, 1986). Nowadays the DOMES number is generally used to provide a unique identifier for each CORS in a particular network.

An IGS site log (IGS, 2011d) has also been generated for each CORSnet-NSW site. It contains detailed station information such as site identification of the GNSS monument (including 4-character ID and DOMES number), site location, and GNSS receiver and antenna information. IGS site logs are international standard for providing up-to-date information on CORS sites, including the history of receiver hardware and firmware updates as well as GNSS antenna changes at the site. This information is frequently updated and the current IGS site log files are available from the CORSnet-NSW website (LPI, 2011a).

6 FUTURE PLANS

CORSnet-NSW will likely grow to over 120 CORS within the next two years, integrating stations from other CORS networks where appropriate in order to help avoid duplication of CORS infrastructure. The planned rollout will focus on substantially expanding availability of the NRTK service from the current coverage area into other regions of the State. It has been shown that NRTK solutions are more robust and achieve accuracies better than those achieved from single-base RTK solutions (e.g. Janssen et al., 2011b; Janssen and Haasdyk, 2011b).

In order to provide the highest possible level of service availability, a second network control centre with full redundancy will be completed in Bathurst in 2011. Such system architecture allows for load balancing and backup between the two control centres in Sydney and Bathurst. Both control centres will use server virtualisation technology to maximise hardware utilisation and at the same time minimise power consumption, space requirements and carbon footprint.

Other forms of redundancy and backup are implemented with uninterruptible power supply (UPS) units being installed at all CORSnet-NSW sites as well as dual communication links at most CORS. The main area of concern continues to be the (generally very reliable) GNSS receivers and antennas themselves since each site is equipped with a single GNSS receiver-antenna pair only and the network density is far less than in other parts of the world. Users continue to be dependent on a reliable connection to the Internet which can be a problem in parts of rural NSW where mobile communications can be limited.

Remote access to the CORS receivers is essential to enable firmware upgrades and troubleshooting in a time-efficient manner. The use of virtual private network (VPN) tunnels will allow direct and secure access to CORS receivers through the firewalls of the host organisations.

In addition to the ongoing monitoring of CORSnet-NSW sites for small, long-term trends (see section 4.5), the commercial Trimble Integrity Manager (TIM) module will be implemented in VRS³Net to enable real-time detection of larger, sudden movements of the CORS antennas. Combining these two state-of-the-art methods ensures that the CORSnet-NSW team is always aware of any motion that the CORS network may be subjected to, whether it is slow tectonic plate motion, land subsidence, or a sudden movement caused by an earthquake or a host removing an antenna from its monument. These tools also ensure continuous validation and quality control of the CORS network and its services, e.g. via the use of continuous RTK/NRTK rover installations and the generation of various quality monitoring statistics.

7 CONCLUDING REMARKS

This paper has presented the current status of CORSnet-NSW, LPI's rapidly expanding CORS network that already provides state-of-the-art GNSS positioning infrastructure across much of NSW. Several issues critical to providing (and achieving) reliable GNSS positioning of homogeneous and high accuracy across the State have been discussed.

These issues include the direct connection to the national datum via GDA94(2010) coordinates, the importance of site transformations and absolute GNSS antenna models, considerations regarding GNSS-based height determination, and the automated monitoring of CORSnet-NSW sites for quality control and integrity monitoring. Efforts to support regional geodesy via the APREF initiative and future developments have also been outlined.

CORSnet-NSW will continue to grow from currently 68 CORS to more than 120 stations within the next two years, providing seamless CORS infrastructure across NSW that is accurate, reliable and easy to use for a wide range of applications. It supports the spatial community across the State, is a valuable resource for major infrastructure projects, and provides stimulus for innovative spatial applications and research using satellite positioning technology.

REFERENCES

Altamimi Z., Collilieux X., Metivier L. (2011). ITRF2008: An improved solution of the International Terrestrial Reference Frame. *Journal of Geodesy*, 85(8), 457-473.

- Brown N.J., Featherstone W.E., Hu G., Johnston G.M. (2011). AUSGeoid09: A more direct and more accurate model for converting ellipsoidal heights to AHD heights. *Journal of Spatial Science*, 56(1), 27-37.
- Dach R., Hugentobler U., Fridez P., Meindl M. (2007). Bernese GPS software version 5.0 user manual. Astronomical Institute, University of Bern, Switzerland, 612pp.
- Dawson J., Hu G. (2010). Improving the geodetic infrastructure of the Asia-Pacific region. Presented at XXIV FIG International Congress 2010, Sydney, Australia, 11-16 April 2010.
- Edwards S.J., Clarke P.J., Penna N.T., Goebell S. (2010). An examination of Network RTK GPS services in Great Britain. *Survey Review*, 42(316), 107-121.
- Featherstone W.E., Kirby J.F., Hirt C., Filmer M.S., Claessens S.J., Brown N.J., Hu G., Johnston G.M. (2011). The AUSGeoid09 model of the Australian Height Datum. *Journal of Geodesy*, 85(3), 133-150.
- Featherstone W.E., Kuhn M. (2006). Height systems and vertical datums: A review in the Australian context. *Journal of Spatial Science*, 51(1), 21-42.
- Feng Y., Li B. (2008). A benefit of multiple carrier GNSS signals: Regional scale network-based RTK with doubled inter-station distances. *Journal of Spatial Science*, 53(1), 135-147.
- GA (2011). AUSPOS – Online GPS processing service, [http://www-b.ga.gov.au/geodesy/sgc/wwwgps/](http://www.b.ga.gov.au/geodesy/sgc/wwwgps/) (accessed Sep 2011).
- Haasdyk J., Roberts C., Janssen V. (2010). Automated monitoring of CORSnet-NSW using the Bernese software. Proceedings of XXIV FIG International Congress 2010, Sydney, Australia, 11-16 April 2010, 19pp.
- Hale M., Ramm P., Wallace J. (2007). Legal positions and Victoria's GPSnet. Proceedings of SSC2007, Hobart, Australia, 14-18 May 2007, 163-176.
- Heo Y., Yan T., Lim S., Rizos C. (2009). International standard GNSS real-time data formats and protocols. Proceedings of IGNS2009, Surfers Paradise, Australia, 1-3 December 2009, 17pp.
- ICSM (2006). Geocentric Datum of Australia technical manual, version 2.3, <http://www.icsm.gov.au/icsm/gda/gdatm/index.html> (accessed Sep 2011).
- ICSM (2007). Standards and practices for control surveys (SP1), version 1.7, <http://www.icsm.gov.au/icsm/publications/sp1/sp1v1-7.pdf> (accessed Sep 2011).
- IGS (2011a). Position of the antenna reference point, <ftp://igs.org/igscb/station/general/antenna.gra> (accessed Sep 2011).
- IGS (2011b). Absolute IGS APCV parameters, <ftp://igscb.jpl.nasa.gov/igscb/station/general/igs08.atx> (accessed Sep 2011).
- IGS (2011c). Naming conventions for IGS equipment descriptions, ftp://igscb.jpl.nasa.gov/igscb/station/general/rcvr_ant.tab (accessed Sep 2011).
- IGS (2011d). Instructions for filling out IGS site logs, ftp://igscb.jpl.nasa.gov/pub/station/general/sitelog_instr.txt (accessed Sep 2011).
- Janssen V. (2009a). Precision rules! How to establish an AusCORS site. *Position*, 44, 64-66.

- Janssen V. (2009b). A comparison of the VRS and MAC principles for Network RTK. Proceedings of IGNSS2009, Surfers Paradise, Australia, 1-3 December 2009, 13pp.
- Janssen V. (2009c). Understanding coordinate systems, datums and transformations in Australia. Proceedings of SSC2009, Adelaide, Australia, 28 Sep - 2 Oct 2009, 697-715.
- Janssen V., Grinter T., Roberts C. (2011a). Can RTK GPS be used to improve cadastral infrastructure?. Engineering Journal, 15(1), 43-54.
- Janssen V., Haasdyk J. (2011a). CORS networks: Absolute antenna models are absolutely vital. Position, 51, 36-40.
- Janssen V., Haasdyk J. (2011b). Assessment of Network RTK performance using CORSnet-NSW. To appear in Proceedings of IGNSS2011, Sydney, Australia, 15-17 November 2011.
- Janssen V., Haasdyk J., McElroy S. (2011b). CORSnet-NSW Network RTK: Same look and feel... only better. Proceedings of Association of Public Authority Surveyors Conference (APAS2011), Bathurst, Australia, 6-7 April 2011, 39-54.
- Janssen V., McElroy S. (2010). Coordinates and CORSnet-NSW: Dealing with distortions in GDA94. Position, 50, 24-27.
- Janssen V., Watson T. (2010). Improved AHD71 height determination from GNSS using AUSGeoid09 in New South Wales, Australia. Journal of Global Positioning Systems, 9(2), 112-121.
- Janssen V., White A., Yan T. (2010). CORSnet-NSW: Towards state-wide CORS infrastructure for New South Wales, Australia. Proceedings of XXIV FIG International Congress 2010, Sydney, Australia, 11-16 April 2010, 14pp.
- Kinlyside D., Yan T. (2005). SydNET – First results. Proceedings of Association of Public Authority Surveyors Conference (APAS2005), Batemans Bay, Australia, 16-17 March 2005, 10pp.
- Lands (2006). Surveying and Spatial Information Regulation 2006, <http://www.legislation.nsw.gov.au/maintop/view/inforce/subordleg+530+2006+cd+0+N> (accessed Sep 2011).
- LPI (2011a). CORSnet-NSW, <http://www.corsnet.com.au/> (accessed Sep 2011).
- LPI (2011b). Guidelines for CORSnet-NSW Continuously Operating Reference Stations (CORS), version 1.0. Available from http://www.lpma.nsw.gov.au/survey_and_maps/corsnet-nsw/reference_material (accessed Sep 2011).
- LPI (2011c). SCIMS online, http://www.lpma.nsw.gov.au/survey_and_maps/scims_online (accessed Sep 2011).
- NGS (2011). GPS antenna calibration, <http://www.ngs.noaa.gov/ANTCAL/> (accessed Sep 2011).
- Ramm P., Hale M. (2004). Realisation of the geodetic datum in Victoria. Proceedings of GNSS2004, Sydney, Australia, 6-8 December 2004, 14pp.
- Rizos C. (1997). Principles and practice of GPS surveying. Monograph 17, School of Geomatic Engineering, The University of New South Wales, Sydney, Australia, 560pp.
- Roberts C., McElroy S., Kinlyside D., Yan T., Jones G., Allison S., Hendro F., Hoffman S. (2007). Centimetres across Sydney: First results from the SydNET CORS network. Proceedings of SSC2007, Hobart, Australia, 14-18 May 2007, 152-161.

- Roelse A., Granger H.W., Graham J.W. (1971). The adjustment of the Australian levelling survey 1970-1971. Technical Report 12, Division of National Mapping, Canberra, Australia, 81pp.
- Rothacher M., Schmid R. (2010). ANTEX: The antenna exchange format, version 1.4, <ftp://igs.org/igscb/station/general/antex14.txt> (accessed Sep 2011).
- Stanaway R., Roberts C. (2010). CORS network and datum harmonisation in the Asia-Pacific region. Proceedings of XXIV FIG International Congress 2010, Sydney, Australia, 11-16 April 2010, 14pp.
- Wang C., Feng Y., Higgins M., Cowie B. (2010). Assessment of commercial Network RTK user positioning performance over long inter-station distances. Journal of Global Positioning Systems, 9(1), 78-89.
- Wilkins G.A., Mueller I.I. (1986). On the rotation of the earth and the terrestrial reference system: Joint summary report of the IAU/IUGG working groups MERIT and COTES. Journal of Geodesy, 60(1), 85-100.
- Yan T., Lim S., Rizos C. (2009). Performance analysis of real-time GNSS data distribution over the internet. Proceedings of SSC2009, Adelaide, Australia, 28 Sep - 2 Oct 2009, 491-502.

How will all the new GNSS signals help RTK surveyors?

Craig Roberts

School of Surveying and Spatial Information Systems
University of New South Wales
Sydney, NSW, 2052
c.roberts@unsw.edu.au

ABSTRACT

In the next few years, surveyors using high productivity, real-time kinematic (RTK) positioning and indeed all other precision positioning users will be faced with a barrage of new global navigation satellite system (GNSS) signals from GPS, GLONASS, GALILEO and possibly COMPASS/Beidou2 as well as Satellite Based Augmentation Systems (SBAS) signals and Regional Navigation Satellite Systems (RNSS). New receivers with enhanced capabilities will enter the market and the old days of choosing which colour GPS receiver will be replaced by a myriad of new considerations.

This paper will attempt to explain the benefits of the new signals that will be on offer in the next 5 or so years and hopefully provide a guide to RTK surveyors and precision positioning users to assist any purchases they intend to make.

KEYWORDS: RTK, GNSS, Surveying, CORS, signals

1 INTRODUCTION

Surveyors are the largest group of professionals dealing with high precision positioning using modern satellite signals. The cm-level demands of surveying has forced the utilisation of the Global Positioning System (GPS) signals beyond their original design by using differential techniques thereby achieving much higher precision and productivity. GLONASS is now a worthy competitor for GPS however many surveyors ignore the politics and simply use combined GPS/GLONASS receivers and benefit from the extra signals. This advance has coined the term Global Navigation Satellite Systems (GNSS). Simultaneously, the Chinese Compass/ Beidou2 system now comprises 9 functioning satellites in its growing constellation (June 2011) and the European Galileo system is set to launch its first operational satellites in October. The Japanese Quasi-Zenith Satellite System (QZSS) has launched its first satellite in November 2010 which heralds the beginning of a GNSS compatible Regional Navigation Satellite System (RNSS) in East Asia. Satellite based augmentation systems (SBAS) such as the US Wide Area Augmentation System (WAAS) (part of GPS), the European Geostationary Navigation Overlay Service (EGNOS) (part of Galileo), the Indian Regional Navigation Satellite System (IRNSS or GAGAN) and the Japanese Multi-functional Transport Satellite (MTSAT) Satellite-based Augmentation System (MSAS) add further complexity to a market already crowded with abbreviations. Add to this a myriad of free Precise Point Positioning (PPP) services and the commercial Wide Area Differential GPS (WADGPS) systems such as Omnistar and Starfire and the modern surveyor can be forgiven for feeling a little confused. This paper will attempt to overview all these services and highlight the benefits for high precision positioning surveyors.

2 SURVEYING WITH GNSS

2.1 GPS Surveying

GPS surveying implies that high quality, dual frequency GPS receivers and antennas are used in a differential mode taking advantage of both pseudorange and carrier phase observables (ie L1 and L2 signals). Initially surveyors performed static GPS whereby two receivers (one called a base and the other a rover) observed the same constellation of satellites simultaneously. Data is returned to the office where it is downloaded and combined in a commercial software package. The software sorts and cleans the data and then employs a double differencing strategy designed essentially to count the number of L1 wavelengths between the satellite and receiver at the initial measurement epoch. This highly complex process is called ambiguity resolution (AR) or initialisation and allows high precision carrier ranging for as long as the signals are logged by the receiver (Rizos, 1997). This process does not occur instantaneously. Early static baseline processing required hours of data to ensure successful AR. The baseline between the base and rover antenna can be routinely measured to better than $10\text{mm} \pm 1\text{ppm}$ over distances up to 20km.

2.2 RTK GPS

Real time kinematic techniques evolved by simply installing a communications link between the base and rover receivers and performing the office processing at the rover in the field. Commercial RTK systems were first released for widespread use in 1993 (Large et al, 2001). At that time many issues such as fast and correct initialisation, reliable communications, usability and range were all in their infancy. As RTK systems evolved many of these issues were addressed, however the limited user range brought about by atmospheric errors between the base and rover receivers remained. Manufacturers developed new techniques to extend the baseline length whilst preserving the precision of positions. Few technical details were released to explain these advances, but it was clear that improvements in communications links as well as GPS signal processing and adaptive combinations of L1 and L2 code and carrier phase measurements using kalman filtering techniques were employed (Large et al, 2001). It allowed initialisation to occur over longer ranges (20-30km) even under mildly changeable ionospheric conditions.

2.3 Continuously Operating Reference Station (CORS) Networks

Concurrently, continuously operating reference station (CORS) networks were being established for primarily scientific applications (Roberts and Stanaway, 2009), to catch an earthquake as it happened, or by governments for datum maintenance such as the Australian Fiducial Network (AFN) expanded to the Australian Regional Geodetic Network (ARGN) (Geoscience Australia, 2011a). Later state governments began to establish their own state-based CORS networks initially to reduce the cost of maintaining their ground mark infrastructure (Roberts, 2011).

This new infrastructure promoted a greater uptake by land surveyors, particularly in rural areas, who began applying GPS to their traditional cadastral and engineering tasks. CORS data is broadcast onto the internet and, providing there is sufficient wireless network coverage, users no longer require their own base station¹. The limitations on base-rover distance brought about by atmospheric errors can also be overcome by using Network RTK techniques (Landau et al, 2002).

However, the original motivation for establishing CORS networks evolved as new applications, new user groups and new business opportunities emerged (Roberts, 2011). Precision agriculture, construction, mining and emergency management all drove high productivity applications and services and shifted the focus for network administrators toward a more commercial outlook. The release of the new AusGeoid09 has also improved GNSS heighting opening the way for potentially more and new applications (Janssen and Watson, 2010). CORS network administrators are now faced with how to continue to service a diverse range of users with a myriad of new signals and GNSS services.

¹ Actually most rural surveyors still use their own base station in areas without mob phone coverage.

3 **SATELLITE NAVIGATION SYSTEMS**

3.1 Global Positioning System

The basic concept of satellite positioning is one-way trilateration. The position of moving satellites must be known at a point in time (ephemeris) and signals sent from the satellites must contain timing information (codes). The receiving device must also have a clock and compare the incoming signals with its own identical codes. Because the wavelength of the codes does not propagate through the Earth's atmosphere, these codes are modulated onto carrier signals. The original intention of GPS positioning was to use the codes for point positioning at an accuracy of around 10m for military users and 100m for civilian users. Engineers developed differential techniques to enable surveyors to use the carrier signal for positioning as well.

The first Block 1 experimental satellites were launched by the US Department of Defence (US DoD) in 1978 broadcasting the L1 carrier, carrying the C/A (civilian access or coarse acquisition) code and P (precise) code, and the L2 carrier (P code only) on frequencies 1575.42 MHz and 1227.60 MHz respectively. The Block II satellites began launching in 1989 and were capable of switching to Selective Availability (S/A) to degrade the accuracy of the civilian signal to provide 100m positioning for civilian users. Anti-Spoofing (A/S) was also switched on for these newer satellites which encrypted the L2 signal with an additional Y code and inhibited tracking of the L2 (since denoted L2(P/Y)) for high precision users (Hofmann-Wellenhof et al, 2008). Commercial GPS manufacturers devised methods of overcoming this A/S encryption by squaring, cross correlation and other more complex techniques. These patented "semi-codeless" techniques increased the cost of equipment especially for dual frequency users. On May 1, 2000 President Clinton ordered that S/A be switched off for all satellites. This was seen as the first step in the GPS modernisation phase. A/S encryption remained.

The Block IIA and Block IIR (replacement) satellites only included minor design changes. The Block IIR-M (replacement – modified) satellites were the first to offer the new L2C signal which is not encrypted with A/S and a more powerful signal. The Block IIF (follow-on) satellites (first launched in May 2010 – currently 2 in orbit) are the first to offer triple frequency; that is L1 (C/A), L2C and the new L5 signal at a frequency of 1176.45MHz. L5 is also an "open frequency" ie is not encrypted (Hofmann-Wellenhof et al, 2008). The code on the L5 signal has a higher "chipping rate" (10.23Mhz) than the L1 C/A and L2C code (1.023Mhz) which means the "least count" of the L5 measurement is 10 times finer².

The US DoD have stated that they will not support the legacy L1 (C/A) / L2 (P/Y) signals after 2020 which means that surveyors will have to upgrade to L2C or L5 capable equipment to guarantee high precision performance (US Federal Register, 2008).

3.2 GLONASS

GLONASS satellites were first launched in 1982, with a full constellation in early 1996. Economic challenges during the transition from the former Soviet Union to the now Russian Federation resulted in a lack of support for the GLONASS system which subsequently declined to 7 satellites in 2002. The Russian government has since strongly supported the revitalization of the GLONASS system with 22 operating GLONASS-M satellites in the current constellation (June 2011).

The GLONASS-M satellites differ significantly from the GPS satellites. Whereas the GPS system uses codes to differentiate between satellites (code division multiple access - CDMA) the Russian system broadcasts two different frequencies for each satellite (frequency division multiple access – FDMA). Consequently surveyors purchasing GPS/GLONASS equipment are also purchasing a more complicated hardware design to overcome this difference. The GLONASS system also operates on a different geodetic datum and time system (see Figure 5) therefore receiver hardware and software must also accommodate these parameters (Hofmann-Wellenhof et al, 2008).

² Surveyors could imagine that the L1 C/A and L2C code is a cm ruler vs the L5C code as a mm ruler.

GLONASS is undergoing a modernisation phase. The first GLONASS-K satellite was launched in Feb 2011. It will transmit a new CDMA signal called L3 (1202.025 MHz) as well as the legacy FDMA signals. Subsequent GLONASS-K satellites will also transmit CDMA signals on the L1 and L2 bands as well as the FDMA legacy signals and will ultimately transition the whole constellation to a CDMA system to be more compatible with other GNSS systems.

GLONASS will also offer an SBAS, called the System for Differential Correction and Monitoring (SDCM), using 24 ground stations in Russian territories and a new geostationary satellite transmitting correction and integrity data on the GPS L1 frequency (Ulrichich, et al 2011).

3.3 Galileo

The European Union (EU) have recognized the benefit of developing their own GNSS system to reduce their reliance on other satellite navigation systems and also to support strategic initiatives for EU member states. Much negotiation has resulted in an interoperability agreement (see section 4.2) with the US GPS system whereby both systems will broadcast L1 and L5-like signals promoting simplicity in receiver design for future GPS/ Galileo devices. Attempts were made to co-fund the system in a European public-private partnership arrangement but this was ultimately unsuccessful and on-going development will be funded from the public sector (Gibbons, 2007).

Galileo has a great opportunity to leap frog the existing GPS and GLONASS systems whose technology is becoming dated and is slow to replenish with an obligation to provide an uninterrupted service. Galileo will provide an Open Service (OS), Commercial Service (CS), Safety-of-Life Service (SoL) and a Public Regulated Service (PRS) (for European security) (ESA, 2011). The Open Service will be of most interest to surveyors and will offer essentially two signals called E1 and E5 (similar to L1 and L5).

The E1 signal, like the GPS L1 C/A code signal will have a chipping rate of 1.023 Mhz. The E5 signal will not only replicate the L5 with a higher chipping rate of 10.23 Mhz, but will also spread the signal over a wider bandwidth (ESA, 2010). The effect of this is that the E5 signal resolution will be as much as 3 times higher than that of the GPS L5 and stronger. Additionally there is less likelihood of multipath with this modern wide-band frequency. A higher chipping rate requires more power usage for the device and will more than likely increase the expense of the receiver (Avila-Rodriguez et al, 2008; Dempster and Rizos, 2009).

The ground segment of the Galileo system supports the European Geostationary Navigation Overlay Service (EGNOS) and provides integrity information for the CS, SoL and PRS.

3.4 COMPASS/ Beidou2

Despite commencing development of their own GNSS after the EU, China currently has 9 satellites in orbit for their COMPASS/ Beidou³ system. Ongoing progress will move away from the Chinese military to the China Satellite Navigation Project Center (CSNPC) who will take charge of the future research, building, and management of the Chinese NSS.

Initially plans are to establish a 12-satellite RNSS with a view to a full GNSS constellation by 2020. The space segment will consist of 30 medium earth orbit (MEO) satellites (similar to all other GNSSs mentioned) and 5 geostationary satellites. They will broadcast on 4 carrier frequency bands: 1195.14–1219.14MHz, 1256.52–1280.52MHz, 1559.05–1563.15MHz and 1587.69–1591.79MHz. It is intended that the Chinese signals will be compatible with GPS, GLONASS and Galileo receivers. Similar to Galileo, two types of services will be offered namely an open service providing positioning accuracy to within 10m (as well as velocity and timing services) and an “authorised service” for security purposes which offers integrity to authorised users (Pace, 2010, Gibbons, 2009, Gibbons, 2011). The Open Service includes a wide area differential positioning service providing 1-metre real time positioning (Li and Dempster, 2010).

³ The original Chinese RNSS Beidou has evolved into the global Compass system which is augmented by the new generation Beidou2 system, hence the double-barrelled name.

3.5 Quasi-Zenith Satellite System (QZSS)

The Quasi-Zenith Satellite System has been designed by the Japanese Space Association (JAXA) as an augmentation to the existing GPS system to provide high elevation satellites over Japan to overcome problems with navigation in urban canyons. The first satellite (denoted GPS PRN 193) was launched in September 2010 and is currently broadcasting L1 and L2C signals in test mode. The L5 and the new L1C GPS signal will also be broadcast in the near future (QZSS, 2010). GPS users can therefore seamlessly track PRN 193 as if it was part of the GPS constellation. There are plans for the QZSS system to comprise 7 satellites in total and become a Regional Navigation Satellite System (RNSS), that is standalone positioning even in the absence of other GNSS signals.

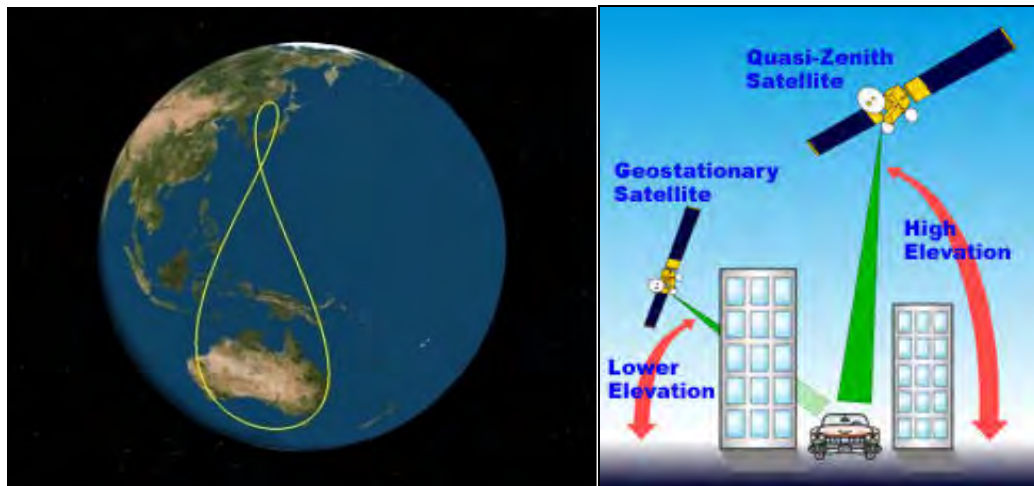


Figure 1: QZSS ground track over Japan and Australia (Kogure, 2007).

The great benefit of QZSS for Australian users is that the ground track passes high over Japan and still remains at a relatively high elevation over Australia. Australia therefore receives the benefit of extra signals by virtue of geography.

3.6 Satellite Based Augmentation Systems (SBAS)

SBASs have evolved due to the modernisation of GPS and the development of Galileo. They comprise essentially a ground-based control segment which provides corrections which are broadcast to geostationary satellites which re-broadcasts these corrections to users. It is a global, pseudorange, DGPS service. In the case of the US Wide Area Augmentation System (WAAS), geostationary satellites not only provide a re-broadcasting communications link over large portions of the USA, but they also broadcast an extra GPS signal with correction messages, based on their monitoring ground based network, modulated onto this broadcast message. Users with a WAAS enabled firmware upgrade of their device can receive essentially DGPS corrected positioning to sub-metre precision within the coverage area (WAAS, 2011). This service is not available in Australia. Similarly EGNOS (EGNOS, 2011) and MSAS (only in limited regions in northern Australia) are also not available to Australian users (MSAS, 2011).

Two commercial services exist namely Omnistar (Omnistar, 2011) and Starfire (Starfire, 2011). The high precision services offered by the commercial operators actually utilise the carrier phase measurement, but require an initialisation time of up to 15 minutes. The repeatability of this initialisation is less robust than using CORS, AUSPOS or precise point positioning techniques (see section 5).

Interestingly, most of these SBAS services will support the L5 signal.

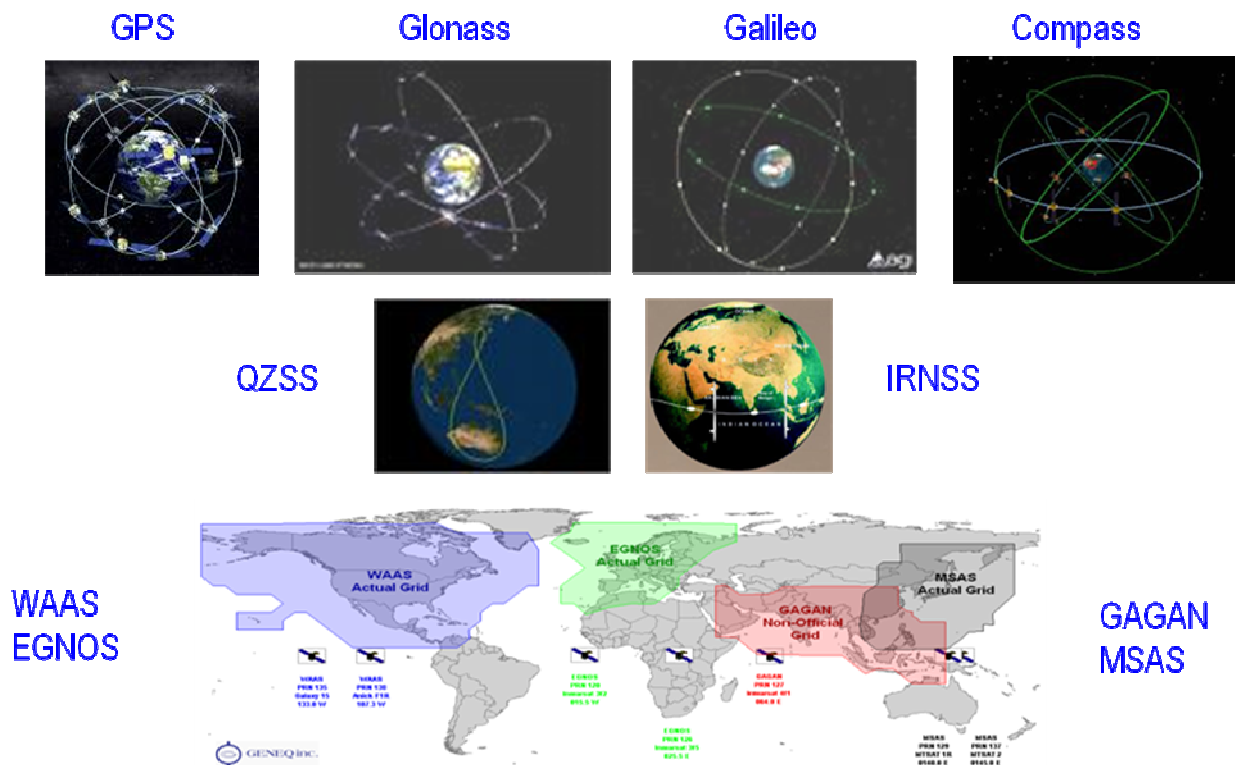


Figure 2: The range of GNSS, RNSS and SBAS globally (Adapted from Gakstatter, 2009).

4 COMBINATIONS OF SIGNALS

Given the number of new and existing satellite positioning services, the days of the RTK surveyor simply buying gear off the shelf knowing that it tracks carrier phase L1 and L2 measurements are gone. As the different GNSS constellations evolve, surveyors will need to have a clearer understanding of the implications of one system versus another and even one signal versus another. However the modern surveyor with an operational grasp of these concepts and good business acumen will benefit from this knowledge.

4.1 Benefits of L2C and L5

The first block IIR-M satellite which broadcast the first L2C signal was available for users in 2005. GPS manufacturers began marketing campaigns to sell new L2C compliant gear, however with just one satellite, the benefits were very limited⁴. Even now in 2011, only 9 L2C capable satellites (7 Block IIR-M and 2 Block IIF) are in orbit.

The block IIF satellites also broadcast the L5 signal. This has significant advantages over the old L2 and even the new L2C signal. The L5 signal is 4 times more powerful than the L2C and has a longer code sequence which means it should exhibit better tracking in difficult environments such as under trees and in urban locations. The signal is also designed to reduce (but not eliminate) the effects of multipath. The original design of the GPS L1/ L2 carrier signals allowed a frequency separation which is exploited to mitigate the effects of the ionosphere over longer baselines. The L1/L5 combination has even greater frequency separation which promotes better ionospheric correction.

⁴ A minor immediate benefit of using just one L2C signal is apparent when RTK surveyors use GPS/GLONASS equipment. The GPS and GLONASS time systems are offset and usually one satellite from a positioning solution is required to solve for the time offset. However the L2C signal accommodates a time difference so that this satellite can also be used in the immediate solution. The GPS/GLONASS time offset changes slowly and is negligible over a 6 hr period.

The L5 frequency is designed for safety of life applications and is therefore located in the highly protected Aeronautical Radio Navigation Services band (Gakstatter, 2011). This guards against future interference and jamming issues currently being experienced by a commercial terrestrial wireless broadband internet provider in the USA (Cameron et al, 2011).

The codes on both the L2C and L5 signals, unlike the existing P/Y code, are not encrypted. This means that complex and patented semi-codeless techniques needed to track the current L2 signal will not be necessary when a full constellation of new signals is available. The cost of GNSS receivers will likely fall dramatically as other manufacturers who do not have (or need) patents to track the new codes will enter the precision positioning marketplace.

Another benefit of the new open codes is that now that the L2C/ L5 signal is easier to track (ie semi-codeless techniques no longer necessary) the signal is likely to be more stable in difficult environments.

However there are only 9 new GPS satellites launched in the past 6 years and only 2 supporting L5. This is the benefit of the new GNSS systems such as Galileo, QZSS, some SBASs and even Compass/Beidou2. These new systems will support the L5 signal thereby accelerating the deployment of the new L5 signal in space.

4.2 Compatibility and Interoperability

The modernisation of GPS and the revitalisation of GLONASS coupled with the likely development of Galileo were the catalyst for the new US Space-Based Position, Navigation and Timing (PNT) policy to define some new terms.

Compatibility refers to the ability of the US and foreign space-based PNT services to be used separately or together without interfering with each individual service or signal, and without adversely affecting navigation warfare.

Interoperability refers to the ability of civil US and foreign space-based PNT services to be used together to provide better capabilities at the user level than would be achieved by relying solely on one service or signal (Hein, 2006).

These two new definitions were the cornerstone of subsequent negotiations between the US and EU when designing future modern Navigation Satellite Systems.

4.3 Benefits of the new NSSs

Galileo will have the capacity to launch multiple satellites in one launch. The first operational Galileo satellites are due to launch in October 2011. It is conceivable that by 2014/15 approximately 12 GPS Block IIF and 18 Galileo satellites will be operational giving a constellation of 30 satellites.

This will give rise to a new interoperable dual frequency RTK combination of L1/ L5 to be exploited by surveyors. Figure 3 illustrates how most of the GNSS/ RNSS providers are accommodating frequencies in the L1 and L5 bands. GLONASS stands out as the only FDMA system and it is now clear why they too will move to CDMA capable satellites. They are also under pressure to not be isolated from all other service providers by continuing to support an L2 signal.

Signal structure and bandwidth are foreign concepts to the RTK surveyor, however a little knowledge is useful. Chipping rate defines the resolution to which a raw measurement can be made. So the L5 (or E5) with a 10.23Mhz chipping rate will be a more precise measurement than a the L1 C/A (or E1) at a 1.023Mhz rate. However the power required to measure the L5/E5 signal is much higher which costs battery life.

Current International Signal Plans

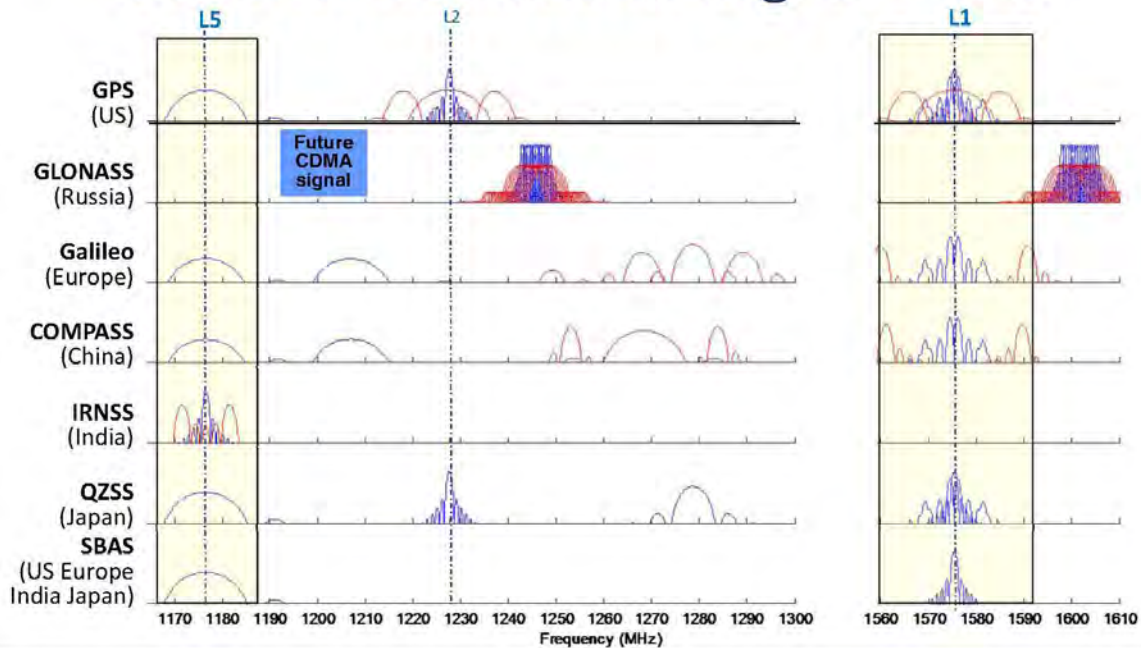


Figure 3: Current International Signal Plans for all planned GNSS and RNSS (Turner, 2010).

The GPS/ Galileo signals are very weak and at the antenna are actually weaker than the background noise. The electronics inside the receiver use correlation techniques to retrieve the code. The GPS L1 C/A and the L5 use a narrow band signal. Galileo uses a wider band signal which improves both the code retrieval and the precision of the measurement on the code by a factor of ~ 3 . The wider bandwidth also means the measurement of the signal is stronger (which is useful in difficult environments) and there is also significantly less likelihood of multipath error. The measurement on the E1 is therefore ~ 3 times more precise than the L1 C/A and the E5 is ~ 3 times more precise than the L5C, however the receiver requires more power to measure this wider band signal and this also presents an electronic constraint on the design of the device (Jinghui Wu, Personal communication, 2011). An improvement in these code measurements assists with carrier phase tracking and more robust initialisation for the RTK surveyor.

Given that Galileo has the advantage of designing a GNSS from current technology (GPS started in 1978), it should be noted that the E1 signal still provides the lower chipping rate to be compatible with the many millions of existing L1 devices. It is conceivable in the future that new devices will operate on the higher resolution L5/E5 band providing higher precision positioning, but again this will require more battery power.

The QZSS signal design is also notable as it provides signals in all bands and is currently a valuable test satellite for future combinations of signals from space. Additionally the ground track of this signal passes over Australia and will be a useful additional signal for surveyors (see Figure 1). Because it aligns with the GPS signal it can be tracked seamlessly with existing GPS equipment.

Another question should be raised about the effectiveness of a GPS-only triple frequency receiver. Currently there are only 2 satellites offering L1 C/A, L2C and L5C signals. The launch schedule for GPS would mean that even by 2020, there may only be 10 – 20 triple frequency capable satellites in orbit with only the QZSS system providing interoperable signals (Figure 3). At the same time, there will be an interoperable constellation of 50+ GPS and Galileo satellites servicing the new L1/L5 linear combination. Will triple frequency offer extra benefit to the modern surveyor or is this a service more suited to a scientist?

In short, triple frequency will allow instantaneous AR over longer distances. This is because traditional dual frequency combinations must either compute a baseline solution whilst ignoring the ionosphere (the next largest error source during double differencing) or vice versa. Triple frequency will allow multiple combinations to compute the baseline and account for the ionosphere thereby providing instantaneous positioning over longer baselines (Rizos, 2008).

The control segment of all the GNSSs is also improving in terms of number and distribution of control stations as well as the resolution of raw measurements, orbit and atmosphere products. This in turn improves the accuracy of positioning for the user.

SBAS accuracy using L1/L5 will be ~10cm as compared to ~60cm today using only L1 (Gakstatter, 2011). What impact will this new capability have on a new geodetic datum for Australia (Johnston and Morgan, 2010; Roberts, 2006; Stanaway and Roberts, 2010)?

One of the great challenges of designing a GNSS receiver for users is marrying the different combinations of signals, datums and time systems to provide positioning services. Figures 4 and 5 illustrate the challenges that receiver designers face.







	System	Develop/ Operation	Satellite Orbit	G/R	Signals		Satellite Launch
					Frequency	MUX	
	GPS	US	MEO	G	L1,L2,L5	CDMA	1978-
	GLONASS	Russia	MEO	G	L1,L2,(L3)	FDMA (,CDMA)	1985-
	Galileo	EU	MEO	G	E1,E5,E6	CDMA	2011-
	Compass	China	MEO+GEO +IGSO	G	B1,B2,B3,L5	CDMA	2007-
	QZSS	Japan	IGSO	R	L1,L2,L5,LEX	CDMA	2010-
	IRNSS	India	GEO+IGSO	R	L5,S	CDMA	2013-?
	SBAS	US, ...	GEO	R	L1,(L5)	CDMA	-

Figure 4: GNSS and RNSS systems compared (Takasu, 2010).

GNSS	Time System	Coordinate System
GPS	GPS Time	WGS84
GLONASS	GLONASS Time	PZ90.01
Galileo	Galileo System Time	GTRF
QZSS	QZSS Time	JGS
Compass	?	?
IRNSS	?	?
SBAS	Own System Time	Own Coordinate System

Figure 5: Comparing the time systems and datums of GNSS services (Takasu, 2010).

The combination of new GNSS and RNSS services presents challenges to CORS network providers who will be under pressure to service their users with suitable combinations of the modern signals. Will they support the new L1/ L5 combination only or provide full triple frequency GNSS services or a strategically distributed mixture of both? RTK surveyors will need to assess what their work requires and ensure their CORS providers support these requirements.

5 IMPACTS FOR RTK SURVEYORS

The US DoD have stated that they will not support the legacy L1 (C/A) / L2 (P/Y) signals after 2020 which means that surveyors will have to upgrade to L2C or L5 capable equipment to guarantee high precision performance (US Federal Register, 2008).

However the new “open” signals will not require patented semi-codeless methods of tracking (such as is currently required for the L2 (P/Y)). This will mean that many more manufacturers will be able to enter the market providing competition to the few manufacturers who currently hold patents. This should drive the price of a high precision GNSS receiver down dramatically. The antenna design will need to be more sophisticated to accommodate the new signals and this may increase but the overall RTK GNSS equipment should be much more affordable.

The L1/L5 combination will supercede the current L1/L2 combination maybe as early as 2015. It is unlikely that RTK surveyors will use triple frequency due to lack of triple frequency equipped satellites (and possibly price) and indeed the requirement of such a long distance solution.

The new signals are more powerful than the existing signals which means more robust tracking in difficult environments such as around buildings and trees which should be an advantage for urban RTK surveyors. The new Galileo signals, by virtue of their wider bandwidth, will also be more resistant to multipath.

The higher chipping rate for the L5/ E5 signal will improve the pseudorange positioning which will increase the time to initialisation for RTK surveyors marginally. This is a minor benefit given that initialisation times are already typically 25 seconds.

More signals however will mean a stronger solution as a result of improved dilution of precision and less likelihood of dropouts due to a lack of satellites. Australia is geographically well situated to benefit from the increase in new signals (ironically more than the US – see Figure 6). Perhaps mission planning will become redundant.

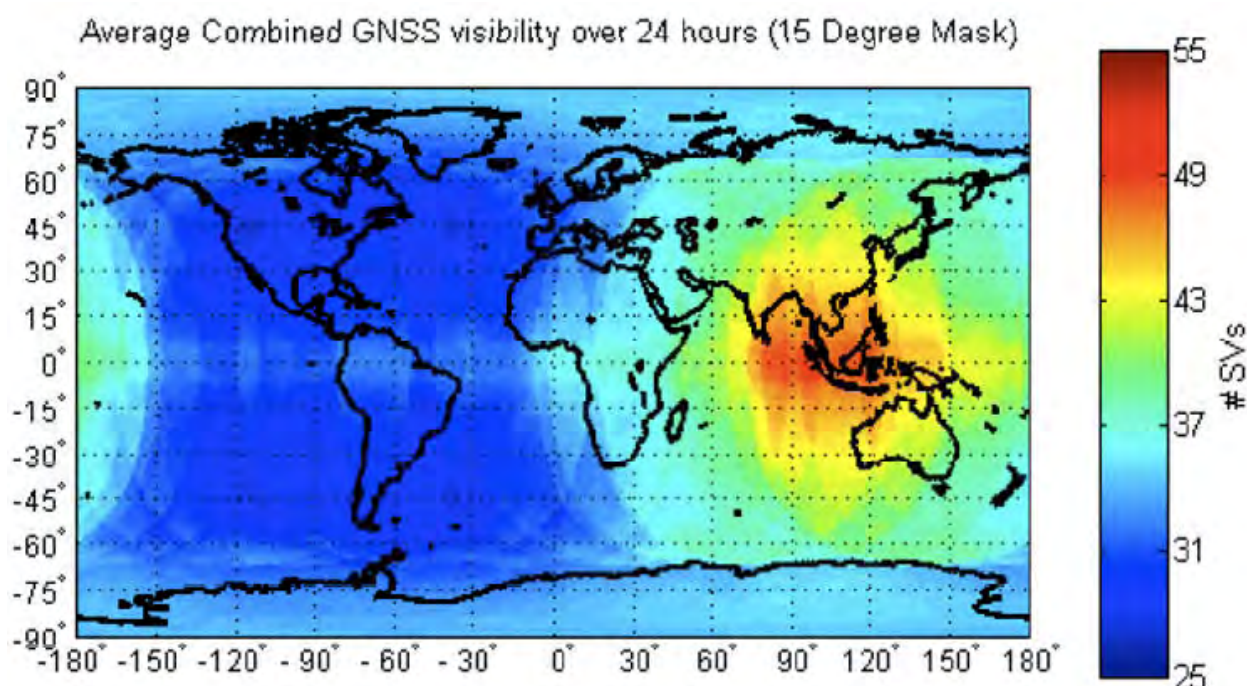


Figure 6: Constellations GPS, Galileo, Glonass, Compass, QZSS, WAAS, EGNOS, MSAS, GAGAN, IRNSS (Dempster and Rizos, 2009).

RTK surveyors working in remote locations can also produce MGA positions to 2cm accurately to support their work more efficiently as a consequence of improvements in the AUSPOS positioning service (Geoscience Australia, 2011b). This is a differential GPS processing service which is offered to users without cost to support high precision positioning. Recent densification of the contributing base stations as part of a large federal government grant (NCRIS project) have decreased the time needed to guarantee 2cm accurate solutions from 6hrs to just 1-2 hours depending on conditions.

Precise point positioning (PPP) services are a worthy competitor for AUSPOS for static positions in remote locations. There are a number of global services provided (IGS, 2011), but the AUSPOS service provides MGA coordinates directly which are of most benefit to RTK surveyors in Australia.

6 CONCLUSION

This paper has attempted to outline the many new GNSS systems and signals which will be available to RTK surveyors in the Australian region in the next few years. These new signals will require both users and CORS network service providers to upgrade their existing GNSS equipment. Hopefully this paper will assist purchasers to more carefully consider which combination of capabilities their new device will require.

This brave new world of high precision positioning also challenges the RTK surveyor to look further at what other business opportunities their expertise may offer. New L1/L5 capable devices will likely be cheaper (as low as \$1000(?) Gakstatter, (2010)) and provide more robust (and slightly more accurate) positioning. Is this a threat or opportunity?

7 ACKNOWLEDGEMENTS

The author would like to thank Dr Jinghui Wu and Professor Andrew Dempster for useful discussion during the writing of this paper.

REFERENCES

Avila-Rodriguez, J., Hein, G., Wallner, S., Issler, J., Ries, L., Lestarquit, L., de Latour, A., Godet, J., Bastide, F., Pratt, T., and Owen, J. (2008) "The MBOC Modulation: The Final Touch to the Galileo Frequency and Signal Plan", NAVIGATION, Vol. 55, No. 1, Spring 2008, pp. 14-28.

Cameron, A., Brown, A., Datta, R. and Gakstatter, E. (2011) LightSquared and GPS: The Sequel, webinar 26 May, <http://www.gpsworld.com/wireless/market-insights-webinars-8423> [Last accessed, 1 Aug 2011]

Dempster, A.G., and Rizos, C. (2009) Implications of a "system of systems" receiver. *Surveying & Spatial Sciences Institute Biennial International Conference*, Adelaide, Australia, 28 September – 2 October, CD-ROM procs.

EGNOS (2011) European Geostationary Navigation Overlay Service, <http://www.esa.int/esaNA/egnoss.html> [Last accessed, 1 Aug 2011]

ESA (2010) Galileo Open Service Signal in Space Interface Control Document, http://ec.europa.eu/enterprise/policies/satnav/galileo/files/galileo-os-sis-icd-issue1-revision1_en.pdf [Last accessed, 29 July 2011]

ESA (2011) European Space Agency, www.esa.int/esaNA/SEM86CSMD6E_galileo_0.html [Last accessed, 1 Aug 2011]

Gakstatter, E. (2009) Using WAAS for sub-metre mapping, 49th Meeting of the Civil GPS Service Interface Committee, Institute Of Navigation GNSS 2010 Conference, Georgia, USA, 21 September.

Gakstatter, E. (2010) What's going to happen when high precision GPS is cheap?, 50th Meeting of the Civil GPS Service Interface Committee, Institute Of Navigation GNSS 2010 Conference, Portland, USA, 20-21 September 2010.

Gakstatter, E. (2011) A closer look at L5: The future of high precision GNSS, webinar 17 March, <http://www.gpsworld.com/wireless/market-insights-webinars-8423> , [Last accessed, 1 Aug 2011]

Geoscience Australia (2011a) <http://www.ga.gov.au/earth-monitoring/geodesy.html> [Last accessed 29 July 2011]

Geoscience Australia (2011b) AUSPOS V2.0 <http://www.ga.gov.au/geodesy/sgc/wwwgps/> [Last accessed, 4 Aug 2011]

Gibbons, G. (2007) Galileo's new PPP: Public-Public Partnership?, Inside GNSS, June/July <http://www.insidegnss.com/node/255> [Last accessed, 1 August 2011]

Gibbons, G. (2009) About Compass, <http://www.insidegnss.com/aboutcompass> [Last Accessed, 3 Aug 2011]

Gibbons, G. (2011) China Plans Another Eight Compass/ BeiDou Launches by End of 2012, <http://www.insidegnss.com/node/2609> [Last Accessed, 3 Aug 2011]

Hein, G. (2006) GNSS Interoperability: Achieving a Global System of Systems or "Does Everything Have to be the Same?", Inside GNSS, Jan/Feb.

Hofmann-Wellenhof, B., Lichtenegger, H. and Wasle, E. (2008) Global Navigation Satellite Systems: GPS, GLONASS, Galileo and more, Springer Verlag.

IGS (2011) Online positioning services of IGS agencies, <http://igsceb.jpl.nasa.gov/overview/links.html> [Last accessed, 4 Aug 2011]

Janssen V. and Watson T. (2011) Performance of AUSGeoid09 in NSW, Proc. APAS 2011, 6-7 April, Bathurst, Australia, 27-36.

Johnston, G. and Morgan, L. (2010) The Status of the National Geospatial Reference System and its Contribution to Global Geodetic Initiatives, FIG Congress 2010, Facing the Challenges – Building the Capacity, Sydney, Australia, 11-16 April

Kogure, S. (2007) QZSS/MSAS Status, 47th Meeting of the Civil GPS Service Interface Committee, Institute of Navigation GNSS 2007 Conference, Fort Worth, USA, 25 Sept.

Landau, H., Vollath, U. and Chen, X. (2002) Virtual Reference Station Systems, Journal of Global Positioning Systems, Vol.1, No. 2: pp 137-143.

Large, P., Goddard, D. and Landau, H. (2001) eRTK: A New Generation of Solutions for Centimetre-Accurate Wide-Area Real-Time Positioning, Trimble Navigation White Paper.

Li, B., and Dempster, A.G. (2010) China's Regional Navigation Satellite System – CAPS. *Inside GNSS*, 5(4), 59-63.

MSAS (2011) Multi-Satellite Augmentation System, http://en.wikipedia.org/wiki/Multi-functional_Satellite_Augmentation_System [Last accessed, 1 Aug 2011]

Omnistar (2011) <http://www.omnistar.com.au/> [Last accessed, 1 Aug 2011]

Pace, S. (2010) Expert Advice: The Strategic Significance of Compass, GPS World, Dec.

QZSS (2010) QZSS Interface specification, <http://qz-vision.jaxa.jp/USE/en/index> [Last accessed, 31 July 2011]

Rizos, C. (1997) Principles and Practice of GPS Surveying, Monograph 17, School of Surveying and Spatial Information Systems, The University of New South Wales, Sydney, Australia, 560pp.

Rizos, C. (2008) Multi-constellation GNSS/RNSS from the perspective of high accuracy users in Australia. *Journal of Spatial Science*, 53(2), 29-63.

Roberts, C. (2006) GNSS and the convergence of Geodesy and the Cadastre in Australia, *FIG 2006, TS33, paper 273*, Munich, Germany 9 – 13 Oct.

Roberts, C. (2011) CORS: Who wants it?, *Position*, No. 51, February – March.

Roberts, C. and Stanaway, R. (2009) CORS Growth and Applications in Australia and Indonesia, South East Asian Survey Congress, Bali, 4 - 7 August.

Stanaway, R. and Roberts, C. (2010) Transformation Parameters from ITRF2008 to Static Geocentric Datums in the Asia-Pacific Region (3773), *FIG2010, Facing the Challenges- Building the capacity*, Sydney, Australia, 11 – 16 April.

Starfire (2011) <http://www.navcomtech.com/StarFire/> [Last accessed, 1 August 2011]

Takasu, T. (2010) Development of an Open Source Multi-GNSS Data Processing Software, 2nd Asia Oceania regional workshop on GNSS, Melbourne, Nov 21-22.

Turner, D. (2010) Update on the US GNSS International Cooperation Activities, 50th Meeting of the Civil GPS Service Interface Committee, Institute of Navigation GNSS 2010 Conference, Portland, USA, 20-21 Sept.

Urlichich, Y., Subbotin, V., Stupak, G., Dvorkin, V., Povalyaev, A. and Karutin, S. (2011) GLONASS: Developing Strategies for the Future, GPS World, April, pp 42- 49.

US Federal Register (2008) <http://www.federalregister.gov/articles/2008/09/23/E8-22197/preservation-of-continuity-for-semi-codeless-gps-applications> [Last accessed, 28 Jun 2011]

WAAS (2011) Wide Area Augmentation System, http://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/waas/ [Last accessed, 1 Aug 2011]

Interpolation of the GNSS Wet Troposphere Delay

Johnny Lo

School of Engineering, Edith Cowan University
Perth, WA, Australia
J.Lo@ecu.edu.au

Ahmed El-Mowafy

Department of Spatial Sciences, Curtin University
Perth, WA, Australia
A.El-mowafy@curtin.edu.au

ABSTRACT

Troposphere delay is one of the main distance-dependent errors in Global Navigation Satellite Systems (GNSS) observations. Precise estimation of the troposphere wet delay is necessary to aid ambiguity resolution and for positioning in network Real-Time Kinematic (RTK) and Precise Point Positioning. Wet tropospheric estimates can also serve as a source of atmospheric information to facilitate weather forecasting. Interpolation of the troposphere wet delay is thus required when its estimation is interrupted for short periods or when data are processed at higher intervals from that of available data. The objective of this research is to compare the performance of several interpolation methods that can be used in order to suggest the most appropriate technique. Six interpolation models were considered. The models ranged from the easy-to-implement linear model, to the more sophisticated Kriging model. Other models considered are the cubic spline interpolation, cubic Hermite polynomial interpolation, Lagrange polynomial interpolation, and Fast Fourier transform interpolation. The performance of these methods was assessed by comparing their results with actual troposphere wet delay data collected at the station Onsala (ONSA) in Sweden. As the number of observations used to generate the interpolation process affects the determination of the model coefficients; the use of different lengths of observations was investigated. The number of missing wet delay values considered for interpolation during testing ranged from one to four in a row.

Test results showed that the linear interpolation, the cubic Hermite polynomial and fast Fourier transform models produce better estimates than splines and ordinary Kriging. The Lagrange polynomials method was the poorest performer. The paper provides explanation of the interpolation results achieved by linking them with autocorrelation of the troposphere wet delays.

KEYWORDS: 1. GNSS 2. Precise Positioning 2. Troposphere wet Delay 4. Interpolation

1 INTRODUCTION

Signal delays induced by the troposphere are generally known as tropospheric refractions or tropospheric delays. A tropospheric delay can be divided into hydrostatic (dry) and wet components. In GNSS data processing at a specific site, instead of dealing with multiple tropospheric delays along line-of-sights between the receiver and each satellite, these delays are usually mapped to a single value along the zenith direction using a mapping function. The zenith hydrostatic delay (ZHD) can be estimated with external models to within a millimetre in accuracy (e.g., Saastamoinen, 1972), and be subtracted from the estimated tropospheric delay, leaving behind the zenith wet delay (ZWD) component, which is mostly due to the atmospheric water vapour. The ZWD can then be used to determine the precipitable water vapour (PWV) for a given

site. A receiver at a nearby location can also make use of this ZWD estimates for accurate positioning.

Although the wet delay accounts for only 10% of the total delay, it is far more difficult to model or remove due to the lack of knowledge regarding the distribution of the water vapour in the atmosphere. The temporal and spatial variability of the water vapour ensures that the wet delay cannot be consistently modelled with millimetre precision by any existing tropospheric model. The Global Navigation Satellite System (GNSS) and water vapour radiometers (WVRs) are two of the most effective tools in estimating atmospheric ZWD. Both techniques are able to estimate the ZWD to within 10 mm or less than 2 mm in PWV (e.g., Haeferle et al., 2004; Liu et al., 2005; Mattioli et al., 2005; Wang et al., 2007). Once estimated, the ZWD can be applied in other GNSS-related areas such as near real-time or real-time kinematic (RTK) GNSS applications. If a network of reference stations (or a single reference station) is able to provide accurate and precise ZWD estimates in a timely manner, these estimates can then be used by a mobile or static user at an unknown location to improve ambiguity resolution, and ultimately, the position solutions.

The benefits of good tropospheric solutions can also extend to aiding Numerical Weather Prediction (NWP) models to provide better weather forecasts. The impact of GNSS PWV estimates on weather forecasting is well documented (e.g., Kuo et al., 1996; Vedel and Huang, 2003; Gutman et al., 2004; Vedel and Huang, 2004; Smith et al., 2006; Macpherson et al., 2007). These studies reported improvements in the humidity and precipitation forecasts when GNSS PWV estimates are assimilated into NWP models. Comparisons between the estimates from a NWP model with and without GNSS PWV estimates assimilation were made and the improvement in relative humidity (RH) forecasts lead to a 40% reduction of forecast errors (Gutman and Benjamin, 2001). The impact of GNSS PWV estimates was further emphasised by a multi-year experiment over the period from 1999-2004 by Smith et al. (2006), whereby improvements were evident in the 6-h and 12-h RH forecasts.

A well-defined statistical description for the GNSS-derived tropospheric estimates is important for NWP modelling. The autocorrelations describe the temporal correlations between pairs of GNSS tropospheric estimates in a time series (TS), as a function of time differences (Borre and Tiberius, 2000). These correlations need to be defined for the eventual assimilation of the GNSS tropospheric estimates into NWP model, especially for the weighting of past data in a bias reduction scheme (e.g., Stoeckli et al., 2007). Furthermore, the autocorrelation time length can be used in recursive data processing procedures such as GM Kalman filtering (KF) with state vector augmentation (e.g., Borre and Tiberius, 2000). Studying the autocorrelation of ZWD can also help in selection of the proper model for its interpolation at a specific instance between known values in the time domain.

2 AUTOCORRELATION OF ZENITH WET DELAY

To better understand the temporal correlations that exist among the tropospheric delay estimates, autocorrelation analysis of the ZWD values estimated from WVR is performed over station ONSA in Sweden, which is also an International GNSS Service (IGS) station. The WVR at ONSA was appropriate for this study as it provided ZWD data at a high frequency (every 8 seconds). In this autocorrelation study, WVR ZWD were sampled every hour and at every 10-min interval. These data were analysed with a 12-h time interval and over three different days on September 10th, 13th and 16th in 2003. A 12-h window ensures a first-order stationarity in the Time Series (TS) of data. Stationary TS refers to a process whose parameters, such as the mean and variance, remain fairly constant over time and space (Wei, 2006). Hence, the corresponding autocorrelation of a stationary TS value can then be deemed constant in any time interval within the 12-h window. The autocorrelation plots are given by Figures 1-3. In these figures, autocorrelation values that lie between the red dotted lines, which represent a 95% confidence interval, are deemed insignificant. Each unit of lag for the left-sided figures represents a 1 hour period and for the right-sided figures the unit lag is a 10-min period. The summary statistics for both the 1-h ZWD and 10-min ZWD data sets are given in Table 1. The results of the autocorrelation analysis between the ZWDs, sampled at different rates, were then compared and summarised in Table 2.

Table 1 Mean and standard deviation (cm) of the WVR ZWD sampled at different time intervals

Sampling Rate	Mean and Standard Deviation (cm)		
	Sep-10	Sep-13	Sep-16
10-min	13.5 (1.3)	10.4 (1.0)	13.0 (2.8)
1-h	13.6 (1.5)	10.4 (1.0)	13.1 (3.1)

Table 2 Comparison between the time lengths for significant autocorrelation of the WVR ZWD sampled at different time intervals

Sampling Rate	Time length with Significant Autocorrelation		
	Sep-10	Sep-13	Sep-16
10-min	1-h 30-min	1-h 50-min	2-h
1-h	1-h	1-h	2-h

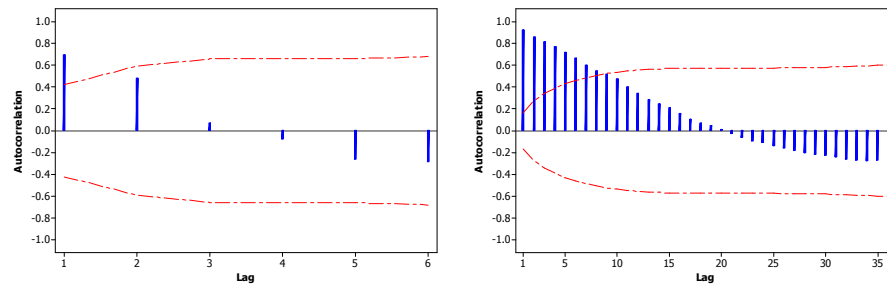


Figure 1: Autocorrelation plot of WVR ZWDs, sampled hourly (left) and at every 10-min interval (right), over ONSA on Sept 10

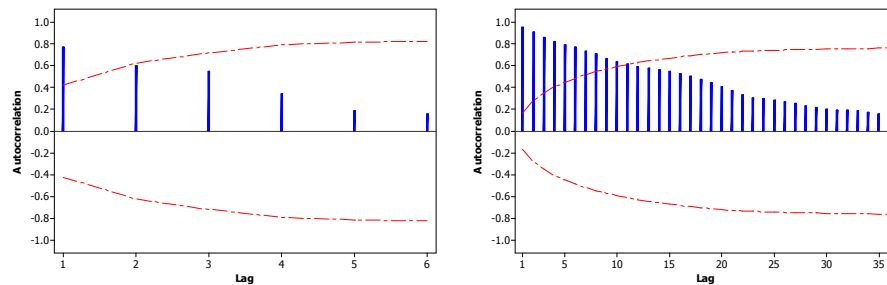


Figure 2: Autocorrelation plot of WVR ZWDs, sampled hourly (left) and at every 10-min interval (right), over ONSA on Sept 13

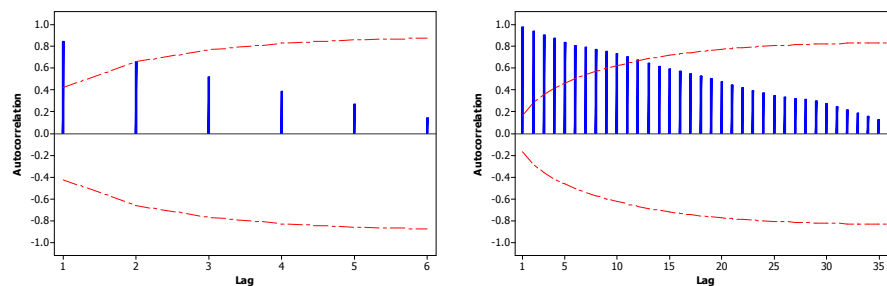


Figure 3: Autocorrelation plot of WVR ZWDs, sampled hourly (left) and at every 10-min interval (right), over ONSA on Sept 16

Although the analysis of 10-min WVR ZWDs involved a greater number of observations than the hourly ZWDs (72 total observations as compared to 12), there are minimal differences between the means and standard deviations of the two sets of data (Table 1). Comparison between Figures 1-3 also shows that the shape of the autocorrelation plots are maintained even when greater number of observations is sampled within the same period. Similarities between the time lengths for significant correlation are also observed for both sets of data in Table 2. On Sept 10th and 13th, both data sets agreed that the autocorrelations are insignificant when the lag is greater than 2-h. However, the 10-min ZWD data set appeared to provide a more precise estimate of the autocorrelation time lag due to the higher sampling rate.

The results above show that the existence of autocorrelations among the tropospheric estimates is evident. The autocorrelations are generally significant for estimates that are within the 1-h to 2-h lag. An average lag value of 1.7 hours is observed at ONSA. Based on these autocorrelation results, the following sections will investigate several possible interpolation methods for ZWD estimates.

3 INTERPOLATION OF ZENITH WET DELAYS

Six models are considered in this study for interpolation of ZWDs. The descriptions of these models are outlined briefly in the following sub-sections. The models range from the easy-to-implement linear interpolant, to an ordinary linear Kriging model. For the purpose of this investigation, the actual and estimated ZWD at time t_i (where $i = 0, 1, \dots, n$) are denoted as $ZWD(t_i)$ and $\hat{ZWD}(t_i)$, respectively, where $(n+1)$ is the total number of ZWD values. A set of $(n+1)$ ZWD observations is denoted by $\{ZWD(t_i)\}_{i=0}^n$; t_k denotes the time at which it is required to interpolate a value for ZWD.

3.1 Linear Interpolation

The linear interpolation (LI) method fits a linear function between each pair of ZWD points $\{ZWD(t_i), ZWD(t_{i+1})\}$ and returns the values of the estimated ZWD, $\hat{ZWD}(t_k)$, at a specified value of time t_k , where $t_i < t_k < t_{i+1}$ (e.g., Benesty et al., 2004). The estimated ZWD can be formulated as:

$$\hat{ZWD}(t_k) = m(t_k - t_i) + ZWD(t_i) \quad (1)$$

where

$$m = \frac{ZWD(t_{i+1}) - ZWD(t_i)}{t_{i+1} - t_i} \quad (2)$$

3.2 Cubic Spline Interpolation

For a set of $\{ZWD(t_i)\}_{i=0}^n$ observations, a cubic spline (CS) ZWD interpolant, $\hat{ZWD}(t_k)$, between the time interval $[t_i, t_{i+1}]$ can be given as (e.g., Burden and Faires, 2004):

$$\hat{ZWD}(t_k) = a_i + b_i(t_k - t_i) + c_i(t_k - t_i)^2 + d_i(t_k - t_i)^3, \text{ for } i = 0, 1, \dots, n-1 \quad (3)$$

where,

$$a_i = ZWD(t_i); b_i = \frac{(a_{i+1} - a_i)}{h_i} - \frac{h_i(2c_i + c_{i+1})}{3}; d_i = \frac{(c_{i+1} - c_i)}{3h_i}; h_i = t_{i+1} - t_i \quad (4)$$

The coefficients $\{c_i\}_{i=0}^{n-1}$ are determined by solving a linear system of equations given by:

$$h_{i-1}c_{i-1} + 2(h_{i-1}h_i)c_i + h_ic_{i+1} = \frac{3(a_{i+1} - a_i)}{h_i} - \frac{3(a_i - a_{i-1})}{h_{i-1}}, \text{ for } i = 0, 1, \dots, n-1 \quad (5)$$

To implement the CS interpolation a minimum of three ZWD observations are needed.

3.3 Cubic Hermite Polynomial Interpolation

For any pair of epochs $[t_i, t_{i+1}]$, the cubic Hermite polynomial (CHP) interpolant, $\hat{ZWD}(t_k)$, between the given points can be estimated as (e.g., Burden and Faires, 2004):

$$\hat{ZWD}(t_k) = ZWD(t_i) + f_1(t_k - t_i) + f_2(t_k - t_i)^2 + f_3(t_k - t_i)^2(t_k - t_{i+1}) \quad (6)$$

where,

$$f_1 = \frac{ZWD(t_i) - ZWD(t_{i-1})}{t_i - t_{i-1}}; f_2 = \frac{f_4 - f_1}{t_{i+1} - t_i}; f_3 = \frac{f_6 - f_2}{t_{i+1} - t_i};$$

$$f_4 = \frac{ZWD(t_{i+1}) - ZWD(t_i)}{t_{i+1} - t_i}; f_5 = \frac{ZWD(t_{i+2}) - ZWD(t_{i+1})}{t_{i+2} - t_{i+1}}; f_6 = \frac{f_5 - f_4}{t_{i+1} - t_i} \quad (7)$$

3.4 Lagrange Polynomial Interpolation

For a set of $\{ZWD(t_i)\}_{i=0}^n$ observations there exists a unique polynomial $P(t)$ of a degree $\leq n$ such that (Burden and Faires, 2004):

$$\hat{ZWD}(t_k) = P(t_k) \text{ for each } i = 0, 1, \dots, n-1 \quad (8)$$

The Lagrange polynomial (LP) is given by:

$$P(t_k) = ZWD(t_0)L_{n,0}(t_k) + \dots + ZWD(t_n)L_{n,n}(t_k) = \sum_{i=0}^n ZWD(t_i)L_{n,i}(t_k) \quad (9)$$

where

$$L_{n,i}(t_k) = \frac{(t_k - t_0)(t_k - t_1) \dots (t_k - t_{i-1})(t_k - t_{i+1}) \dots (t_k - t_n)}{(t_i - t_0)(t_i - t_1) \dots (t_i - t_{i-1})(t_i - t_{i+1}) \dots (t_i - t_n)} = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{(t_k - t_j)}{(t_i - t_j)} \quad (10)$$

3.5 Fast Fourier Transform Interpolation

To use the fast Fourier transform (FFT) method for the interpolation of the ZWDs, a vector of ZWD observations $[ZWD] = \{ZWD(t_i)\}_{i=1}^n$ of length n (sampled at equally spaced points) is firstly transformed to the discrete Fourier transform vector F_{ZWD} using the algorithm given by (Frigo and Johnson, 1998):

$$F_{ZWD} = \sum_{i=1}^n ZWD(t_i) v_n^{(i-1)(t-1)} \quad (11)$$

where v_n is the complex n^{th} root of unity (with $j = \sqrt{-1}$) defined by:

$$v_n = e^{-2\pi j/n} \quad (12)$$

The next step of the process is to calculate the inverse Fourier transform vector $[Z\hat{W}D] = \{Z\hat{W}D(t_i)\}_{i=1}^N$, i.e. the interpolated value) by using the following expression for a user-specified value of N :

$$[Z\hat{W}D] = \left(\frac{1}{N}\right) \sum_{i=1}^N F_{ZWD} v_N^{-(i-1)(t-1)} \quad (13)$$

If $n < N$, the vector F_{ZWD} is padded with trailing zeros to a length of N , prior to applying the inverse transformation defined by Eq. (13). If $n > N$, F_{ZWD} is truncated to the specified length. In this investigation, N is given as:

$$N = n \times (\text{number of interpolated observations}) \quad (14)$$

3.6 Ordinary One-dimension Kriging Interpolation

Kriging's method is known as a best linear unbiased estimator as it estimates the value of a random function at a point as a linear combination of the values at the sample points whilst minimizing the error variance. The method assumes that the closer the input parameters are, the more correlated the observations are. With this concept, it is then worthwhile exploring whether Kriging is appropriate as ZWD interpolator whereby time t is the input parameter. More precisely, the use of *ordinary* Kriging is investigated in its simplest one-dimensional form to determine its usefulness for interpolating ZWD.

Ordinary Kriging interpolation is performed by using a two-component predictor. The first component can be viewed as the generalised least-squares (LS) estimate while the second component is treated as the realisation of a Gaussian process. The ZWD can be modelled at time (t) as (Sacks et al., 1989):

$$ZWD(t) = \sum_{j=1}^p \beta_j h_j(t) + Z(t) \quad (15)$$

where h_j 's are the pre-determined functions of time; p is the number of unknown parameter; β_j 's are unknown coefficients to be estimated. The Gaussian process, $Z(t)$, is assumed to have zero mean and a covariance that can be estimated as:

$$V_t = \text{Cov}(t_{i1}, t_{i2}) = \sigma^2 R(t_{i1}, t_{i2}) \quad (16)$$

between times t_{i1} and t_{i2} ; σ^2 is the a-priori variance of the model in Eq. (16), and $R(t_{i1}, t_{i2})$ is the correlation, whose form can be given by:

$$R_t = R(t_{i1}, t_{i2}) = e^{(-\theta |t_{i1} - t_{i2}|^q)} \quad 0 < q \leq 2 \quad (17)$$

In this study, the variable q is selected to equal two to indicate Euclidian norm, whilst the unknown parameter θ is to be estimated. Additionally, the first component of Eq. (15) can be simplified as an unknown coefficient $\hat{\mu}$, and the ordinary Kriging model can be formulated as (Morris et al, 1995):

$$Z\hat{W}D(t) = \hat{\mu} + Z(t) \quad (18)$$

The use of $\hat{\mu}$, instead of $\sum_{j=1}^p \beta_j h_j(t)$, will result in less computational effort with no significant model degradation (Sacks et al., 1989).

Given a set of times $t = \{t_0, t_2, \dots, t_n\}$ and the corresponding $(n+1)$ vector of ZWD estimates, $ZWD(t) = \{ZWD(t_0), ZWD(t_2), \dots, ZWD(t_n)\}^T$, the best linear unbiased predictor (BLUP) at time t_k can be written as (Sacks et al., 1989):

$$\hat{ZWD}(t_k) = \hat{\mu} + v_{t_k}^T V_t^{-1} (ZWD(t) - H\hat{\mu}) \quad (19)$$

where

$$(V_t)_{ij} = \text{Cov}[Z(t_i), Z(t_j)], v_{t_k}^T = \{\text{Cov}[Z(t_k), Z(t_1)], \dots, \text{Cov}[Z(t_k), Z(t_n)]\}$$

$$\hat{\mu} = ((H^T V H)^{-1} H^T V) ZWD(t) \quad (20)$$

and H being a $(n+1)$ vector of ones. In general, σ^2 and θ in Eqs. (17) and (18) are unknowns. They can be estimated by a method equivalent to the empirical Bayes approach (Koehler and Owen, 1996), which finds the parameters that are most consistent with the observed data. Since $Z(t)$ is Gaussian, the maximum likelihood estimation (MLE) method can be used to estimate σ^2 and θ (Koehler and Owen, 1996). The MLE of σ^2 is given as:

$$\hat{\sigma}^2 = \frac{(ZWD(t) - \hat{\mu}H)^T R_t^{-1} (ZWD(t) - \hat{\mu}H)}{n+1} \quad (21)$$

The maximum likelihood estimation of θ is a one-dimensional optimisation problem of the form:

$$\max_{\theta \in \mathbb{R}^1} (-1/2) [n \ln \hat{\sigma}^2 + \ln(\det(R_t))],$$

subject to $0 \leq \theta \leq \infty$ (22)

A nonlinear optimisation subroutine can usually solve Eq. (22) with respect to the parameter θ (Koehler and Owen, 1996). Once the optimal value of θ is obtained, it can then be substituted back into Eq. (17), and be used to determine V_t and $\hat{\mu}$. The predictor $ZWD(t_k)$ in Eq. (19) can then be completely determined.

4 TESTING AND ANALYSIS OF METHODS FOR ESTIMATING MISSING ZENITH WET DELAY OBSERVATIONS

4.1 Test Description

The purpose of this section is to identify the best method of interpolating ZWD data for missing periods or when processing data at a different interval from which ZWD are available. The performances of all the aforementioned interpolation methods given in Sections 3.1-3.6 are assessed and inter-comparisons between the models are made using the ONSA ZWD data for the period September 10-22, 2004. The ZWD data is determined from a solution of GNSS data of a baseline between the IGS stations Onsala- Wetzell (ONSA-WZTR). The baseline length between the two stations is approximately 920km, which allows the “absolute” tropospheric estimation to be determined. The two stations were constrained to within 0.0001m. The data were collected for a 3-hr window session with sampling interval of the data was 5 minutes. The data were processed with the Bernese GNSS software package (Hugentobler et al., 2001). Data included in the processing also comprises the IGS products concerning the monitoring stations, satellite ephemerides, Earth Orientation Parameters, coordinates and velocity of ground stations, antenna phase centre offsets and variations. During processing, satellite and receiver clock offsets and the tropospheric zenith delay were estimated. The processing parameters include a cut-off elevation angle of 15 degrees,

the use of Niell mapping functions (Niell, 1996) and the Saastamoinen tropospheric model (Saastamoinen, 1972), which was employed to provide a-priori ZTD estimates. The observations were weighted using the elevation-angle dependent model. The ionosphere-free linear combination was implemented to mitigate the ionospheric residual errors. The final GNSS solution produced ZWD with RMSE of 12 mm. Figure 4 shows the time sequence of the ZWD data used in this study.

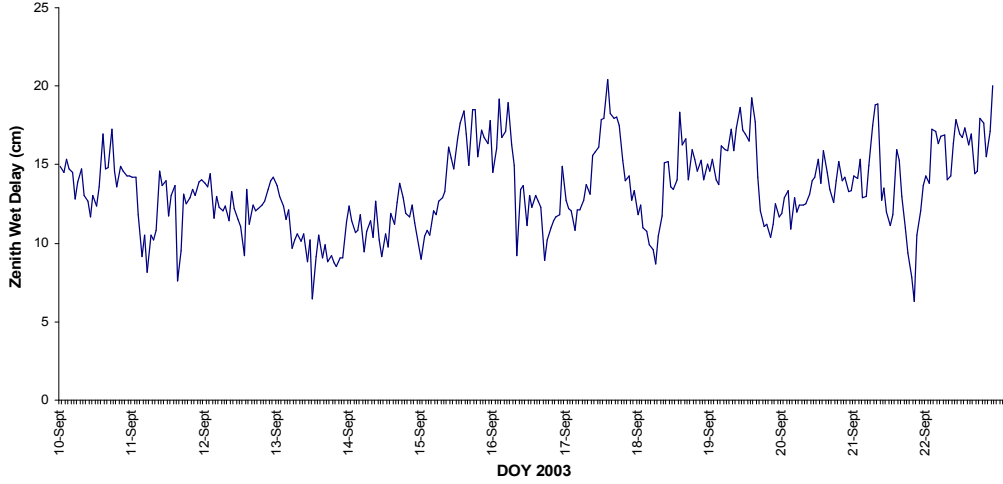


Figure 4: GNSS ZWD estimates at ONSA

The number of ZWD observations used to generate the interpolation models impacts the determination of the model coefficients, and consequently the accuracy of the interpolated ZWD observations. Thus, in this investigation, different sets of observations were used to construct the tested models. The number of pre-determined data points, m_{ZWD} , used to generate these models ranged from 4 to 48, i.e. $m_{\text{ZWD}} \in \{4, 6, 8, 12, 18, 24, 30, 36, 40, 44, 48\}$. Additionally, in each of these runs, the tested models were used to estimate one, two-consecutive, three-consecutive and four-consecutive missing observations, i.e. $1 \leq k_{\text{mis}} \leq 4$. The models were generated and analysed using the following procedure:

- Assuming a total of n observations in the data set, let k_{mis} be the pre-determined number of missing data points, where the interpolation is assumed performed to recover missing data, and m_{ZWD} the number of data points used to generate the model.
- Set $i = \frac{m_{\text{ZWD}}}{2}$.
- Let $\text{ZWD}_{\text{mis}}(t) = \{t_j, \text{ZWD}(t_j)\}_{j=i+1}^{i+k_{\text{mis}}}$ be the consecutively selected missed ZWD data set.
- Let $\text{ZWD}_{\text{obs}} = \{\text{ZWD}(t_{i-k_{\text{mis}}+1}), \text{ZWD}(t_i), \text{ZWD}(t_{i+k_{\text{mis}}+1}), \dots, \text{ZWD}(t_{2i+k_{\text{mis}}})\}$ be the selected data set used to generate the models.
- Generate the IM or LS model based on the data set ZWD_{obs} and estimate the wet delay $\hat{\text{ZWD}}(t_j)$ for $\{t_j\}_{j=i+1}^{i+k_{\text{mis}}}$.
- To assess the model used at any epoch, the difference between the known ZWD at station ONSA, assumed as “truth”, and the estimated ZWD, i.e. $\Delta\text{ZWD}_{\text{diff}}(t) = \hat{\text{ZWD}}(t) - \text{ZWD}_{\text{mis}}(t)$, is computed.
- Similarly, the next missing data points are estimated by shifting one position in time, i.e. $\{t_{j+1}, \text{ZWD}(t_{j+1})\}$ becomes $\{t_j, \text{ZWD}(t_{j+1})\}_{j=i+2}^{i+k_{\text{mis}}+1}$, until the last missing data point has been reached.

The above procedure places the set of missing ZWD observations, $ZWD_{\text{mis}}(t)$, in the centre of the modelling data set, $ZWD_{\text{obs}}(t)$. The first set of missing data begins at time $\{t_{i+1}, \dots, t_{i+k_{\text{mis}}}\}$ and the last set finishes at time $\{t_{n-i-k_{\text{mis}}+1}, \dots, t_{n-i}\}$. In all, a total of $(n - m_{\text{ZWD}} - k_{\text{mis}})$ missing data sets are considered. Given that there are k_{mis} missing observations in each of these sets, the total number of comparisons is therefore, $k_{\text{mis}}(n - m_{\text{ZWD}} - k_{\text{mis}})$. When all cases of missing data sets for a given model have been considered, the RMSE of the $k_{\text{mis}}(n - m_{\text{ZWD}} - k_{\text{mis}})$ ZWD differences are then calculated by:

$$RMSE_i = \sqrt{\frac{\sum_{j=1}^{k_{\text{mis}}(n-m_{\text{ZWD}}-k_{\text{mis}})} [\Delta ZWD_{\text{diff}}(t_j)]^2}{k_{\text{mis}}(n-m_{\text{ZWD}}-k_{\text{mis}})}} \quad (23)$$

Models with low RMSEs were considered as the most ideal interpolation models. For these models, a repeated-measures ANOVA (e.g. Walpole et al., 2007) was implemented as a follow-up test to determine whether there is a significant difference among them. For the repeated-measures ANOVA test, a p-value of less than 0.05 indicates a significant difference among the models.

4.2 Comparisons Between the Interpolation Models

In an effort to determine a suitable model for the purpose of estimating missing ZWDs, the interpolation models outlined in Section 3, were tested. The RMSEs, calculated via Eq. (23), of these models for $m_{\text{ZWD}} \in \{4, 6, 8, 12, 18, 24, 30, 36, 40, 44, 48\}$ are summarised in Tables 3-6. The results show that the Lagrange polynomials (LP) method is the poorest performer. As the number of data points increases, the LP exhibits what is known as Runge's phenomenon (Runge, 1901; Fornberg and Zuev, 2007), and thus produces poor outcomes. Runge's phenomenon is an error problem for a high-order polynomial interpolant on equidistant intervals, whereby the polynomial oscillates towards the end of the interval, as shown in Figure 5, resulting in poor ZWD estimation between the intervals. This effect was more prominent when estimating two, three and four missing ZWD observations.

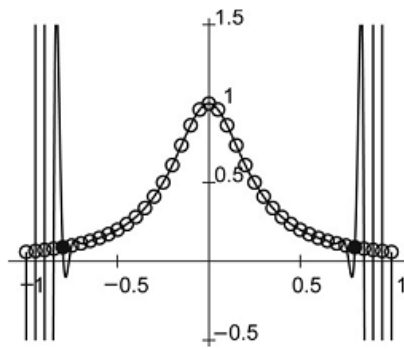


Figure 5 Runge's phenomenon (Fornberg and Zuev, 2007)

Table 3 RMSEs (cm) of the interpolated ZWDs for the case of a single missing observation

Num of Data Pts	Linear	Spline	CHP	FFT	Lagrange	Kriging
4	1.27	1.39	1.30	1.35	1.39	1.27
6	1.27	1.45	1.30	1.36	1.49	1.35
8	1.27	1.47	1.30	1.36	1.56	1.40
12	1.27	1.47	1.30	1.35	1.65	1.37
18	1.27	1.47	1.30	1.33	1.72	1.40
24	1.27	1.47	1.30	1.32	1.76	1.41
30	1.27	1.47	1.30	1.32	1.78	1.41
36	1.27	1.47	1.30	1.32	1.81	1.43
40	1.27	1.47	1.30	1.32	1.82	1.43
44	1.27	1.47	1.30	1.32	1.84	1.42
48	1.27	1.47	1.30	1.32	1.85	1.42

Table 4 RMSEs (cm) of the interpolated ZWDs for the case of two-successive missing observations

Num of Data Pts	Linear	Spline	CHP	FFT	Lagrange	Kriging
4	1.41	1.66	1.46	1.48	1.66	1.41
6	1.41	1.79	1.46	1.51	1.96	1.58
8	1.41	1.87	1.46	1.52	2.28	1.54
12	1.41	1.89	1.46	1.54	2.88	1.58
18	1.41	1.89	1.46	1.54	3.65	1.59
24	1.41	1.89	1.46	1.54	4.26	1.65
30	1.41	1.89	1.46	1.54	4.78	1.68
36	1.41	1.89	1.46	1.54	5.22	1.71
40	1.41	1.89	1.46	1.54	5.48	1.73
44	1.41	1.89	1.46	1.54	5.73	1.75
48	1.41	1.89	1.46	1.53	5.97	1.75

Table 5 RMSEs (cm) of the interpolated ZWDs for the case of three-successive missing observations

Num of Data Pts	Linear	Spline	CHP	FFT	Lagrange	Kriging
4	1.50	1.76	1.53	1.55	1.76	1.54
6	1.50	1.88	1.53	1.55	2.12	1.53
8	1.50	1.94	1.53	1.55	2.54	1.55
12	1.50	1.96	1.53	1.55	3.59	1.66
18	1.50	1.96	1.53	1.54	5.57	1.72
24	1.50	1.96	1.53	1.54	7.76	1.82
30	1.50	1.96	1.53	1.54	10.01	1.88
36	1.50	1.96	1.53	1.54	12.29	1.93
40	1.50	1.96	1.53	1.54	13.81	1.95
44	1.50	1.96	1.53	1.54	15.31	1.98
48	1.50	1.96	1.53	1.54	16.81	1.99

Table 6 RMSEs (cm) of the interpolated ZWDs for the case of four-successive missing observations

Num of Data Pts	Linear	Spline	CHP	FFT	Lagrange	Kriging
4	1.62	2.12	1.68	1.68	2.12	1.65
6	1.62	2.35	1.68	1.70	2.99	1.71
8	1.62	2.48	1.68	1.71	4.02	1.73
12	1.62	2.51	1.68	1.71	6.45	1.81
18	1.62	2.51	1.68	1.71	11.03	1.87
24	1.62	2.51	1.68	1.70	16.62	1.98
30	1.62	2.51	1.68	1.70	23.04	2.06
36	1.62	2.51	1.68	1.71	30.20	2.10
40	1.62	2.51	1.68	1.71	35.42	2.13
44	1.62	2.51	1.68	1.70	41.04	2.16
48	1.62	2.51	1.68	1.70	47.04	2.17

Tables 3-6 also indicate that the linear interpolation (LI), the cubic Hermite polynomial (CHP) and fast Fourier transform (FFT) models (interpolants that are dependent only on the most recent pair of data points) produce better estimates than splines and ordinary Kriging, which estimate the missing data points by giving greater weights to more recent data points, and lesser weights to those that are further away. Kriging did, however, produce comparable results to these models when the number of modelling data is low. The LI model, which was the simplest of all to use, produced the best results across all scenarios. The LI provided, on average, ZWD estimates to within 1.3 cm to 1.6 cm from the actual ZWD data, which corresponds to a PWV error of about 2 mm to 2.5 mm. This level of discrepancy is comparable to many GNSS PWV studies (e.g., Basili et al., 2002; Snajdrova et al., 2006; Wang et al., 2007). Note that both LI and CHP are methods that utilise the two most recent observations, with one on either side of the missing data set.

The favourable results for the LI, CHP and FFT models can be explained by the autocorrelation study in Section 2, whereby significant correlations occur among the estimates within a 1h to 2-h period. Successive 1-h ZWD estimates have an autocorrelation value as high as 0.8 h. Inclusion of several data points that are, time-wise, distant from the estimation time may have introduced noise into the splines and ordinary Kriging models.

Although the LI models appears to be best interpolation method based on its RMSE value, a repeated-measures ANOVA test was necessary for investigating whether it is statistical superior to the CHP and FFT methods. The results of the ANOVA test are given in Table 7. The table shows that the p-values for all the cases considered here were significantly greater than 0.05, which strongly suggest that there is no statistical difference in the performances among the LI, CHP and FFT interpolation models. Hence, the overall RMSE and the ANOVA results summarise that the LI is marginally, but not statistically, better than the CHP and FFT models.

Table 7 P-values of repeated-measure ANOVA test for LI, CHP and FFT interpolation methods

Number of data points used	Number of Missing Observations			
	One	Two	Three	Four
4	0.772	0.893	0.869	0.973
6	0.752	0.832	0.896	0.999
8	0.806	0.866	0.913	0.999
12	0.806	0.863	0.922	0.998
18	0.824	0.866	0.925	0.997
24	0.823	0.87	0.925	0.997
30	0.824	0.87	0.926	0.997
36	0.824	0.873	0.927	0.996
40	0.824	0.873	0.927	0.996
44	0.825	0.874	0.927	0.996
48	0.825	0.875	0.927	0.996

5 CONCLUSIONS

In the efforts to determine the most appropriate models for the interpolation and prediction of ZWD estimates, an autocorrelation analysis of the tropospheric estimates time series was initially carried out on the WVR ZWD estimates at ONSA. The autocorrelation study was restricted to a 12-h time period to ensure a degree of stationarity in the time series. It was found that the time lag for significant autocorrelation was observed within 2 hours. If a Gauss Markov model is assumed, this value can be of significant help in dynamic modelling of ZWD in recursive techniques such as Kalman Filtering.

An investigation into the performance of several modelling techniques was carried out to determine the best approach for estimating missing data points for a set of ZWD observations. Such interpolation processes are needed for post processing applications. For the investigated data set, the RMSE results indicate that the LI model generated the best interpolation results and thus no complex models are needed for the interpolation of ZWD. The favourable results for the LI model, which only depends on the two most recent data points, were reflected in the autocorrelation plot of the ZWD estimates, whereby significant autocorrelation values were observed for up to 2-h only. Although a follow-up ANOVA test indicate that there is no statistical difference in the performances of the LI model in comparison to the CHP and FFT models, the ease-of-use of the LI model ensures that it is still the recommended interpolation method.

REFERENCES

- Basili, P., Bonafoni, S., Mattioli, V., Ciotti, P. and Fionda, E. (2002). A ground-based microwave radiometer and a GPS network for the remote sensing of atmospheric water vapour content: a year of experimental results. 1st COST720 Workshop: Integrated ground based remote sensing stations for atmospheric profiling. l'Aquila, Italy, 18-21 June 2002.
- Benesty, J., Chen, J. and Huang, Y. (2004). Time-delay estimation via linear interpolation and cross correlation. *IEEE Transactions on Speech and Audio Processing*, 12 (5), 509-519.
- Borre, K. and Tiberius, C. (2000). Time series analysis of GPS observables. *Proceedings of the 11th International Technical Meeting of the Satellite Division of The Institute of Navigation*, Salt Lake City, Utah, 19-22 September 2000.
- Burden, R. L. and Faires, J.D. (2004). *Numerical Analysis*. 8th ed. Pacific Grove: Brooks/Cole Publishing Company.
- Fornberg, B. and Zuev, J. (2007). The Runge phenomenon and spatially variable shape parameters in RBF interpolation. *Computers & Mathematics with Applications*, 54 (3), 379-398.

- Frigo, M., and Johnson, S.G. (1998). FFTW: An adaptive software architecture for the FFT. Proceedings of the International Conference on Acoustics, Speech, and Signal Processing 3, Seattle, USA, 12-15 May 1998, 1381-1384.
- Gutman, S. I., and Benjamin, S.G. (2001). The Role of Ground-Based GPS Meteorological Observations in Numerical Weather Prediction. *GPS Solutions*, 4 (4), 16-24.
- Gutman, S. I., Sahm, S.R., Benjamin, S.G., Schwartz, B.E., Holub, K.L., Stewart, J.Q. and Smith, T.L. (2004). Rapid retrieval and assimilation of ground based GPS precipitable water observations at the NOAA forecast systems laboratory: Impact on weather forecasts. *Journal of the Meteorological Society of Japan*, 82 (1B), 351-360.
- Haefele, P., L. Martin, M. Becker, E. Brockmann, J. Morland, S. Nyeki, C. Matzler, and M. Kirchner. (2004). Impact of radiometric water vapor measurements on troposphere and height estimates by GPS. ION GNSS-04 Proceedings, Long Beach, California, 21-24 September 2004.
- Hugentobler, U., S. Schaer, and P. Fridez, eds. (2001). Bernese GPS software version 4.2: Astronomical Institute University of Berne.
- Liu, J., Sun, Z., Liang, H., Xu, X. and Wu, P. (2005). Precipitable water vapor on the Tibetan plateau estimated by GPS, water vapor radiometer, radiosonde, and numerical weather prediction analysis and its impact on the radiation budget. *Journal of Geophysical Research*, 110, D17106.
- Macpherson, S. R., Deblonde, G., Aparicio, J.M. and Casati, B. (2007). Impact of NOAA ground-based GPS observations on the Canadian regional analysis and forecast system. *Monthly Weather Review*, 136 (7), 2727-2745.
- Mattioli, V., Westwater, E.R. Gutman, S.I. and Morris, V.R. (2005). Forward model studies of water vapor using scanning microwave radiometers, global positioning system, and radiosondes during the cloudiness intercomparison experiment. *IEEE Transactions on Geoscience and Remote Sensing*, 43 (5), 1012-1021.
- Niell, A. E. (1996). Global mapping functions for the atmosphere delay at radio wavelengths. *Journal of Geophysical Research*, 101 (B2): 3227-3246.
- Koehler, J. R. and Owen, A.B. (1996). Computer experiments. In *Handbook of Statistics*, ed. S. Ghosh and C. R. Rao, 261-305. New York: Elsevier Science.
- Kuo, Y. H., Zou, X. and Guo, Y. R. (1996). Variational assimilation of precipitable water using a nonhydrostatic mesoscale adjoint model. *Monthly Weather Review* 124 (1): 122-147.
- Runge, C. (1901). Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. *Zeitschrift für Mathematik und Physik*, 46, 224-243.
- Saastamoinen, J. (1973). Contributions to the theory of atmospheric refraction. *Bullétin Géodésique* 105, 106, 107: 279-298, 383-397, 13-34.
- Sacks, J., Welsh, W.J., Mitchell, T. J. and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical Science*, 4 (4), 409-423.
- Smith, T. L., Benjamin, S. G., Gutman, S. I. and Sahm, S. (2006). Short-range forecast impact from assimilation of GPS-IPW observations into the Rapid Update Cycle. *Monthly Weather Review*, 135 (8), 2914-2930.
- Snajdrova, K., Boehm, J., Willis, P., Haas, R. and Schuh, H. (2006). Multi-technique comparison of tropospheric zenith delays derived during the CONT02 campaign. *Journal of Geodesy*, 79 (10-11), 613-623.
- Stoew, B., Nilsson, T., Elgered, G. and Jarlemark. (2007) P. O. J. Temporal correlations of atmospheric mapping function errors in GPS estimation. *Journal of Geodesy*, 81 (5), 311-323.
- Vedel, H. and Huang X. Y. (2003). A NWP impact study with ground based GPS data. The International Workshop on GPS Meteorology. Tsubaka, Japan, 14-17 January 2003.

- Vedel, H. and Huang, X.Y. (2004). Impact of ground-based GPS data on numerical weather prediction. *Journal of Meteorological Society of Japan* 82 (1B), 459-472.
- Walpole, R. E., R. Myers, and S. L. Myers (2007). *Probability and Statistics for Engineers and Scientists*. 8th ed. London, Prentice-Hall.
- Wang, J., Zhang, L., Dai, A., van Hove, T. and van Baelen, J. (2007). A near-global, 2-hourly data set of atmospheric precipitable water from ground-based GPS measurements. *Journal of Geophysical Research*, 112, D11107.
- Wei, W. (2006). *Time series analysis - Univariate and multivariate methods*. 2nd ed. USA, Pearson Education, Inc.

Schema Element Dependencies in a Federated Spatial Database System

Xiaoying Wu; Jianhong (Cecilia) Xia; Geoff West; Bert Veenendaal

Department of Spatial Sciences, Curtin University
Kent St, Bentley, WA, 6102, Australia

wxy_1975@hotmail.com; c.xia@curtin.edu.au; g.west@curtin.edu.au; b.veenendaal@curtin.edu.au

Lesley Arnold

Landgate
1 Midland Square, Midland WA 6056, Australia
lesley.arnold@landgate.wa.gov.au

ABSTRACT

A federated spatial database is the integration of multiple spatial data sources and the realisation of effective spatial data sharing. However, in a federated database environment, database schemas are subject to change.

Some schema elements depend on other schema elements and therefore schema changes result in dependent schema elements becoming invalid in a federated spatial database. Schema element dependencies are more complex in a federated database environment because dependencies include both intra-dependencies (within a database) and inter-dependencies (between databases). This makes management of schema evolution in a federated environment far more difficult than traditional systems.

Schema element dependency (SED) is an important component of managing schema evolution. This paper systematically reviews the schema element dependencies in a federated spatial database system, proposes a SED metamodel to identify schema element dependencies and a SED metadata schema to store schema element dependencies metadata, and explains the methods and steps to generate SED metadata. This paper also presents a SED tool that is used to generate SED metadata and conduct schema change impact analysis. SED provides a foundation for schema evolution management in a federated spatial database.

KEYWORDS: Database schema changes; Federated Spatial Database System; Schema Element Dependency; Schema Element Dependency Metamodel; Impact Analysis

1 INTRODUCTION

A database schema describes the structure of the database. A spatial database schema is an extension of the traditional schema and includes spatial types, indexes and functions (Yeung & Hall, 2007). Spatial database schemas, like traditional schemas, are subject to change due to changes in representation of reality and application requirements. Changes can also occur as a consequence of integration with other systems, compliance to new regulations and the implementation of new security requirements.

Schema change can invalidate the applications that are based on the schema through both the changes in a schema element itself and the dependencies of the changed element. That is because a change of one database schema element might also affect other schema elements and result in other schema elements being invalid. The task of managing schema evolution is to ensure the consistency and integrity of data and application after change.

A Federated Spatial Database System (FSDBS) integrates various geographically distributed spatial data sources and provides a unified data access mechanism that facilitates spatial data sharing (Yeung & Hall, 2007). In a FSDBS, spatial data are shared by multiple organisations and applications across a number of databases. This has been achieved through advancements in networks and communications, distributed computing technologies, and conformance with standards and policies. FSDBS has been an active research subject in dealing with database heterogeneity (Litwin et al., 1990) and can be seen as an important part of Spatial Data Infrastructure (SDI).

A basic FSDBS consists of (Yeung & Hall, 2007): (1) A Federated Spatial Database Management System - an ordinary DBMS that includes a server, a federated database and the system catalogue. It communicates with applications, receives requests from applications, analyses queries and discomposes them into subqueries, sends subqueries to relevant data sources and composes results from data sources then sends back the end users and application. The federated database here is a virtual database which integrates multiple spatial data sources. System catalogue contains schema information of the federated database as well as matching and mapping information between schemas; (2) Spatial data sources which are autonomous and heterogeneous; and, (3) Applications. Applications here could be desktop applications, and, web applications. For example, the federated server can work with a web feature server to provide web mapping services. Figure1 illustrates the basic architecture of a FSDBS.

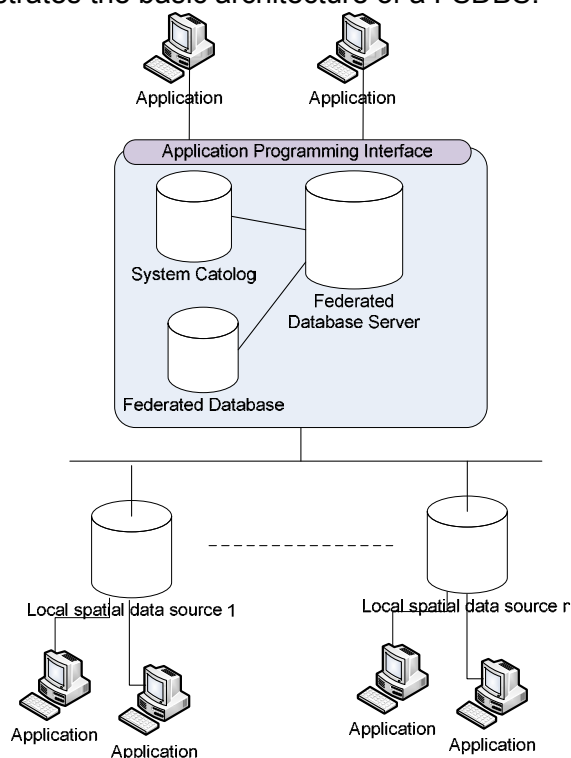


Figure1. Architecture of A FSDBS

In a FSDBS, a local database schema change may affect not only the database itself but also other databases involved in the system. Applications are affected by both local schema changes well as changes elsewhere in the federated schema. This is because the federated schema is dependent on local schemas. Managing schema evolution in a FSDBS environment overcomes the mismatch between applications and the evolved schemas, and maintains consistency of the correspondences between the local schemas and the federated schema before and after schema changes.

Schema Element Dependency (SED) details the dependencies between attributes in the various relations in the databases in a FSDBS. It is important for managing schema evolution. It can help to conduct schema change impact analysis, and to identify affected schema elements after schema changes so as to modify them to adapt to the new schema. Schema element dependencies are more complex in the FSDBS environment as dependencies in such a system

include both inner-dependencies (within a database) and inter-dependencies (between databases). Therefore, management of schema evolution in this system is more challenging.

Currently schema management is performed manually and there is a need to develop methods that can be used to deal with schema evolution automatically or semi-automatically. This research has developed such methodologies, and the conceptual Automatic Schema Evolution (ASE) framework to manage schema evolution in a federated spatial database system in an automatic manner. This paper focuses on Schema SED in the ASE which includes the concept of SED and its functions, SED Metamodel, SED schema in the metadata repository, the methodologies to generate SEDs, as well as the SED tool developed.

2 BACKGROUND

This section briefly introduces some concepts and background information in relation to schema element dependencies.

2.1 Architecture of Automatic Schema Evolution

An Automatic Schema Evolution framework has been developed for managing schema evolution in a FSDBS (Wu et al., 2011b). Figure 2 illustrates the architecture of the ASE. There are five main components included in the architecture: spatial databases, a metadata repository, a view rewriting tool, a schema mapping tool and a schema element dependency tool (Wu et al., 2011a).

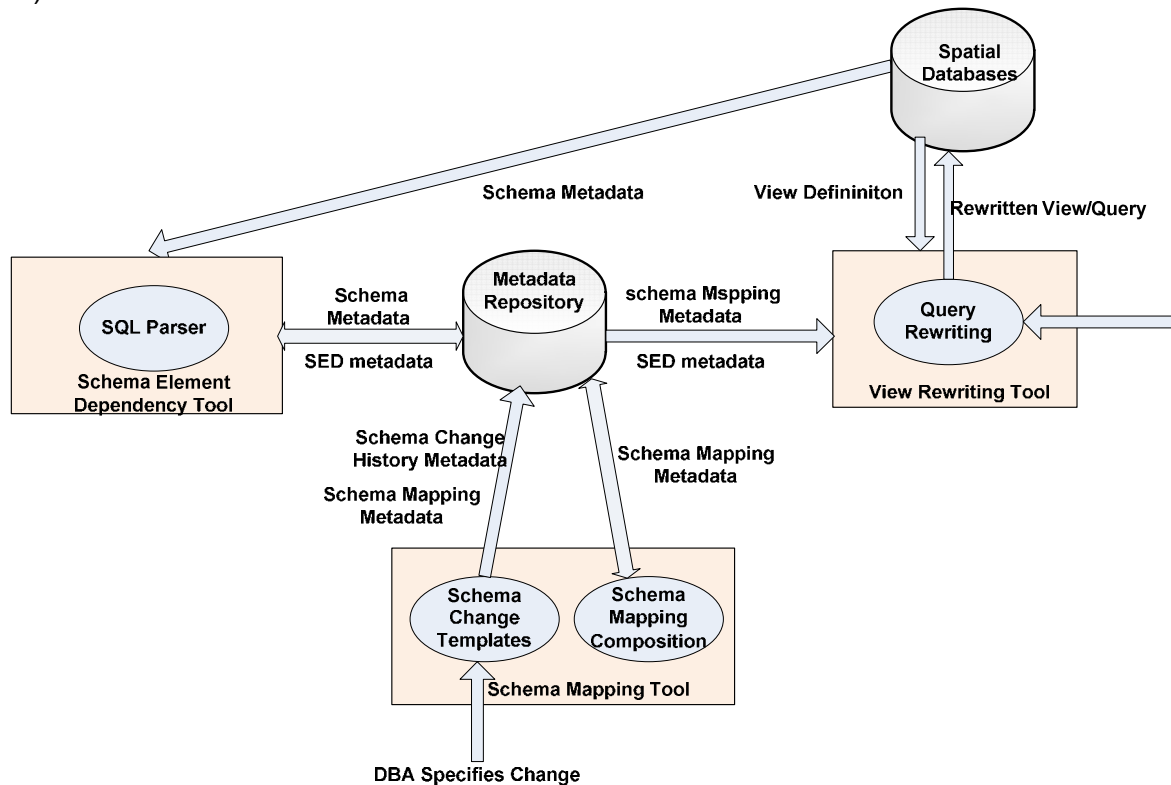


Figure 2. The ASE Architecture

Spatial databases are all databases involved in the FSDBS that include all local spatial data sources and the federated or virtual spatial database as illustrated in Figure 1.

The Metadata Repository is the central repository that stores various types of metadata. Metadata is data about data (Sen, 2004). In this research, metadata is treated as the first class for schema evolution management. The metadata repository provides a consistent and united access mechanism to improve the effectiveness of schema evolution management. Four types of metadata are included: (i) schema metadata; (ii) schema element dependency metadata; (iii) schema mapping metadata; and (iv) schema change history metadata.

The Schema Mapping Tool includes two modules: Schema Change Templates (SCTs) and Schema Mapping Composition. The functionality of this tool is to generate and update schema mapping metadata which then can be used for view/query rewriting.

The SED tool is used to generate SED metadata in the FSDBS, and to update the SED metadata when a schema changes. Based on the SED metadata generated by the SED tool, the schema change impact can be analysed and affected views can be identified when database schema changes.

The View Rewriting Tool that includes the query rewriting module is used to rewrite views against the old schemas into views against the new schema as determined by the schema mapping metadata.

2.2 Spatial Schemas

In order to model the spatial aspects of real world objects, different approaches have been proposed and developed. For example, Modeling Application Data with Spatio-temporal features model (MADS) defines constructs and a language to describe an application schema including thematic and spatial-temporal data structures, and multi-representation aspects (Parent et al., 2006). Spatial object types supported in MADS include both feature and coverage types. ISO 19109 - Rules for application Schema (2005) defines a general feature model to describe the feature object types and their properties including attributes, operations and associations. A Universe of Disclosure described by the constructs and modelling languages defined in MADS or ISO 19101 is a conceptual spatial schema that is platform independent. The defined spatial data types including points, lines and polygons can represent spatial attributes in the conceptual spatial schemas.

With the development of database technologies, spatial object types and their properties, described by the conceptual spatial schema, can be implemented in a relational database. For example, the implementation of feature object types is supported by contemporary DBMSs. The Simple Features Access (SFA) specification (defined in ISO 19125) defines the Structured Query Language (SQL) implementation of the geometry object model that is consistent with ISO 19109 - Rules for application Schema (2005). SFA specifies the SQL schema for storage, query, retrieval and update of spatial data via a SQL interface. SQL implementations include the implementation of primitive data types including the binary data type, and support for extended geometry data types (OGC, 2006a, 2006b). However, coverage objects are not included in the SFA specification so storage, query and management of coverage data are not supported by the SQL interface.

The relationship between the conceptual spatial schema and the SQL spatial schema is explained in Section 3.2: SED Metamodel.

2.3 View and Spatial View

In a traditional database, a view is defined as a stored query, usually as SQL statements. Views can be used to provide a flexible representation of a database by providing users with only data of interest and concealing the rest (Date, 2003). A spatial view extends the traditional view by including a spatial field. A spatial view (or map view) is a live window to the underlying geographic database (Arnold, 2005). Such spatial views are created in the same way as traditional non-spatial views and stored as spatial queries that contain spatial attributes and operations such as spatial selection and spatial join.

2.4 Schema Metadata

In a database, schema metadata is stored in the system catalogue. According to the SQL standard, there are Definition Schema base tables and Information Schema views stored in the database catalogue (ISO/IEC, 2006). The Definition Schema provides a data model for the Information Schema. Examples of Definition Schema tables include *TABLES* and *COLUMNS*. The *TABLES* table stores all tables including views and the *COLUMNS* table has one row for each column. Information Schema views “are viewed tables defined in terms of the base tables of the Definition Schema”(ISO/IEC, 2006). Information Schema views can be accessed the same way as

other tables in the database. Users can query the views if they are granted the privilege. Examples of Information Schema views include *TABLES* view and *COLUMN* view (ISO/IEC, 2006).

In commercial DBMSs, system views are often used to display metadata. For example, in SQL Server, system catalogue views are most widely used for obtaining metadata that is used by the SQL Server Database Engine. All database schema metadata can be derived through catalogue views (MSDN, 2009).

In a federated database, the system catalogue stores both database schema metadata and mapping metadata between the federated database schema and local database schemas.

3 SCHEMA ELEMENT DEPENDENCY

3.1 Schema Element Dependency

Conventionally, a SQL schema consists of schema objects like tables, views, columns, keys and constraints in a relational database (Stephens & Plew, 2001). Some schema objects refer to others when they are defined. For example, a view refers to one or more tables and views. Normally, the object that is being referred to is called a referenced object. The one which references other objects is called a dependent object (Oracle, 2011). The dependency can occur between different types of schema objects such as table to table, view to table, view to view, and trigger to table etc.

If one schema of an object is changed, the dependent objects may become invalid. For example, a change to a base table can affect the view when a view references the base table: (1) If the base table is renamed or dropped, the view will become invalid; (2) if the columns in the base referenced by the view are renamed or dropped, the view will become invalid too; (3) however, any changes on the columns that are not referenced by the view will have no impact on the view. Dependency information at the level of schema objects, however, can't provide a clear answer to the impact of schema changes. In order to better manage schema evolution, dependencies of interest in the research are taken to be column level dependencies which are on a finer level.

In order to represent column level dependencies, the concept of schema element has been introduced in this paper. Schema elements include both schema objects and columns. Schema elements of interest in the research are mainly tables, views and columns. With schema elements, column dependencies and view dependencies can be represented.

A column dependency represents the dependency between two columns. For example, a computed column is derived from other column(s) in the same table by an expression and is a virtual column. Therefore, the computed column is dependent on the columns in the expression. Column dependencies also occur when one column in a table references the primary key or candidate key in another (or the same) table via a foreign key, also called referential integrity.

In a spatial database, a view (traditional and spatial view) is defined in the form of a SQL statement. Dependency between a view and a column of a table/view occurs when the query defined in the view references the column of the table/view. Dependencies between a view and columns of tables/views are called view dependencies in this paper.

In a federated database system, schema element dependencies occur not only within a database, but also across databases. Since the federated database is a virtual database based on other data sources, tables in the federated database refer to tables in other databases.

Finding all the dependencies in a database system is very important in terms of minimising error, avoiding missing information and ensuring consistency in the database. With accurate and detailed schema element dependency information, a schema change impact analysis can be conducted before schema changes; and affected schema elements can be identified and then modified accordingly to adapt to the changed schema.

3.2 SED Metamodel

The SED metamodel (illustrated in Figure 3 using UML syntax) is used to represent column level schema element dependencies in a FSDBS. This conceptual metamodel consists of three

parts: the schema elements of a conceptual spatial schema, the schema elements of a SQL schema, and the SEDs.

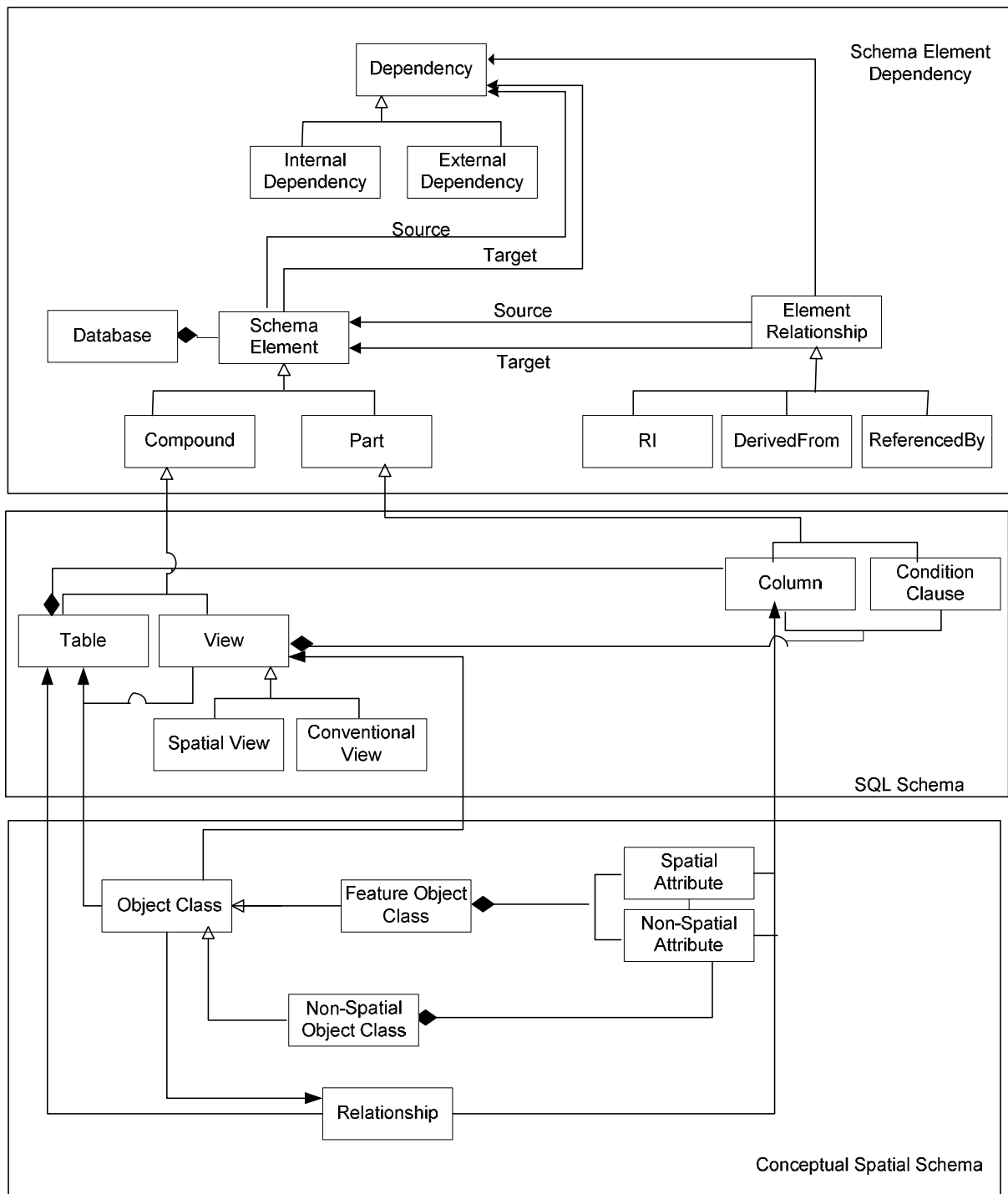


Figure 3. SED Meta-model

The conceptual spatial schema is a conceptual data model adapted from the generic feature model from ISO 19109 (2005). Schema elements defined in the model are object classes consisting of feature and non-spatial object classes, attributes including spatial and non-spatial attributes, and relationships. A feature object class represents a collection of vector objects of the same type. A non-spatial object class represents a category of objects with no spatial characteristics. An object class has a name and attributes including spatial (for a spatial object class) and non-spatial attributes. Each attribute has its own properties including name and data

type etc. Spatial attribute data types are points, lines or polygons. Object classes can have relationships between them.

The SQL schema is an implementation of a conceptual spatial schema according to SFA specifications (OGC, 2006a, 2006b). Schema elements include tables and views that are supported by contemporary relational or object-relational DBMSs. A table contains the name of the table, names of all columns and data types for these columns. A view includes the name of the view and names of columns. These are required because a view is a virtual table that is derived from base tables or views by means of stored queries.

There are three types of table in a DBMS: feature table, geometry table and non-spatial table. A feature table stores a collection of features of the same type. Attributes of a feature object class including the spatial attributes are columns of a feature table. A geometry table stores the geometric information of the geometric objects. Tables without any spatial element are called non-spatial tables.

The relationship between the conceptual spatial schema and the SQL schema can be seen in Figure 3. Object classes are stored as tables in the relational database. More specifically, the implementation of a non-spatial object class is a non-spatial table while the implementation of a feature object class varies depending on the data types supported by the underlying DBMS. For example, a feature object class uses a feature table and a geometry table, with a key reference between them if the predefined data types in the DBMS are used. However, only the feature table is needed if geometry data types are supported by the DBMS. Attributes of an object class are stored as columns of the corresponding table. A relationship between two object classes is implemented as either a table or the primary key and foreign key column of the two tables.

Schema element dependencies consist of schema elements and the relationships between them. Schema elements are divided into *Compound* and *Part*. The *Part* component denotes atomic elements such as a column. The *Compound* component represents the elements that contain parts such as tables and views. Relationships include *Referential Integrity (RI)*, *DerivedFrom* and *ReferencedBy*. A relationship connects two schema elements. The dependency information, termed a column level dependency, can be derived from the relationships between the source and the target elements that are connected by the relationship. Depending on the databases containing the source and target elements, the dependency will be identified as an internal or external dependency.

3.3 Schema and SED Metadata Schema

As mentioned in the background section, there are four types of metadata stored in the metadata repository: (i) schema metadata; (ii) schema element dependency metadata; (iii) schema mapping metadata; and (iv) schema change history metadata. SED metadata in the metadata repository, as shown in Figure 2, is used for schema change impact analysis before schema changes and detection of affected schema elements after schema changes. Figure 4 is the schema of database schema metadata and SED metadata.

Among the seven tables shown in Figure 4, five tables are used to store database schema metadata: SED_Server, SED_Database, SED_Owner, SED_Tabview and SED_Column. Based on the names of these tables in the schema, it is not hard to understand what metadata is stored in each table: SED_server stores information on all servers involved in the FSDBS; SED_Database is for information on all spatial databases participating in the FSDBS; SED_Owner is used to store details of the owners of tables/views in each database; SED_Tabview is for all tables and views; and, SED_Column is for all columns in each table/view.

The other two tables in Figure 4: SED_Viewd and SED_Columnnd are used to store different types of SED metadata. The SED_Viewd table stores dependencies between views and columns of tables/views. In this table, the dependent element is the view while the referenced element is the column in a table/view. The SED_Columnnd table stores dependencies between columns generated by computed columns or referential integrity. In SED_Columnnd, the dependent and referenced elements are both columns and these columns may be in the same table or in different tables.

Schema metadata can be extracted from the original database system catalogue. The generation of SED metadata varies depending on the type of dependency. Details of the generation of SEDs are explained in section 4: Generation of SEDs.

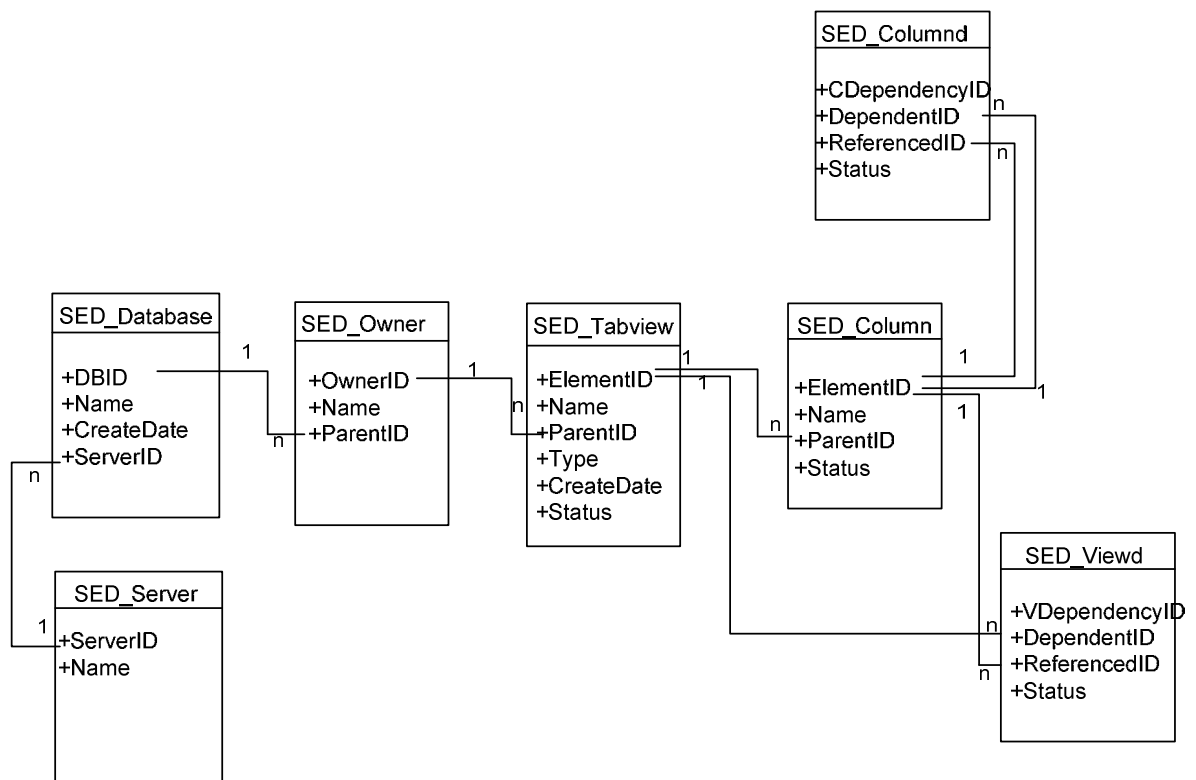


Figure 4. Schema and SED Schema

4 GENERATION OF SEDS

Different approaches have been used to generate different types of dependencies.

4.1 View Dependencies

To generate dependencies between a view and base tables (or views) the view definition is analysed. View definitions are SQL statements stored in the system catalogue that can be extracted using a SQL statement.

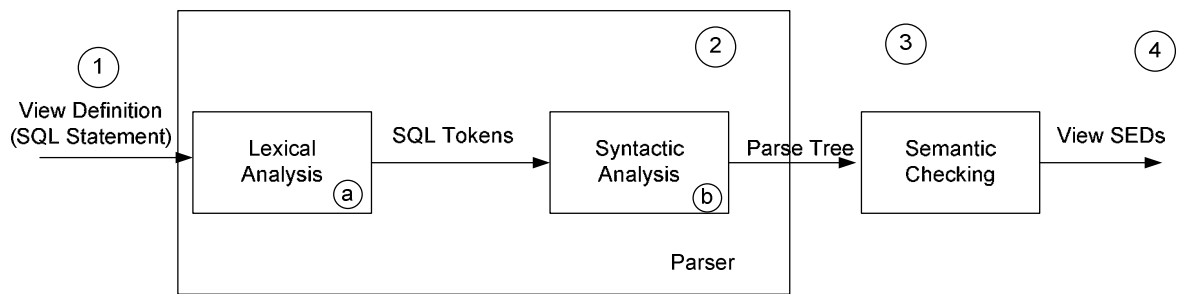


Figure 5. Steps to generate View SEDs

The steps to generate view SEDs are shown in Figure 5:

1. View metadata extraction: View definitions are stored in the system catalogue and a view can be derived by a SQL statement.
2. Parsing the view definition SQL statement to generate a parse tree. Two stages are involved in this step: lexical analysis and syntactic analysis:
 - a) Lexical analysis is the process of breaking text into a set of tokens. The input of this stage is the text written in a certain language, and, the output is a set of tokens. In SQL, tokens include ordinary and delimiter tokens. Ordinary tokens include numeric constants, ordinary identifiers, host identifiers and key words. Delimiter tokens include string constants, delimited identifiers, operator symbols and other special characters (IBM, 2011). Examples of key words are SELECT and FROM. Identifiers are schema elements such as tables, columns, databases. After tokenisation of the SQL statement a sequence of SQL tokens is formed, which are the input of the syntactic analysis.
 - b) Syntactic analysis is the process to generate a grammatical structure (normally called a parse tree) based on the tokens formed by lexical analysis according to the grammar of the language.
3. Semantic checking. The parse tree is the syntactic representation of the view definition. In order to generate view SEDs, semantic checking has to be conducted on the parse tree. This can be done by consulting the database schema metadata to determine the relationships between schema elements and hence the dependency relationships between schema elements.
4. Insert the view SEDs information in the SED-Viewd table in the metadata repository.

View dependencies generated with this approach include both internal and external database dependencies. A federated database is a virtual database that can be seen as views built against component databases. Therefore, dependencies between schema elements in the federated database and component databases can be similarly generated.

This approach can also be extended to generate dependencies caused by triggers, procedures or function since they are mainly composed of SQL statements. They are outside the scope in this paper.

4.2 Column Dependencies

Column dependency occurs when there is a computed column or referential integrity. Steps to generate column dependencies caused by a computed column is similar to view dependencies and include: (1) extraction of the definition of the computed column from the system catalogue; (2) analysis of the definition according to the SQL standard and extract schema elements in the definition; and (3) insert the dependent element (the computed column) and referenced elements (extracted from the definition) into the SED_Column table in the metadata repository.

The generation of column dependencies caused by integrity is relatively straightforward. It can be achieved by extracting the dependency data from the system catalogue and inserting it into the SED_Column in the metadata repository.

4.3 SED tool

In order to generate and update SEDs automatically, a SED tool has been developed in the .NET framework using C# based on the methodologies described above. The functions of the tool include, as shown in Figure 6: (1) extraction of schema metadata from the system catalogue and insert into metadata repository; (2) generation or update of SEDs including view SEDs and column SEDs; and (3) analysis of schema change impact.

The SED tool can be used to extract different ranges of schema metadata depending on what is specified in the interface. For example, if a database is specified, all schema metadata in the database can be extracted, and if a table is specified, only schema metadata related to the table will be extracted. By these means, the schema metadata in the metadata repository can be synchronised with the related spatial databases.

The main function of the tool is to generate SEDs metadata, as indicated by the name of the tool. This is mainly achieved by the SQL parser imbedded with the tool, as shown in Figure 5. The SQL parser is included to parse the SQL statements such as the view definition and column definition. SEDs can be generated when a new schema element is created and updated when a schema element is changed. For example, when a new view is created, the SEDs for the view can be generated by specifying the view in the user interface.

With the SED metadata stored in the metadata repository, the SED tool can also be used to conduct impact analysis of schema changes when the changing schema element is specified. For example, when a column of a table is specified in the interface, all schema elements that depend on the column will be displayed.

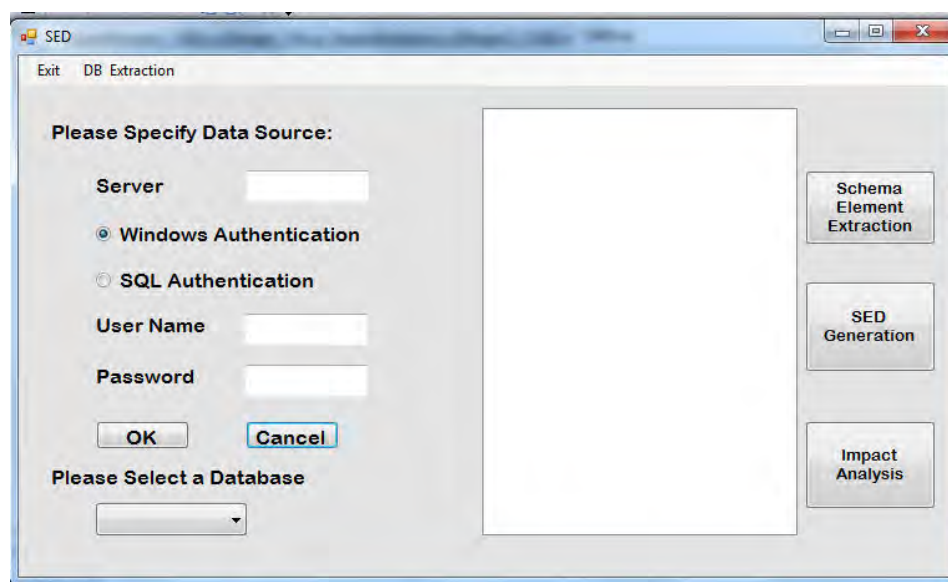


Figure 6. SED Tool

4.4 Implementation of SED tool in a FSDBS

A test environment involving a sample federated spatial database system has been set up to test the SED tool. The system comprises a federated ArcGIS geodatabase G and two component geodatabases S1 and S2. Figure 7 illustrates the schemas of geodatabases involved in the sample system. The underlying relational DBMS used is SQL Server 2008. S1 consists of two feature object classes - *ExistingMainRoads* and *ProposedMainRoads*, and one non-spatial object class - *RoadType*. A View Highway displays all highways including both existing and proposed. S2 consists of one feature object class - *LocalRoads* and one non-spatial object class - *FeatureNames*. The federated schema G includes one feature object class - *RoadSegments* that displays all road segments including both main roads and local roads.

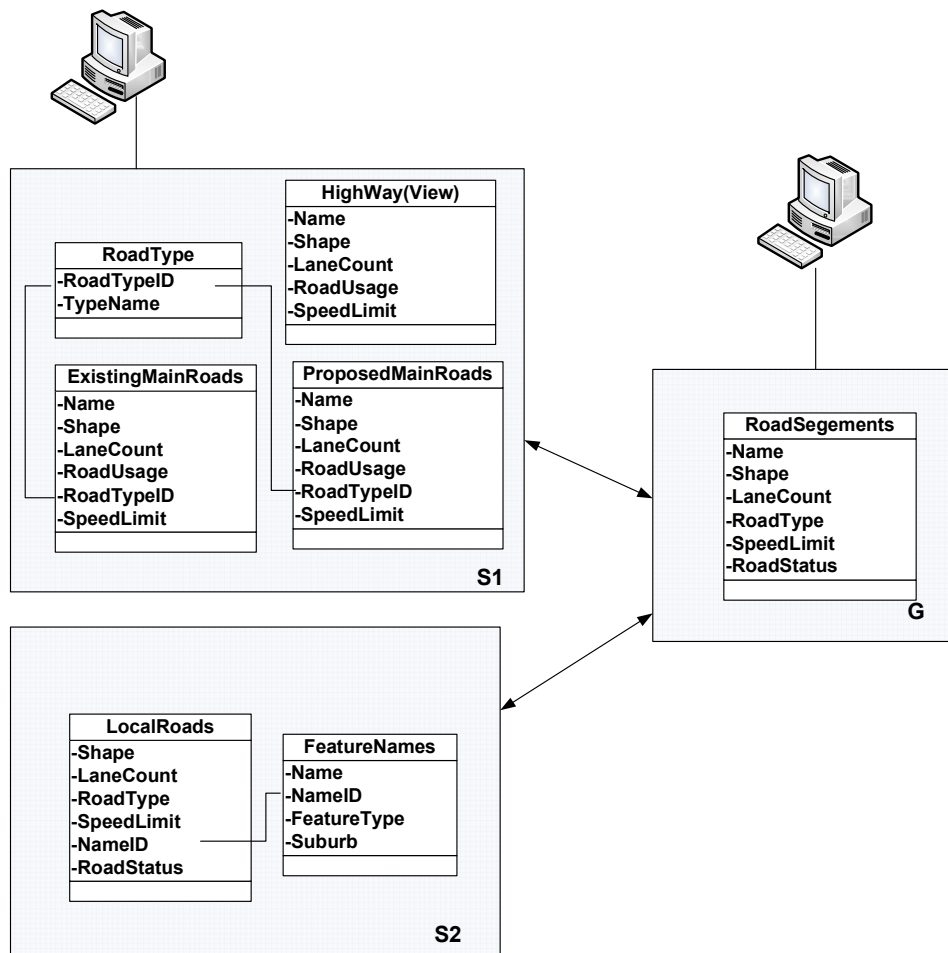


Figure 7. Schema of the Sample Federated Spatial Database System

The tool extracted schema elements from the two local spatial databases and the federated spatial database, generated view dependencies and column dependencies including both inner-dependencies and inter-dependencies, and detected the affected elements if a schema element is specified. Snapshots of view dependencies and column dependencies generated by the tool are shown in Figure 8 and Figure 9 respectively. Figure 8 shows the dependencies for each view. For example, the first 14 rows are the dependencies of the view “Highway” on the columns of other tables, and the first two rows mean the view is dependent on column “LaneCount” from table “PROPOSEDMAINROADS” and “EXISTINGMAINROADS”. Figure 9 shows the columns dependencies e.g. the first row shows there is a dependency of column “RoadTypeID” of table “PROPOSEDMAINROADS” on column “RoadTypeID” of table “ROADTYPE”.

DependentView	Downer	Ddatabase	Dserver	ReferencedColumn	Rtabview	Rowner	Rdatabase	Rserver
Highway	dbo	S1	GIS\ARCSDE_SQL	LaneCount	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	LaneCount	EXISTINGMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	Name	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	Name	EXISTINGMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	RoadTypeID	ROADTYPE	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	RoadTypeID	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	RoadTypeID	EXISTINGMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	RoadUsage	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	RoadUsage	EXISTINGMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	SHAPE	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	SHAPE	EXISTINGMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	SpeedLimit	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	SpeedLimit	EXISTINGMAINROADS	dbo	S1	GIS\ARCSDE_SQL
Highway	dbo	S1	GIS\ARCSDE_SQL	Type Name	ROADTYPE	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	LaneCount	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	Name	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	RoadTypeID	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	RoadUsage	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	SHAPE	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	SpeedLimit	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	RoadTypeID	ROADTYPE	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	Type Name	ROADTYPE	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	LaneCount	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	Name	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	RoadTypeID	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	RoadUsage	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	SHAPE	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	SpeedLimit	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	Name	FEATURENAMES	dbo	S2	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	NameID	FEATURENAMES	dbo	S2	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	LaneCount	LOCALROADS	dbo	S2	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	NameID	LOCALROADS	dbo	S2	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	RoadStatus	LOCALROADS	dbo	S2	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	RoadType	LOCALROADS	dbo	S2	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	SHAPE	LOCALROADS	dbo	S2	GIS\ARCSDE_SQL
RoadSegment	dbo	G	GIS\ARCSDE_SQL	SpeedLimit	LOCALROADS	dbo	S2	GIS\ARCSDE_SQL

Figure 8. View Dependencies in the Sample FSDBS

DColumn	Dtabview	Downer	Ddatabase	Dserver	Rcolumn	Rtabview	Rowner	Rdatabase	Rserver
RoadTypeID	PROPOSEDMAINROADS	dbo	S1	GIS\ARCSDE_SQL	RoadTypeID	ROADTYPE	dbo	S1	GIS\ARCSDE_SQL
RoadTypeID	EXISTINGMAINROADS	dbo	S1	GIS\ARCSDE_SQL	RoadTypeID	ROADTYPE	dbo	S1	GIS\ARCSDE_SQL
NameID	LOCALROADS	dbo	S2	GIS\ARCSDE_SQL	NameID	FEATURENAMES	dbo	S2	GIS\ARCSDE_SQL

Figure 9. Column Dependencies in the Sample FSDBS

5 CONCLUSIONS

Schema evolution is essential in a FSDBS because it is dependent on databases owned, maintained and, importantly, modified by other organisations. Effective management of schema evolution is an integral part of a SDI. It ensures that discovery and access to spatial data and services is consistent even though schema changes have occurred.

Managing schema evolution in such an environment is a significant challenge. To manage this effectively, it is important to identify all existing schema element dependencies and to maintain this over the lifetime of schema changes. This paper discusses the SEDs including the concept, types and function of SED and methodologies to generate SEDs in a FSDBS. A SED tool has also been developed to extract and synchronise schema metadata, generate and update SEDs, and analyse schema change impact and detect affected schema elements when schema changes.

With the SED metadata generated, further work can be conducted on the components of the Schema Mapping Tool and View Rewriting Tool, as illustrated in Figure 2, which includes developing schema change templates, generating and storing schema mapping metadata, schema mapping composition, and developing tool to rewrite views. These aspects combined will further automate and enhance the management of schema evolution in a FSDBS.

ACKNOWLEDGEMENT

The authors would like to thank Landgate and Curtin University for providing funds to support this PhD research.

REFERENCES

- Arnold, L. (2005). *Dynamic Spatial Updating: A Scale-independent Approach to Data Management and Thematic Map Revision*. PhD Research, Curtin University of technology.
- Date, C. J. (2003). *An Introduction to Database System - 8th Edition*: Addison Wesley.
- IBM. (2011). Tokens Retrieved 03/07, 2011, from <http://publib.boulder.ibm.com/infocenter/dzichelp/v2r2/index.jsp?topic=/com.ibm.db2.doc.sqlref/rch2tok.htm>
- ISO. (2005). ISO 19109 *Geographic information -- Rules for application schema*
- ISO/IEC. (2006). Information technology - Database Languages - SQL - Part 11: Information and Definition Schemas (SQL/Schemata).
- Litwin, W., Mark, L., & Roussoupoulos, N. (1990). Interoperability of Multiple Autonomous Databases. *ACM Computing Surveys*, 22(3), 267-293.
- MSDN. (2009). Catalog Views (Transact-SQL) Retrieved 14/01, 2010, from <http://msdn.microsoft.com/en-us/library/ms174365.aspx>
- OGC. (2006a). OpenGIS® Implementation Specification for Geographic information - Simple feature access - Part 1: Common architecture. Retrieved from http://portal.opengeospatial.org/files/?artifact_id=25355
- OGC. (2006b). OpenGIS® Implementation Specification for Geographic information - Simple feature access - Part 2: SQL option. Retrieved from http://portal.opengeospatial.org/files/?artifact_id=25354
- Oracle. (2011). Oracle® Database Concepts --11g Release 1 (11.1) Retrieved 10/05, 2011, from http://download.oracle.com/docs/cd/B28359_01/server.111/b28318/dependencies.htm#i1234
- Parent, C., Spaccapietra, S., & Zimányi, E. (2006). Conceptual Modeling for Traditional and Spatio-Temporal Applications, the MADS Approach. *Springer*.
- Sen, A. (2004). Metadata management: past, present and future. *Decision Support Systems*, 37(1), 151-173.
- Stephens, R., & Plew, R. (2001). *Database Design*: Sams Publishing.

- Wu, X., Xia, J., West, G., Arnold, L., & Veenendaal, B. (2011a). *An Architecture of Managing Schema Evolution in a Federated Spatial Database System*. Paper presented at the 7th International Symposium on Digital Earth, Perth.
- Wu, X., Xia, J., West, G., Arnold, L., & Veenendaal, B. (2011b). Managing Schema Evolution in a Federated Spatial Database System *Discovery of Geospatial Resources: Methodologies, Technologies and, Emergent Applications*: IGI Global.
- Yeung, A. K. W., & Hall, G. B. (2007). *Spatial Database Systems - Design, Implementation and Project Management* (Vol. 87): Springer.

ISBN: 978-0-478-11042-1



Land Surveying / Remote Sensing & Photogrammetry / Engineering & Mining / Hydrographic Surveying / Spatial Information & Cartography