ARTIFICIAL INTELLIGENCE INFORMED SIMULATION OF DISSOLVED INORGANIC NITROGEN FROM UNGAUGED CATCHMENTS TO THE GREAT BARRIER REEF

A Thesis submitted by

Cherie O'Sullivan

BEnvSci(Hons)

For the award of

Doctor of Philosophy

2023

ABSTRACT

Land based sources of nutrient loads impact the health and resilience of the Australian Great Barrier Reef, however, the current methods used to quantify and prioritise nutrient reduction to the reef need improvement to increase certainty in estimates of Dissolved Inorganic Nitrogen (DIN) from ungauged catchments. Catchment Scale Water Quality Models are currently the primary tools used to quantify the influence of landscapes towards receiving waters and are effective for communication of the influences of the landscape and its management towards the Great Barrier Reef. The design and development of these models rely on extensive observed water quality data for development and calibration of the models, however, the collection of the data are both expensive and not possible in all areas. This PhD project has developed new knowledge in simulating DIN from ungauged catchments, to overcome the challenge and knowledge gaps associated with data voids that afflict water quality modelling. Research herein has coupled catchment classification, a method demonstrated by existing the literature to effectively overcome data voids for flows, with Artificial Intelligence pattern matching and techniques to identify corroborating catchment matches for both DIN patterns and spatial data. Additionally, this research, *for the first time*, has used spatial datasets for Original Vegetation, as a proxy dataset to the drivers of DIN. This research has found that the Original Vegetation data represents the variability in biological response to the drivers of heterogeneity in DIN patterns across the landscape. Explainable artificial intelligence approaches were then developed to identify landscape features most influential in the classification results. Development of these methods ultimately facilitated satisfactory simulation of DIN for a pseudo ungauged catchment as well as identifying catchments that are unsuitable to share data and others that need prioritisation for future gauging programs. Together, these approaches have enabled the development of knowledge to classify ungauged catchments of the Great Barrier Reef using spatial data as a proxy for absence of observed DIN data. The findings of this doctoral study have provided new insights into water quality modelling and the selection of catchments as well as classifying the catchments and performing DIN simulations.

CERTIFICATION OF THESIS

I, Cherie Moira O'Sullivan, declare that the PhD Thesis entitled *Artificial Intelligence informed Simulation of Dissolved Inorganic Nitrogen from Ungauged Catchments to the Great Barrier Reef* is not more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references, and footnotes.

This Thesis is the work of Cherie Moira O'Sullivan except where otherwise acknowledged, with the majority of the contribution to the papers presented as a Thesis by Publication undertaken by the student. The work is original and has not previously been submitted for any other award, except where acknowledged.

Date: 25 October 2023

Endorsed by:

Professor Ravinesh Deo
Principal Supervisor (January 2023-Submission)

Dr Afshin Ghahramani
Principal Supervisor (August 2020-January 2023)

Associate Professor Keith Pembleton
Principal Supervisor (July 2019- August 2020)

Student and supervisors' signatures of endorsement are held at the University.

STATEMENT OF CONTRIBUTION

The doctoral research thesis has produced three quartile 1 (*Q*1 ranked) journal publications completed during the PhD candidature. Articles 1, 2, and 3 are the primary (core) parts of this thesis. The following presents the student contributions and the contributions of the co-authors of the publications.

Paper 1:

O'Sullivan, C. M., Ghahramani, A., Deo, R., Pembleton, K., Khan, U., Tuteja N. (2022). Classification of catchments for nitrogen using Artificial Neural Network Pattern Recognition and spatial data, *Science of The Total Environment*. *Volume 809,* Article 151139. https://doi.org/10.1016/j.scitotenv.2021.151139

O'Sullivan, C. M contributed 70% to this paper. Collectively Ghahramani, A., Deo, R., Pembleton, K., Khan, U., Tuteja N. contributed the remainder.

Paper 2:

O'Sullivan, C. M., Ghahramani, A., Deo, R., Pembleton, (2023). Classification of catchments for nitrogen using Artificial Neural Network Pattern Recognition and spatial data, *Science of The Total Environment*. *Volume 861,* Article 160240. https://doi.org/10.1016/j.scitotenv.2022.160240

O'Sullivan, C. M contributed 80% to this paper. Collectively Ghahramani, A., Deo, R., Pembleton, K., contributed the remainder.

Paper 3:

O'Sullivan, C. M., Deo, R., Ghahramani, A. (2023). Explainable AI approach and original vegetation data classifies spatio-temporal nitrogen in flows from ungauged catchments to the Great Barrier Reef, *Scientific Reports* Volume 13, Article 18145 (2023). https://doi.org/10.1038/s41598-023-45259-0

O'Sullivan, C. M contributed 90% to this paper. Collectively Ghahramani, A., Deo, R., contributed the remainder.

ACKNOWLEDGEMENTS

My journey into this PhD candidature has been a fantastic opportunity, eye opening experience and colliding of many circumstances. At the organisational level, this research has been supported by the Australian Government Research Training Program Scholarship, therefore my sincere gratitude to the University of Southern Queensland and the Australian Federal Government who facilitate that support. Further appreciated support was also provided in the form of top-up funding and professional development opportunities by the State of Queensland Government via Queensland Water Quality Modelling Network Initiative.

On the personal level – I want to start by thanking Simon Newcombe whose enthusiasm for engaging all youth in rowing, regardless of their background, amplified my connection with our local river, shaped my desire to look after it, and led to my drive for this research. Ben McMullen opened the door to my professional Environmental Science career and connected me to the Healthy Land and Water steering committees. This insight kept me up to date and inspired by water quality management in Queensland. Dr Paul Maxwell and Alan Hoban's engaging communication style regarding water quality models, as well as all the other committee members positive energy has been motivating. Cr Brian Stockwell is integral to my passion to overcome the limitations of water quality models for informing land management investments. Concurrently, Dr Mike Olsen's enthusiastic communication of ecological knowledge contained within remnant vegetation became foundational to this research.

In delivery of this thesis, foremost I wish to sincerely thank the time, dedication, commitment, and passion of all my supervisors. Dr Afshin Gharamani and Associate Professor Keith Pembleton for collaborating their skills and connections to embed the project within the QWMN Innovation Associate program, engage exceptional expertise of Professor Ravinesh Deo, and welcome me to the team. I greatly appreciate Assoc. Prof. Keith's professional advice, level approach and unbiased support throughout all phases, particularly the foundational confirmation of candidature to ensure all research can pass the quality test. Indebted to Dr Afshin Ghahramani for believing in me, passionate drive for outcomes, as well as genuine investment towards evolving my systems approach, research communication skills and confidence to ensure delivery of high-quality research for this PhD, and future endeavours. Ultimately grateful to Prof. Ravinesh for foundational training in Artificial Intelligence, continual positive encouragement of all students in the Data Analytics team to reach for the stars, support each other and create ongoing opportunities throughout. Very grateful for Prof. Ravinesh's energising positivity always, and particularly his support on the home run.

This PhD is a collaboration between the University of Southern Queensland, State of Queensland Department of Environment and Science Queensland Water Modelling Network (QWMN), and Bureau of Meteorology who together developed and supported this PhD research. Exceptional personalities in QWMN driven to achieve common goals include Associate Professor Brian McIntosh, for advocating for industry embedded research in Australia, Dr Piet Filet for the many motivating professional development workshops and research communication opportunities, and Sarah Cochrane for making these happen. Jenny Riches for passionate and caring support for the QWMN program. The Innovation Associates cohort Chinenye Ani, Filipe Pinhati, Chao Deng and Laura Bellis, great to share the experience together. Heartfelt dedication to both Dr John Brodie and Mark Pascoe for their passion and taking the time to instil knowledge of water quality management to all new researchers, but also setting the scene for the warm, unique and passionate culture within in the industry. I am grateful to have crossed both their paths enough for inspiration, but devastated it was so short. From Bureau of Meteorology as industry partner, Dr Urooj Khan, Richard Laugesen, Dr Narendra Tuteja and Dr Mohammad Hasan. In addition to data sharing, the collaboration, informed discussions, welcome to the project, conferences and professional insights were always motivating.

UniSQ is a powerhouse of many motivated and inspirational students and staff outside the supervisory team also. This big data research was made possible on the High Performance Computer thanks to the tireless assistance of Richard Young and Dr Francis Gacenga from the UniSQ Office of Research HPC. Also Chris Davey and Dr Mohanad Shakir Khalid Al-Musaylh, who's coding assistance became invaluable to this research. Dr Mohammad Kaisb, Dr's Adham and Fatima Ali, whose attitude and passion are inspirational. Ibrar Hussain, my first uni class mate and now very dear friend, our competitiveness was hilarious, your drive with a sense of humour through anything thrown at you is admirable. Bidhan Nath, wise and very dear from the orientation week astronomy workshop to the stars & beyond. Your insightful, genuine inclusiveness opened my world to enjoy the full UniSQ PhD student experience, and what matters most – family - invaluable. So glad you can be reunited with yours again. Jake the positive motivating star in the Astrophysicist office era. Dr Douglas Eacersall and Professor Patrick Danaher for providing introductory grounding during the confirmation period. The UniSQ refec staff for friendly welcomes to campus, coffee and nourishment. Hanlu Zhang for the experiences we shared including with Dr Aram Ali welcoming me to share enjoying your fieldwork. Jo Owens for the check-ins and industry updates, Dr JP Emery, Krystal Watson, Professor Robbie Girling for keeping up my motivation on the tail run. Liz Neary for supportive and active approach - inspiring. Sharon for always warm admin, systems and travel booking support. Dr Kathryn Reardon-Smith and Dr Jarrod Kath for kind and constructive engagement from the beginning of the candidature process. GRS staff including Professor Charlotte Brownlow, Charlie Nelson, Tina Crow and Kathryn Schmacker, whose responsive and understanding support is highly appreciated also.

Outside the research industry, those I am indebted for their patience with me over my research journey. All affected staff at Noosa Council, particularly my bosses, Craig Doolan, Ian Florence, Kerri Coyle, Kim Rawlings, and more recently Camille Oliver who supported me to follow my passion and challenge myself. Marie Wilson for planting a seed for change, and Richie Mc Cormack who passionately believed, motivated and instilled confidence in me to achieve more than I thought possible. Housemates Kylie, Annie, Teghan, Sam, Niesha, Surooj, Lauran, Maram, Tay and Maggie for reality checks and understanding. My long time friends for supportive and motivating check ins throughout, especially Anna, Cliona, Matt, Conor, Evelyn, Kari, Tabitha, Jono, Kara, Colleen, Scienta, Barry, Capper, Greg, Conor, Ben, Kiel, Kate, Cate, The UQ Uni gang Kristie, Haylee, Scott, Ango, Nathe, Ali, Jock, Trin, Alicia, Corinna, Pippi, Danielle for wise insight and my brother SeanO for the laughs. Last but definitely not least, my parents for their cautious excitement about this unfamiliar journey. Also for taking me in during some parts of Covid, excessively feeding me and, along with my dog Benji, reminding me to sleep and exercise.

DEDICATION

All my Grandparents who instilled my love of the ocean & nature.
Dad's infectious passion, and strong desire to take best care of our planet.
Mum for challenging status quo, and inspiring me to apply my mind always.
My entire heart Benji, who endured many late nights sticking by my side.
Paul Randall's passion for vegetation, that led me to appreciate the immense knowledge its growth patterns contain.
The opportunity to combine these influences together in a way that can enhance water quality management approaches.

TABLE OF CONTENTS

LIST OF FIGURES

Note: *All other Figures in Chapters 3-5 (or published papers) are found in the respective chapters and not included in the above*.

ABBREVIATIONS

A=All

AI=Artificial Intelligence

Ammonia=NH_3

ANN = Artificial Neural Network

ANN-PR= Artificial Neural Network Pattern Recognition

ANN-WQ= Artificial Neural Network Water Quality

BF=Baseflows

BFSF=Baseflow and Streamflow

C=Catchment

Category 1=Catchments with similar DIN patterns during increasing flows and rainy season.

Category 2=Catchments with year round similar DIN patterns.

Category 3= Catchments with similar DIN patterns during retreating flows and dry season.

CM=Mary Catchment

CSWQM= Catchment Scale Water Quality Simulation Models

d=Willmotts Index

DDM= Data driven model

DIN = Dissolved Inorganic Nitrogen

Eq=equation

etc.= et cetera

EU=Ecounit

EULUOV=Ecounit, Land Use and Original Vegetation

F=Flows

F1=Category 1 flows (Wet season/increasing flows)

G=Gauged

GBR=Great Barrier Reef

HPC=High Performance Computing

IUCN= International Union for Conservation of Nature

km=kilometres

L=Litre

LU=Land Use

mg=milligrams

ML=Machine Learning

Match=Catchments paired together for their similarities

MSE=Mean Square Error

Nitrate= NO^-_3

Nitrite=NO^-_2

NSE=Nash Sutcliffe Efficiency

obs=observed data

OV=Original Vegetation

OW=Open Woodlands

PDA=Production on Dryland Agriculture

pde=Peak Percentage Deviation

PhD= Doctor of Philosophy

PIA=Production on Irrigated Agriculture

PR=Pattern Recognition

PUB= Predictions in Ungauged Basins

PBM=Process based models

Q1= Top 25% of journals in the field, based on Impact Factor

QWMN=Queensland Water Modelling Network

R^2=regression coefficient

ReLU=Rectified Linear Units

RMSE=Root Mean Square Error

ROC=Receiver Operator Characteristic

SHAP= Shapley Additive exPlanations

SHAP-AD=Shapley Additive Deviations

sim=simulation

SF=Streamflow

UG=Ungauged

UN=United Nations

UniSQ=University of Southern Queensland

UNESCO=United Nations Educational, Scientific and Cultural Organization

USQ=University of Southern Queensland

W=Water

WQ=Water Quality

WT=Wet Tropics

XAI=eXplainable Artificial Intelligence

XAI-SHAP= eXplainable Artificial Intelligence using SHAPley Additive Deviations

CHAPTER 1: INTRODUCTION

1.1. Research context

Australia's Great Barrier Reef (GBR) is the world's largest coral reef system. It evolved in low nutrient waters off Northeast Queensland's tropical coastline (Furnas, 2003). During the late 1900s intensification of anthropogenic land use within catchments that drain to the GBR coincided with symptoms of elevated nutrients including algal and crown of thorns starfish (*Acanthaster planci*) blooms, (Kroon et al., 2012). These symptoms of nutrient imbalances critically threaten the resilience of coral reefs to withstand and recover between major disturbances such as regular tropical cyclones, and bleaching events (Baker, 2003; Furnas, 2003; GBRMPA, 2001). In response, extensive action plans aiming to reduce anthropogenic impacts towards the health and resilience of the World Heritage Listed GBR, were committed to (Anonymous, 2003; State of Queensland, 2011, 2018).

Despite the investments and efforts to reduce nutrients, as of March 2022 a report by United Nations Educational, Scientific and Cultural Organization (UNESCO) and International Union for Conservation of Nature (IUCN) recorded that the GBR met criteria for the "in-danger" list due in part to slow progress towards nominated water quality targets (Brassington et al., 2017; Carter & Thulstrup, 2022; Chen et al., 2011; State of Queensland, 2018; Steven et al., 2019). Progress to water quality targets and prioritisation of actions are measured qualitatively by models and certainty in those models increases with observed data. Given the precarious state of the water quality balance, and impracticality to gain observed data for all ungauged areas, reliable and explainable State of the Art approaches for modelling the ungauged areas are an ethical necessity to justify decisions that impact both humans and the GBR (Baird et al., 2021, Creighton et al. 2021, Di Baldassarre et al., 2019).

Hereafter, lands that contribute runoff to the waters surrounding the coral, lagoon, fore and back reefs of the GBR Marine Park are referred to collectively as the GBR Catchments. Areas that drain each of the 35 rivers whose surface hydrology are separated are referred to as gauged or ungauged catchments. For the purpose of this thesis gauged means catchments which have continuous flow and sporadic water quality data monitored and systematically recorded by the Queensland State Government and sufficient, and ungauged don't. These areas are shown, along with the gauging features, in Figure 1.

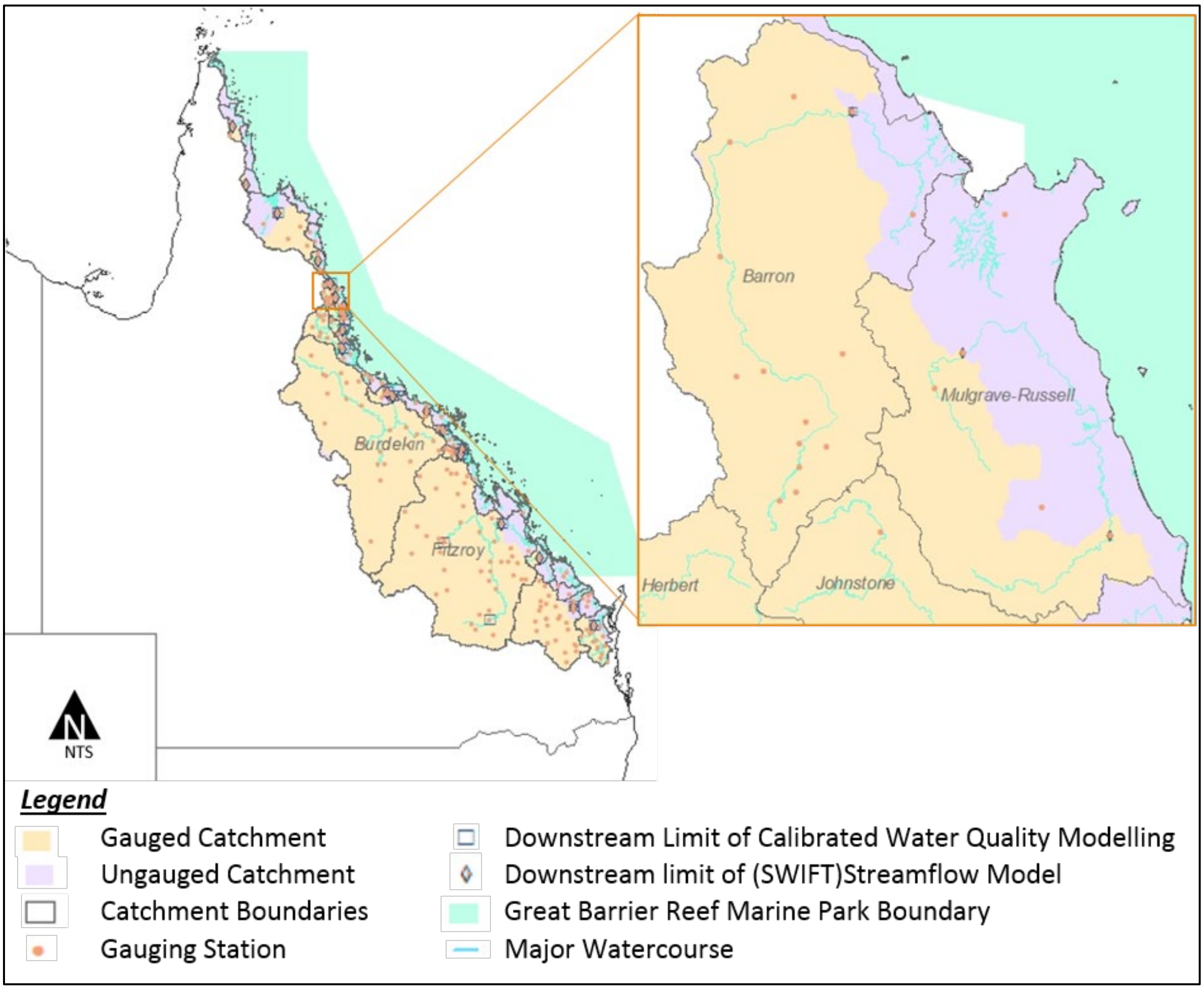


Figure 1.1: Location of Gauging Stations throughout the Great Barrier Reef Drainage Basin. Inserted close up shows the river networks in relation to the gauged and the ungauged catchments.

1.2. Statement of the challenge and knowledge gaps

Nutrient quantification is made possible by Catchment Scale Water Quality Simulation Models (CSWQM), which are the tools that quantify water quality constituents throughout those systems (Baker, 2003; Fu et al., 2019). Catchment Scale Water Quality Simulation Models are developed, calibrated, and verified using observed water flow and observed water quality data collected in the corresponding catchment (Dadson et al., 2019). While flow and water quality data suitable for CSWQM development are collected in gauged catchments, approximately 30% of all flows to the GBR source from ungauged catchments (Khan et al., 2019; Wells et al., 2017). This substantial lack of observed data, limits calibration, validation, and training of catchment specific models in those ungauged/unmonitored circumstances and affects certainty regarding collective impacts of land-based sources of nutrients towards the health and resilience of the reef (Aerts et al. 2023, Bartley et al., 2017, McCloskey & Waters 2017). Such lack of data afflicts CSWQMs worldwide (Do et al. 2018, Kratzert et al., 2019, Kreibich et al., 2023). An entire decade of research called Predictions in Ungauged Basins (PUB) was dedicated to a resolution. The overall conclusion of the PUB decade was that regionalisation, was the best method overcome data shortages in catchment scale hydrological modelling for ungauged areas (Hrachowitz et al., 2013). The recommendation to overcome data deficiencies resonates the seminal principles of catchment modelling also recommended by Nash & Sutcliffe (1970) whereby inclusion of data that quantifies drivers of the constituent being simulated generates the most reliable results. This widely validated approach for modelling ungauged areas underscores the necessity for underlying processes of the constituent being modelled to be understood and demonstrates the need to identify catchments that share similar process drivers.

Catchment classification is the fundamental precursor to regionalisation which matches catchments by similar process drivers and is therefore the foundational topic for this thesis. While catchment classification methods such as nearest neighbour and physical similarity are effective to reflect the empirically linear drivers of flow and total suspended solids, heterogeneous performance results for Dissolved Inorganic Nitrogen (DIN) simulations highlight that process drivers of constituents differ (Hrachowitz et al., 2013; Merz & Blöschl, 2004; Narbondo et al., 2020). Although studies demonstrate DIN patterns are detectible in highly monitored areas (Ebeling et al., 2021, Zhang et al., 2022), the methods are not transferrable to ungauged catchments that drain to the GBR. Drivers of DIN are dynamic and fluctuate depending on a combination of cryptic and other biological influences (Lintern et al., 2018, Lintern et al., 2021, Liu et al., 2021). Exploration of the influence of catchment scale drivers towards heterogeneous nitrogen to date have found that soil water depth, topography, climate as well as underlying geology are principal drivers (Zhi et al., 2020). While some of the driving influences on DIN patterns are known, all studies agree that relationships appear non-linear, heterogeneous, and changeable over space and time (Rafiei et al. 2022). Methods that enable catchments to be matched based on the combination of heterogeneous response to the wide and varying combination of drivers of DIN are necessary to facilitate the increased certainty in existing water quality modelling methods and ensuing sound stakeholder decisions.

Traditional statistical regression methods typically used for classifying the catchment similarity require prior knowledge of the influential drivers towards the constituent of interest (Liu et al., 2021). Additionally, any method intended to classify all catchments that flow to the Great Barrier Reef necessitates for ubiquitous availability of that same data across all respective catchments. However, changeability in relationships between the drivers of DIN and the water quality response over space and time is undefined and inhibits the suitability of statistical regression methods to be transferred to new areas. For these situations, Harris (2012), Parrott (2010), Prinzio et al. (2011) and Toth (2013) all highlighted the need to consider new approaches to evaluate the complex and adaptive interrelationships of the biological and ecological aspects of catchments on the water quality responses. Improved computing power coupled with machine learning skills overcome existing paradigms and can provide new insights to inform classification (Goodwell et al, 2020; Kitchin, 2014). The Artificial Neural Network discipline of machine learning is characterised by automated forward and back propagation which overcomes operator knowledge bias associated with a priori selection of dominant catchment descriptors for classification and has track record in detecting the heterogeneous water quality patterns associated with DIN (Husic et al. 2023). Together these machine learning abilities can enhance the feasibility of interrogating larger datasets and uncover non-linear functions (Merz et al., 2020; Saadi et al., 2019, Tung & Yaseen, 2020).

For this doctoral research project, new methods to classify catchments using ubiquitously available data as a proxy for the similarity of patterns in DIN in ungauged areas are explored. Dissolved Inorganic Nitrogen has been selected due to the key influence it has on algal growth on the reef (Anonymous, 2003; Baker, 2003; Bartley et al., 2017; Kroon et al., 2012). Dissolved Inorganic Nitrogen has different drivers to the other key constituent of Total Suspended Solids, which has similar physical drivers as flow and is heavily researched (Sivakumar et al.,2015, Moliere et al. 2009, Ali et al., 2012). Further, DIN is part of the dynamic nitrogen cycle, which undergoes constant transformation (Gao et al. 2023) and in its Nitrous Oxide form N_2O form is a greenhouse gas with 273 times the equivalent temperature change impact of carbon (Jones et al., 2023). Therefore, this doctoral research thesis, for the first time, has focused on artificial intelligence (AI) and machine learning (ML) methods to classify the catchments for DIN as a priority to maximise the potential impact the research can influence to both water quality as well as climate change influences towards the resilience of the reef.

1.3. Thesis outline

This doctoral thesis contains 6 chapters in total, inclusive of this introduction chapter. The thesis chapters include:

Chapter 2: provides an overview of the importance of Classification towards water quality modelling worldwide, but particularly for Great Barrier Reef Catchments In this chapter water quality model approaches are reviewed and evaluated in relation to their suitability for use on ungauged catchments. Gaps in knowledge as they relate to application of models to land use decision making are identified, and opportunities that exist to fill the gaps are introduced.

Chapter 3: investigates suitability of alternative spatially relevant data sources, that reflect the spatial drivers of DIN in catchments, as a proxy for missing water quality data, necessary to validate water quality models. In this chapter Original Vegetation data is coupled with Land Use data and Artificial Intelligence is used to determine whether catchments that share the same patterns in Original Vegetation and Land Use data also share the same patterns in water quality data such that the spatially relevant data could become a proxy for water quality classification, where the water quality data is lacking.

Chapter 4: builds on the findings of Chapter 3 to explore whether spatial temporal variabilities influence the suitability of certain catchments for classification or water quality datasets for data transfer in water quality models. Spatial data features are also identified to provide explanation for the results.

Chapter 5: Applies the findings of chapter 3 and 4 to classify gauged catchments that flow to the Great Barrier Reef, to the catchments that are ungauged. The suitability of the classification as a proxy for water quality patterns under different dataset partitions is evaluated via a case study to confirm whether the classification method and dataset partitioning explored in Chapters 4 and 5 can be applied to Artificial Intelligence model frameworks to forecast Dissolved Inorganic Nitrogen flowing to the Great Barrier Reef.

Chapter 6: discusses the overall results, potential application, and overall outcomes and conclusion of how this thesis contributes towards new knowledge that can be used to deliver practical outcomes to improve communication material and decision-making tools that influence anthropogenic impacts on the health of planet earth. Further research to address the identified limitations of this research are also recommended.

Appendix A-C: Supplementary material as published for Chapters 3-5.

An Outline of the Doctoral Thesis

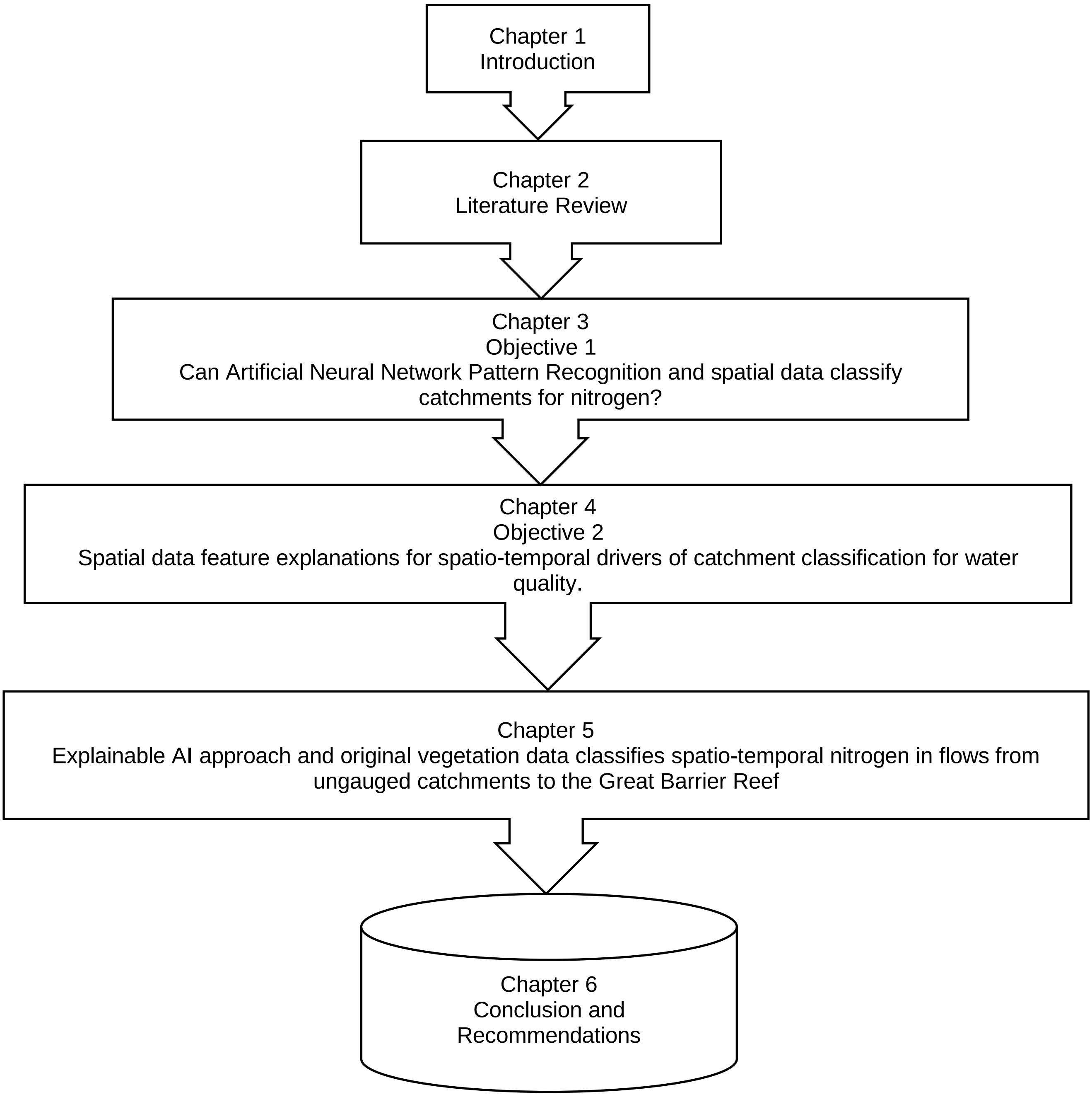


Figure 1.2: Schematic flow chart of various organisational components of this doctoral research project.

CHAPTER 2: LITERATURE REVIEW

2. Foreword

This chapter builds on the background introduced for this doctoral thesis, providing a critical evaluation of knowledge published outside this thesis, for simulating nutrients from ungauged catchments. In building the epistemology of this thesis, this chapter details the current state of knowledge and approaches, knowledge gaps and evaluates opportunities in the emerging paradigms of data evaluation.

An overview of the importance of classification to water quality modelling worldwide is provided in this chapter. In particular, this chapter critically evaluates the current knowledge on water quality simulations in ungauged areas, modelling of Dissolved Inorganic Nitrogen in catchment runoff; the importance of classification for data transfer to ungauged basins flowing to the Great Barrier Reef and elsewhere. Gaps in knowledge as they relate to application of models to land use decision making are identified, and opportunities that exist to fill the gaps are introduced.

2.1. Water quality modelling in ungauged areas: current knowledge

Catchment Scale Water Quality Models are a series of empirical equations developed and extrapolated to the catchment scale to quantify water quality responses to catchment features (Trancoso et al., 2016; Wagener et al., 2007; Di Prinzio et al., 2011). While parsimonious model design refined to dominating drivers is recommended to reduce error propagation (Andréassian et al., 2012; Gazzaz et al., 2015, Nash & Sutcliffe, 1970), the model complexity must be sufficient to answer questions that instigated the model development (Bell et al.,2007; Cole et al.,2006; Yaseen et al., 2018). Observed data is necessary to calibrate these models to overcome errors that can result via extrapolation of empirical principles to the catchment scale, however, suitable calibration data is not available in all areas such as ungauged or unmonitored catchments (Niroula et al. 2023; Sivapalan et al 2003). The International Association of Hydrological Sciences - Predictions in Ungauged Basins – “PUB Decade” final summary paper recommended regionalisation of model parameters as the most appropriate alternative method to overcome data shortages (Hrachowitz et al., 2013).

Classification is a necessary precursor to regionalisation because it identifies the most similar catchments to transfer parameter data between (Hrachowitz et al., 2013; Merz & Blöschl, 2004; Goodarzi & Navardi, 2019). Catchments are classified based on dominant physical catchment processes, or nearest neighbour (Ayana et al., 2015; DeLancey et al., 2020; Hrachowitz et al., 2013; Ebeling et al. 2021), and this facilitates data sharing between the most similar catchments in CSWQMs. Empirical relationships that exist between catchment drivers and water balance response is attributed to the success of classification underpinning regionalisation for simulating flow and suspended solids for ungauged catchments (Hrachowitz et al., 2013; Merz & Blöschl, 2004; Narbondo et al., 2020). However, it becomes limited in overly parsimonious models. Inductive and deductive classification approaches overcome limitations associated with parsimonious empirical model architecture by using deductive approaches of spatial data as a proxy to representing the full catchment response (Olden et al., 2012). While effective, these studies are focussed on catchment response to abiotic drivers of physical features relevant to flow.

Variability in nutrient patterns is observed across different catchments of the Great Barrier Reef (Liu et al. 2021) and this is reflected in heterogeneity of nutrient simulation results where existing classification approaches are applied (Swain et al. 2019, Merz et al., 2020; Sivapalan, 2018). Notable amounts of DIN in receiving waters of the Great Barrier Reef catchments are consistent with water quality simulation results for gauged areas, however the drivers are not fully quantified to effectively inform classification of gauged to ungauged areas (Cheng et al., 2018; Soltani-Gerdefaramarzi et al., 2021; Kroon et al., 2012; Snelder et al., 2018; Khan et al., 2020). It is well established that modelling capabilities can be improved by inclusion of unique catchment characteristics within model architecture as well as classification that reflect the modelling purpose (Merz et al., 2020, Saadi et al., 2019). Techniques to classify catchments based on the catchment response to the combination of nutrient processes drivers is limited, overlooks the fact nutrient cycle includes a biotic response component, and therefore an opportunity to explore for improving nutrient modelling of ungauged catchments, particularly for Dissolved Inorganic Nitrogen for the Great Barrier Reef catchments.

2.2. Classification for Dissolved Inorganic Nitrogen

There is limited research regarding hydrological catchment classification for nutrient responses (Giesbrecht 2022), and inconsistent performance of simulations for nutrients is found where existing classification approaches designed for flow are applied (Buzacott et al, 2019, Liu et al., 2018). A key difference between the drivers of nutrients vs flows and their response is the non-linearity as well as lack of empirical relationships for nutrient responses (Liu et al., 2021). While parsimonious models are criticized for reducing system knowledge and catchment heterogeneity within the ungauged areas, it is the drivers of nutrients are that are also overlooked (Hallouin et al., 2020; Waterhouse et al., 2017).

Land use is reported to have the greatest influence on Dissolved Inorganic Nitrogen inputs towards the Great Barrier Reef, and land use adjustments are a primary target for management of anthropogenic impacts on water quality (Fu et al., 2019; Liu et al., 2023). Consequently, within architecture of water quality simulation models, nitrogen constituents are most influenced by land use (Alnahit et al., 2022; Liu et al. 2021; Li et al., 2023). Likewise, regardless of the benefit of using Land Use data as a proxy for DIN similarities across catchments, research continues to find heterogeneity across areas with identical land uses, despite each Land Use type parameterised with homogenous contribution (Park & Lee, 2020). Although both biological and physical catchment attributes influence nutrient processes, biological influences have not specifically been included in classification and influence towards and the driver of the observed inconsistencies is overlooked (Merz et al., 2020; Turak et al., 2017). Inclusion of biotic drivers into classification approaches is not straightforward, however, because the spatial representation of the combined nutrient drivers across the Great Barrier Reef catchments are ambiguous, and not established (Liu et al., 2018).

Classification methods that consider variable drivers of water quality separately, can overlook the combined influence of catchment features in the ungauged system that infield measurements, used for parameter calibration of models for gauged catchments, otherwise capture (Buzacott et al., 2019; Booker and

Woods, 2014; Kuentz et al., 2017; Teutschbein et al., 2018; Oudin et al., 2010). Newall & Tiller (2002) introduced the concept of bio-regionalisation proposing that the biological response can indicate the productivity, and hence nutrient cycling in an area. While Zhang et al. (2022) has demonstrated that appropriate spatial data can be used to classify and explain the drivers of Nitrogen flowing to Chesapeake Bay this is limited to gauged catchments only and using datasets that are not applicable to the Great Barrier Reef catchment area. A spatial dataset suited as a proxy for classification of nitrogen and therefore transferrable to ungauged catchments in the Great Barrier Reef catchments is the fundamental gap in the literature to classify catchments more appropriately for the purpose of modelling nutrients flowing from ungauged areas.

2.3. Data availability for inductive vs deductive classification of DIN

Inductive classification is classification informed by observed data, regarding that data, while deductive classification is classification informed by an alternate data source, to deduce the same conclusions. Where observed data is missing in some areas, deductive classification methods use spatial data patterns known to be a proxy to identify catchments that theoretically share corresponding catchment responses (Olden et al. 2012). For catchments draining to the Great Barrier Reef, spatial data of Land Use are available as static datasets across all areas of the Great Barrier Reef catchments (ABARES, 2016). The ubiquitous availability of this data makes it suitable for deductive classification where water quality observations are missing. However, due to the existing heterogeneity observed in DIN patterns from areas with similar Land Use patterns (Liu et al., 2021), an alternative parsimonious dataset is required to capture the heterogeneity and increase trust in the classification method that informs data transfer to ungauged areas. In addition to land use, soil water depth, topography, climate as well as underlying geology are other known drivers toward nitrogen releases in water across landscapes (Lintern et al., 2018; Zhi et al., 2020). While these features are considered as input variables within the CSWQMs, classification methods were not found in the literature that represent the combined catchment response to those same nutrient drivers.

Mapping available for each nutrient driver exists at varying scales, is considered in isolation of each other, and multiple mapping layers reduce parsimony and increase complexity (Dadson et al., 2019; Fu et al., 2019). In contrast, the original evolutionary response of natural vegetation growth to the varied combination of landscape features, that are also the above-mentioned nitrogen drivers, are represented in the Queensland State Government open access Broad Vegetation Group Mapping. The heterogeneity in vegetation communities shown on these maps are informed by differing combinations of bioregion, geology, aspect topography etc, and therefore parsimoniously capture the required information ubiquitously across all gauged and ungauged catchments that flow to the Great Barrier Reef (Neldner et al., 2017). Jay & Neumann (2021) found the vegetation mapping reveals site quality that is influential towards vegetation productivity and therefore nutrient and water demands across large landscapes. This productivity is also influenced by biotic process drivers such as aspect, geology, natural water balances, radiation, elevation, soil biota etc even after the original vegetation has been removed. For this reason, the open access Broad Vegetation Group Mapping, referred to herein as Original Vegetation mapping, has been identified as an alternative dataset with merit to explore for its suitability as a proxy spatial dataset for deductive classification approaches for DIN.

2.4. Data Driven vs. Process Based models

Different model abilities suit different applications, yet the trade of parsimony vs parameter detail render current applications of CSWQMs unsuitable for predictions and forecasting of ungauged catchments (Dadson et al., 2019; Fu et al., 2019). CSWQMs may be either processes based, or data driven. Process based models (PBMs) virtually represent key drivers of hydrological systems, so are the preferred simulation tool where data are scarce (Fatichi et al., 2016). They use generic catchment attributes coupled with fundamental mathematical formulae established from hydrological principles tested elsewhere, i.e. laboratories, to establish the models parameter inputs. Process Based Models then require a posterior calibration with field data (i.e., as initial and boundary conditions) to ensure parameterisation reflects the overall combination of system processes. In the presence of calibration data, this architecture facilitates effective estimates of water quality constituents in runoff (Dadson et al., 2019; Salas et al., 2014). Where data is unavailable, parameters from the classified catchment can be transferred to facilitate refinement of the model from fundamental to empirical catchment responses (Merz and Blöschl, 2004; Narbondo et al., 2020; Pagliero et al., 2019). However suitable catchment classification approaches that represent the drivers of the constituent are necessary for trust in the simulation outputs (Nash & Sutcliffe et al., 1970).

Despite ability of PBMs to operate in data scarce situations, their computational requirements, fidelity, and uncertainty drawn from estimates of initial conditions increase as the number of processes increase. For these reasons, well trained parsimonious data driven models are better suited for large scale hydrological modelling, compared to process-based models (Dupas et al., 2013; Fatichi et al., 2016). Data driven models require catchment specific data a priori to operate, so are not suited to modelling of ungauged catchments which lack observational flow and water quality data (Dadson et al., 2019).

While some process-based models have achieved satisfactory simulations of DIN in receiving waters, these are established via calibration meaning that consistent relationships between water quality patterns and the catchment drivers are not empirical, reducing certainty in transfer of parameter data to inappropriately classified ungauged areas (McCloskey et al., 2021, Zhang et al., 2022). Previous statistical approaches to detect heterogeneous DIN patterns have been limited by linearity, need for prior knowledge of relationships in the data and inability to extend methods to unmonitored areas (Huang et al., 2019; Khan et al., 2020; Lintern et al., 2018; Liu et al., 2021; Snelder et al., 2018). The absence of linearity on the DIN datasets has also limited the ability to establish relationships between DIN in the receiving water quality and catchment drivers meaning that classification of ungauged to gauged areas using spatial data as a proxy for DIN patterns has not been possible.

2.5. Machine Learning for classification purposes

Improved computing power and machine learning skills have been explored for their potential to enhance selection of system influences (Lu et al., 2023, Singh et al., 2019) but these benefits have not previously been explored for informing classification for purpose of catchment drivers of DIN in the GBR. Machine learning enhances the feasibility of interrogating much larger datasets, uncovers non-linear functions for parameter calibration and can therefore overcome existing paradigms (Goodwell et al., 2020; Kitchin 2014). This is demonstrated for flow and suspended solids simulation in large ungauged catchments across France, Germany and Canada (Merz et al., 2020; Saadi et al., 2019). While machine learning requires a priori data to operate, it is not limited to static co-efficient weightings of traditional regression and other data driven model (DDM) methods. Forward and back propagation design of machine learning provides greater flexibility of data input sources compared to both DDMs and PBMs. Additionally, machine learning delivers superior pattern recognition abilities over traditional DDMs (Tyralis et al., 2019; Worland et al., 2018; Yaseen et al., 2019). While pattern recognition abilities of machine learning is useful to identify catchments that share similar dataset patterns, the black box nature affects transparency and ability to scrutinise model results necessary for interrogability and trust in results (DeLancey et al., 2020).

Dadson et al. (2019) found different process based, data driven and machine learning modelling strategies are complementary rather than mutually exclusive, i.e. they serve and help the other. Where machine learning can enhance relationship mining ability, process-based models maintain the ability to interrogate the system. A hybridised approach to couple CSWQM process-based principles with relationship mining abilities of machine learning can therefore offer a solution to overcome the data deficiencies, however, prior to commencement of this research, remained unexplored for nutrient simulations for ungauged catchments (Dupas et al., 2013; Merz et al., 2020; Oehler & Elliott, 2011; Sharifi et al., 2017; Valizadeh et al., 2017).

2.5.1. Artificial Neural Network Models for classification purposes

Artificial Neural Network (ANN) is a type of machine learning method that has demonstrated track record for classification based on shared patterns, as well as effectively simulating nitrogen patterns in monitored catchments throughout the world (Ighalo et al., 2021; Khalil et al., 2011; Tung & Yaseen, 2020). Although ANN has demonstrated success in calculating instream nutrient concentrations using other instream water constituents as predictors, prior to the doctoral research, there was no demonstrated success in its application to calculating instream nutrients in areas lacking data, such as ungauged areas (DeLancey et al., 2020; Hameed et al., 2017; Jin et al., 2019; Khalil & Adamowski, 2014; Khalil et al., 2019; Tabari & Talaee, 2015; Tao et al., 2019). The reason is consistent, long term, accessible instream data of any kind, which is necessary for the development of the ANN models, does not exist in the ungauged areas. Further, research demonstrates that the relationships between DIN and its drivers are not consistent in all catchments (Liu et al., 2018), which may explain the inability to transfer ANN models directly to any ungauged catchment. Process based CSWQM principles demonstrate that transfer of data from gauged catchments with most similar drivers of the constituents being measured can be effective, where the classified catchments share similar drivers. However, apart from Zhang et al. (2022) demonstrating the benefits of machine learning techniques to find relationships between catchment response and some catchment features and prior to this doctoral research, the application of spatial data proxies for simulating DIN, and impact on model performance has not been explored. Exploitation of the expanded pattern recognition and classification abilities of ANN is an unexplored opportunity to establish relationships between spatial data proxies expected to be an indicator of DIN patterns, and the observed catchment response.

2.5.2. Explainable Artificial Intelligence-based classification models

While the forward and back propagation of Artificial Neural Networks have potential to overcome limitations of existing classification methods, benefits have been met with caution due to lack of explainability of results (Nearing et al., 2021; Tocchetti & Brambilla, 2022). The ability to explain the decisions influenced by water quality modelling results is an ethical expectation of stakeholders and is the reason for the dominance of process based and traditional data driven approaches in CSWQMs (Arrieta et al., 2020). Since completion of the PUB decade, computing powers and further research on explaining the results of Artificial Intelligence are developing and therefore is prudent to explore for the benefit of simulating DIN in ungauged areas.

Explainable Artificial Intelligence, known as XAI is a discipline of various approaches to evaluate and identify the likely drivers behind results generated from otherwise black box model architecture of AI (Vilone & Longo 2021). This XAI ability therefore can facilitate for the drivers of classification to be identified and support inclusion of AI within CSWQM architecture. Shapley game theory approaches provide explainability and transparency to human end users of artificial intelligence model outputs and further builds trust when used for decision-making (Lundberg et al., 2020) and referred to herein as XAI-SHAP. Unlike hydrological classification studies which consider each variable having equal weighting (Jehn et al., 2020), game theory recognises up front that each variable in datasets used for classification is influenced by the additive influence of all variables with each other (Cohen et al., 2007; Lundberg and Lee, 2017). Contributions of each feature are considered individually to overall system outcomes, and synonymous with biotic interactions, each feature variable is uniquely influenced by the presence of the other features in the dataset (Arrieta et al., 2020; Lundberg and Lee, 2017; Wang et al., 2022). To date the driver's influencing similarities in nutrient responses in Great Barrier Reef catchments remains unexplained (Liu et al. 2021), and therefore XAI-SHAP provides opportunity for insight to new understandings of data relationships exposed by the ANN classification method.

2.6. Summary of knowledge gaps and opportunities for research

In CSWQMs for ungauged areas, classification is an effective technique to identify the most suitable catchments to transfer data and overcome data shortages. While the method is effective for flow and total suspended solids which are driven by abiotic drivers only, the same catchment classifiers are ineffective for CSWQMs which are designed to simulate nutrients. Nutrients are driven by more dynamic abiotic influences, however methods to classify catchments for response to these drivers using methods transferrable to ungauged areas is not yet established.

Current classification approaches consider physical influences separately and in isolation, so do not account for the variability of influences towards nutrient responses observed across different catchments of the Great Barrier Reef. Nutrient processes are driven by a combination of biotic interactions that vary in response over space and time depending on the full combination of all environmental influences. Unlike physical drivers, the full suite of nutrient drivers, which include biotic drivers, are not additive, or specifically captured in one mapping dataset and therefore inhibits the ability for catchment feature response relationships to be established within parsimonious model architecture. Techniques to classify catchments based on the combination of nutrient processes drivers is a fundamental gap as well as opportunity to explore for improving nutrient modelling of ungauged catchments. Datasets that are available for all gauged and ungauged areas and reflect the combined influence towards nitrogen outputs are needed. Heterogeneity in original vegetation communities are informed by differing combinations of bioregion, geology, aspect topography etc, and therefore parsimoniously capture the required information ubiquitously across all gauged and ungauged catchments that flow to the Great Barrier Reef.

Non-linearity of nutrient responses has previously inhibited the evaluation of relationships between inductive catchment responses and ability to deduce the contributing spatial data related to catchment drivers. Forward and back propagation design of machine learning provides greater flexibility of data input sources and delivers superior pattern recognition to overcome non-linearity in datasets. Exploitation of the expanded pattern recognition and classification abilities of ANN provides opportunity to establish relationships between spatial data proxies expected to be an indicator of DIN patterns, and the observed catchment response. Although ANN establishes relationships within a black box architecture, XAI provides opportunity for insight to new understandings of data relationships exposed by the ANN classification method to better evaluate the suitability of catchments classified using ANN informed approaches.

2.7. Research Hypothesis

The three hypothesises of this PhD thesis are based on prior knowledge that:

a) Classification is fundamental to CSWQMs because it informs the most suitable catchment to transfer data between where one is lacking, but current classification methods result in inconsistent results for dissolved nutrient simulations;

b) Consistent satisfactory performance is achieved for hydrological model simulations where data transfer is from catchments classified based on process drivers of the water constituent under investigation. The method is effective for linear relationships, but drivers of dissolved nutrient responses are ill poised and non-linear;

c) Forward and back propagation abilities of machine learning algorithms can facilitate evaluation of non-linear relationships in datasets, are relatively under-explored for classification of catchments for nutrients.

d) Exploiting the unique abilities of Artificial Intelligence to identify and classify catchments with similar drivers has potential to enhance certainty in data transfer for CSWQMs.

Based on this knowledge, the overarching hypotheses guiding this research are:

1. Inductive classification of catchments for flow pattern similarities differs from inductive classification for DIN pattern similarities, and spatially relevant data can deductively classify the same catchments together for DIN patterns. Spatially relevant data means data ubiquitously available across both gauged and ungauged catchments in the study area, and contains knowledge of the combined influence of all known drivers towards DIN.
2. Heterogeneity in DIN patterns in water quality discharges can be explained by biological knowledge contained in the appropriate spatial data.
3. Spatially relevant data identified as a suitable proxy for DIN driver classification, enables all ungauged catchments to be classified to gauged catchments, and can inform improved DIN simulation performance in a pseudo-ungauged environment.

2.8. Research Aim, Objectives, and Outreach Actions

The aim of this PhD project has been attained by a publication's pathway.

The purpose of this research is therefore to develop and share new and applied knowledge exposed by artificial intelligence and machine learning regarding classification of ungauged catchments for DIN. This overall aim of the doctoral research project, to develop and deliver stakeholder-relevant knowledge, is achieved by means of three high quality *Q*1-ranked publications, prefaced by a publication ready literature review section. Together, the resulting publications aim to develop the case that spatial data can be an appropriate a proxy for DIN water quality responses and can inform classification of ungauged catchments to gauged equivalents where water quality data is lacking.

This aim is delivered by a presentation of the following research objectives as publishable papers.

1. Develop a novel method to apply Artificial Neural Network Pattern Recognition (ANN-PR) approach, to evaluate the suitability of nominated spatial data as a proxy for drivers of DIN in flows to the Great Barrier Reef. *The research article has been published in **Science of the Total Environment** (Volume 809, Article 151139).*

2. Uncover new knowledge regarding the interactions of patterns in the water quality and spatial datasets for the study region by evaluating variability in DIN patterns for different season and the flow regimes. Also adapt explainable artificial intelligence-Shapley (XAI-SHAP) principle to add transparency by adding physical understandings of the proposed artificial intelligence models to explain the trends. *The research article has been published in **Science of the Total Environment** (Volume 861, Article 160240).*

3. Provide a demonstration of the practical application of the newly developed ANN-PR and XAI methods to classify all ungauged catchments that flow into the Great Barrier Reef to the Gauged catchments and validate the novel method

using a simulation case study in a pseudo-ungauged catchment. *The research article has been published in Nature Portfolio's journal* ***Scientific Reports*** *Volume 13, Article 18145 (2023)*. *The article was lodged and accepted by the editors of special collection – "New Advances in Ecological Modelling".*

In fulfilment of these research objectives under the doctoral research program, a copy of the three (Q1) papers will form the primary body of the *PhD Thesis by Publications*. The workflow for delivery of these objectives is shown in Figure 2.1 below.

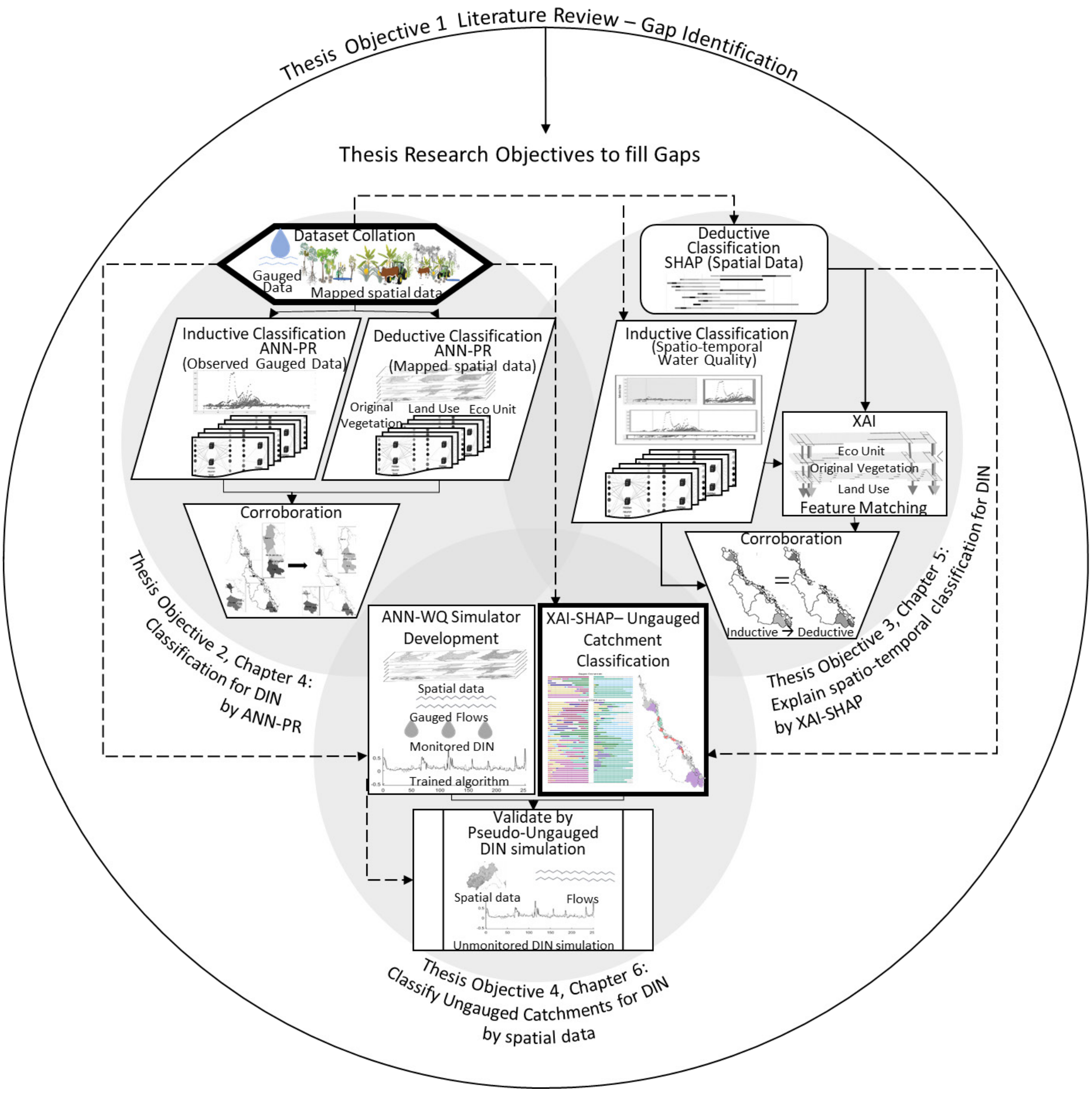


Figure 2.1: Workflow for classification of ungauged catchments using spatial data as a proxy for DIN.

CHAPTER 3: JOURNAL PAPER 1 – What Drives DIN Patterns?

3. Foreword

This chapter presents an exact copy of the published article in ***Science of the Total Environment*** (Volume 809, Article 151139 (2022)). This journal was selected due to the relevancy of the scope that integrates all elements of the natural and anthropogenic environment with the hydrosphere. The title of the published research paper is:

Classification of catchments for nitrogen using Artificial Neural Network Pattern Recognition and spatial data

This research piece establishes the foundations for the classification approach developed throughout this doctoral thesis, and specifically evaluates Hypothesis 1 of this doctoral thesis which is as follows:

Catchments classified using spatially relevant data can classify the same catchments together for DIN patterns in water quality discharges. Spatially relevant data means data ubiquitously available across both gauged and ungauged catchments in the study area and contains knowledge of the combined influence of all known drivers towards DIN.

A set of Artificial Neural Network pattern recognition (ANN-PR) techniques were developed and applied for the first time to explore the ability for spatial datasets to be a proxy classification tool for catchments that share matching DIN patterns. In particular, the spatial datasets evaluated were Original Vegetation (referred to in the paper as Original Ecosystem), Land Use, and then Ecounits, which combine the patterns in both Original Vegetation and Land Use. ANN-PR was used to match catchments together that share the most similar spatial patterns, then repeated to identify catchments that share the most similar DIN with flow patterns. A control dataset was established to match catchments that share the most flow records with no DIN data included. The number of DIN and flow, or flow only records ANN-PR matched to each catchment was then recorded, and the catchment with most records for each dataset deemed inductively classified. This result was compared to the catchment ANN-PR matched together for spatial data, referred to as deductively classified. Results were then evaluated using Kruskal Wallis test for independence (p>0.05) and found that classification scores for Flow datasets are independent of the spatial data classification scores (p=0.09), whereas DIN datasets are not (p=0.01-0.02). The lack of independence for spatial data classification from DIN dataset classification validated the merit of the spatial data used as a proxy for classification of catchments that share DIN drivers. New findings of this research are that classification using Original Vegetation spatial datasets classified additional catchments that Land Use data was unable to classify.

3.1. Graphical abstract

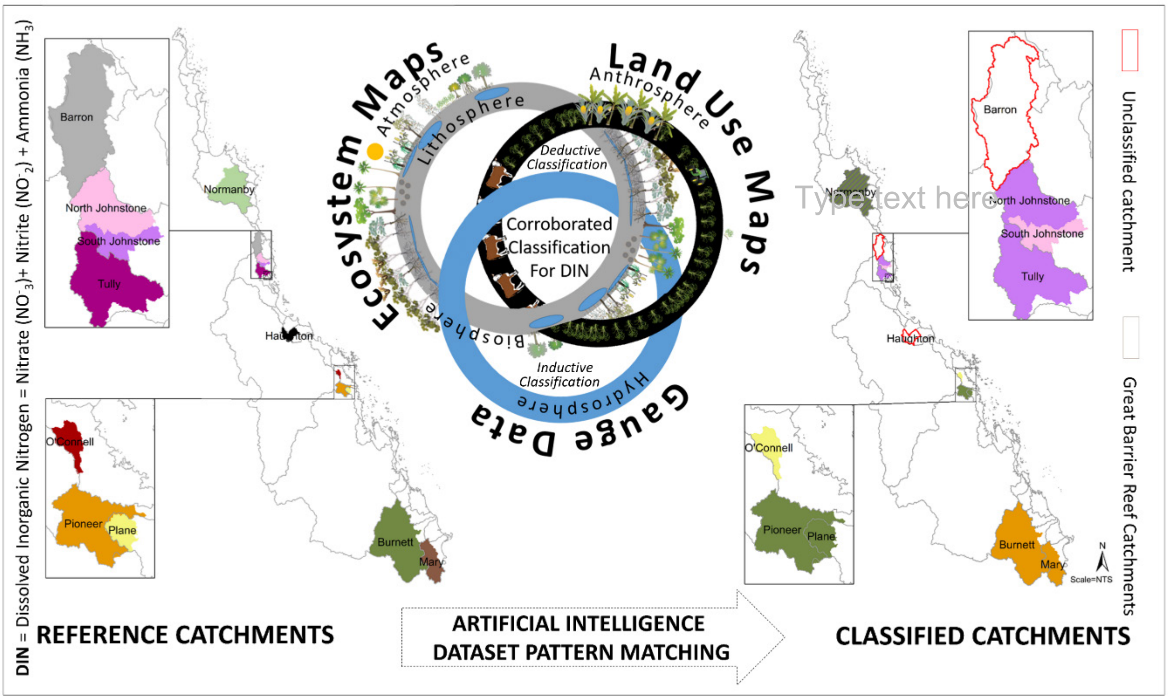


Figure 3.1 Graphical Abstract for Objective 1 – Classifying Catchments for DIN

3.2. Published journal paper

The paper published for this chapter is provided below. The supplementary material in support of the paper is provided in Appendix A.

3.3. Links and implications

This paper forms the foundation of the doctoral thesis by demonstrating that the ANN-PR model coupled with Land Use and Original Vegetation (referred to in the paper as Original Ecosystem) spatial data is an efficient tool to classify the catchments together for DIN. The published results affirm both parts of Hypothesis 1 that:

a) *Inductive classification of catchments for flow pattern similarities differ from inductive classification for DIN pattern similarities; and*
b) *spatially relevant data can deductively classify the same catchments together for DIN patterns.*

The finding that inductive classification for DIN differs from inductive classification for flow demonstrates that classifying catchments based on similarities in drivers of flows are not appropriate for classifying catchments for similarities in the drivers of DIN. Overall, four general DIN pattern groups were observed for water quality data. Interestingly, the results also identified that while the majority of water quality records in each dataset were classified to one catchment, there were a notable number of records matched to the water quality patterns for an alternative catchment. This heterogeneity suggested that classification to only one other catchment at all times may not be appropriate and suggests drivers of DIN may vary.

Notably for advancement of this thesis, the results affirmed that corroboration exists between inductive classification using observed water quality data, with deductive classification using Land use and Original Vegetation spatial datasets. The observations confirmed the hypothesis that ANN-PR approach coupled with spatial data has merit for classifying catchment for drivers of DIN. Nevertheless, application of the method for classifying ungauged catchments for purpose of data transfer in water quality simulation models requires understanding of the drivers of the heterogeneity that could affect model performance. Furthermore, identification of the spatial dataset variables common to the four groups identified for similar DIN patterns was not revealed by the ANN-PR method in isolation and is needed to explain the drivers of the four different DIN patterns. While this chapter identified that ANN-PR

Original Vegetation and Land Use data have merit for classifying catchments for the DIN pattern similarities, the drivers of the classification, and variability in DIN patterns are not known. The next chapter will investigate potential sources of variability observed in the DIN patterns and trial methods to find knowledge in the spatial data as it corroborates with classification, and variability on the DIN patterns.

CHAPTER 4: JOURNAL PAPER 2 – Why is there Heterogeneity in DIN patterns?

4. Foreword

This chapter presents an exact copy of the published article in ***Science of the Total Environment*** journal (Volume 861, Article 160240 (2023)). As with paper 1 reported in Chapter 3, and for consistency, this journal was selected due to the relevancy of the scopes integrated consideration of all elements of the natural and anthropogenic environment with the hydrosphere. The title of the published research paper is:

Pattern recognition describing spatio-temporal drivers of catchment classification for water quality.

This research piece builds of the findings of Chapter 3, by refining the classification method in a way to isolate the cause of heterogeneity in DIN patterns. In particular, this chapter is designed to specifically validate Hypothesis 2 of this Doctoral thesis which is:

Heterogeneity in DIN patterns can be explained by biological knowledge contained in the appropriate spatial data.

This research chapter contributed two additional novel approaches to the classification technique. Firstly, this research split the hydrograph by seasonality and flow regime to evaluate classification responses to these drivers. The classification approach developed in Objective 1 (Chapter 3) was then repeated on all split hydrograph datasets. Secondly, this research piece developed an eXplainable Artificial Intelligence technique using the Shapley approach. This XAI-SHAP approach identified variables with greatest deviation from all other variables in the spatial datasets, and therefore most likely influential in the ANN-PR classification results. Brief literature reviews were undertaken to provide a process-based explanation for results.

The research in this chapter that deviated variables in the Original Vegetation datasets corroborate with DIN classification for various aspects of the hydrograph, and validated Hypothesis 2 via provision of a biological explanation for the relationship. In particular, the doctoral research in this chapter found that catchments classified based on DIN observed in the dry season or below average flows also corroborated with catchment classifications explained by Open Forest vegetation types. Catchments classified on DIN patterns in the Increasing Flows or Wet Season hydrograph corroborated with catchment classifications explained by Open Woodlands, or Woodlands respectively. Finally, catchments that consistently classified together for DIN patterns in all hydrograph combinations corroborated with spatial data catchment classifications explained by greatest deviation of vine forests.

The novel findings of this piece of research are the first time, to our knowledge, that the drivers of heterogeneity in observed DIN for areas with the same Land Use have been explained.

4.1. Graphical abstract

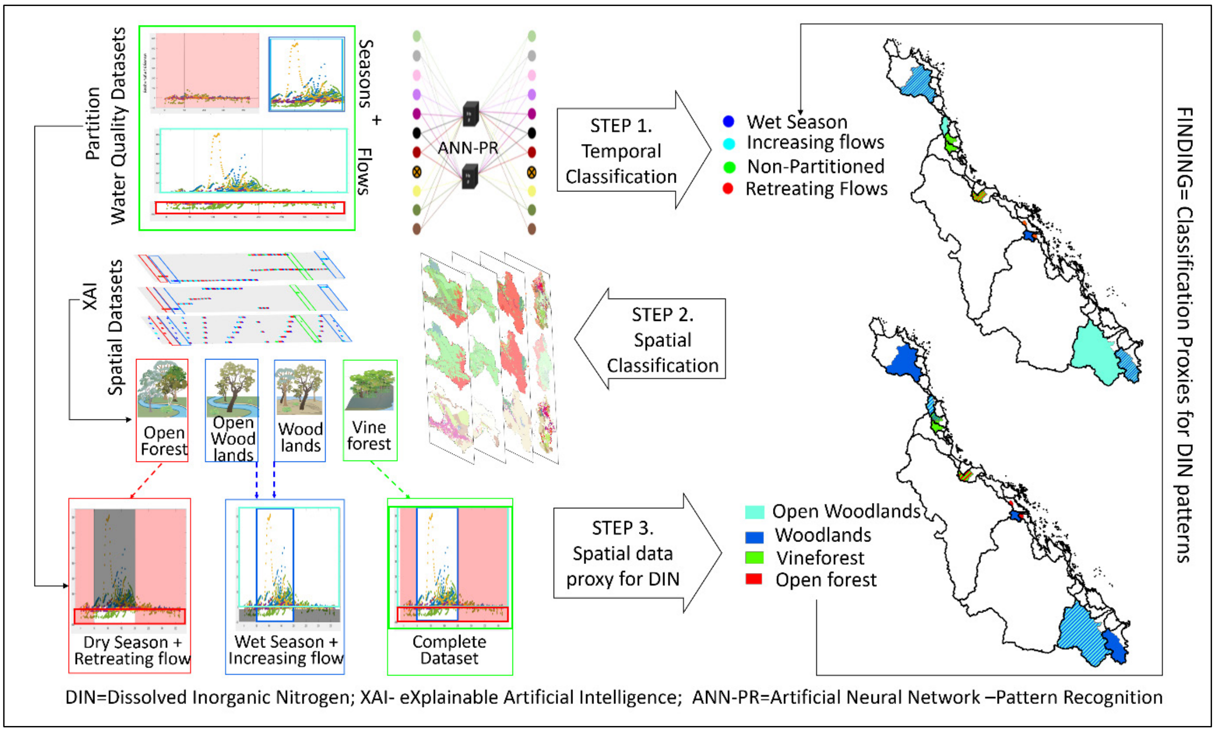


Figure 4.1 Graphical Abstract for Objective 2 – Drivers of DIN pattern heterogeneity

4.2. Published journal paper

The paper published for this chapter is provided below. The supplementary material in support of the paper is provided in Appendix B.

4.3. Links and implications

This chapter (Objective 2, Paper 2, Chapter 4) built on the findings published in Objective 1 (Paper 1, Chapter 3) to provide insight to the drivers of heterogeneity in DIN patterns, relate the drivers of heterogeneity to spatial data variables, and therefore justify why Original Vegetation spatial data can be a proxy for catchment classification for DIN. Observed datasets for DIN patterns were split based on spatio-temporal drivers that isolate DIN records to flow and season drivers, then XAI-SHAP identified spatial data variables that corroborate for inductively classified catchments. Together, results generated by the ANN-PR and XAI-SHAP methods affirmed hypothesis 2 that:

> *Heterogeneity in DIN patterns in water quality discharges can be explained by biological knowledge contained in the appropriate spatial data.*

The results in this chapter provide an explanation for the heterogeneity observed in inductive classification results in Objective 1. This Objective 2 research found for some catchments, inductive classification only applied for DIN records collected on certain season or flow regime. These catchments are consistent with the four DIN pattern groups identified in Objective 1. The XAI-SHAP method developed for paper 2 enables deviated variables in the spatial datasets for each catchment to be identified as a proxy driver for inductive classification. Ecological knowledge of the variables as they relate to heterogeneity in classification patterns then facilitate communication of the spatial data as a proxy for the DIN classification. In summary, findings of Objective 2 suggest that data transfer in water quality simulation models may only be suitable during specific flow or seasons, and Original Vegetation spatial data can be a proxy indicator for the suitable flow or season data to transfer and catchments to classify. Such evaluation of the spatial datasets using XAI-SHAP allows for the ANN-PR inductive catchment classification using spatio-temporal datasets to be logically communicated. This chapter has identified the variables in the Original Vegetation dataset suitable as a proxy to classify catchments to the four differing DIN patterns flowing from the GBR catchments. The next chapter can now apply the findings of this research to classify ungauged catchments to gauged ones using Original Vegetation as a proxy for DIN patterns.

CHAPTER 5: JOURNAL PAPER 3 – Classifying Ungauged Catchments that flow to the GBR

5. Foreword

This chapter presents an exact copy of the article in ***Scientific Reports*** journal. The journal was selected due to the relevancy of this research to the advertised special collection ***New Advances in Ecological Modelling***. The title of the research paper in review is:

Explainable AI approach and Original Vegetation data classifies spatio-temporal nitrogen in flows from ungauged catchments to the Great Barrier Reef

This penultimate research piece applies the findings of ANN-PR and XAI-SHAP techniques developed in journal paper 1 and journal paper 2 to classify, for the first time, ungauged catchments that drain to the Great Barrier Reef, to the gauged ones for the purpose of developing models for DIN simulations. Journal paper 3 forms objective 3 of this doctoral thesis research and is designed to specifically validate Hypothesis 3 which is:

Spatially relevant data identified as a suitable proxy for DIN drivers of classification, enables all ungauged catchments to be classified to gauged catchments, and can inform improved DIN simulation performance in a pseudo-ungauged environment.

Spatial data is ubiquitously available across all Great Barrier Reef catchments and therefore can inform the classification method to classify all catchments to gauged ones. As a continuation of the findings and methods developed in Journal paper 2/Objective 2, the XAI-SHAP method matched ungauged to gauged catchments based on similarity of variable deviations of Original Vegetation datasets. The approach exposed catchments that did not share similar Original Vegetation data variable deviations with the gauged catchments. This was an important finding because the corroboration of classification deduced from spatial data with classification induced by

DIN data was limited to gauged catchments only, therefore classification for informing data transfer for DIN simulation purposes may only be limited to catchments that share the same spatial data deviations as the gauged catchments. The ability of the XAI-SHAP to identify those catchments that do and don't share similar spatial data deviations, and therefore suitability of ungauged catchments as a data recipient, allow for the potential limitations of using ANN-PR classification to be identified.

The suitability of using spatial data as a proxy for DIN classification was validated via development of an Artificial Neural Network Water Quality (ANN-WQ) Simulator. In addition to its validation role, the ANN-WQ simulator unexpectedly made a novel contribution to the knowledge of water quality patterns flowing to the Great Barrier Reef.

Development of the ANN-WQ simulator found, for the first time, that the performance of algorithms, trained to simulate water quality, change depending on the combination of catchment data included in the training datasets. While journal paper 2/Objective 2 identified three categories of DIN patterns in flows to the Great Barrier Reef, trial and error development of the ANN-WQ simulator in this journal paper 3 discovered segregation of training data based on these same three categories affect the ANN-WQ simulation performance.

5.1. Workflow

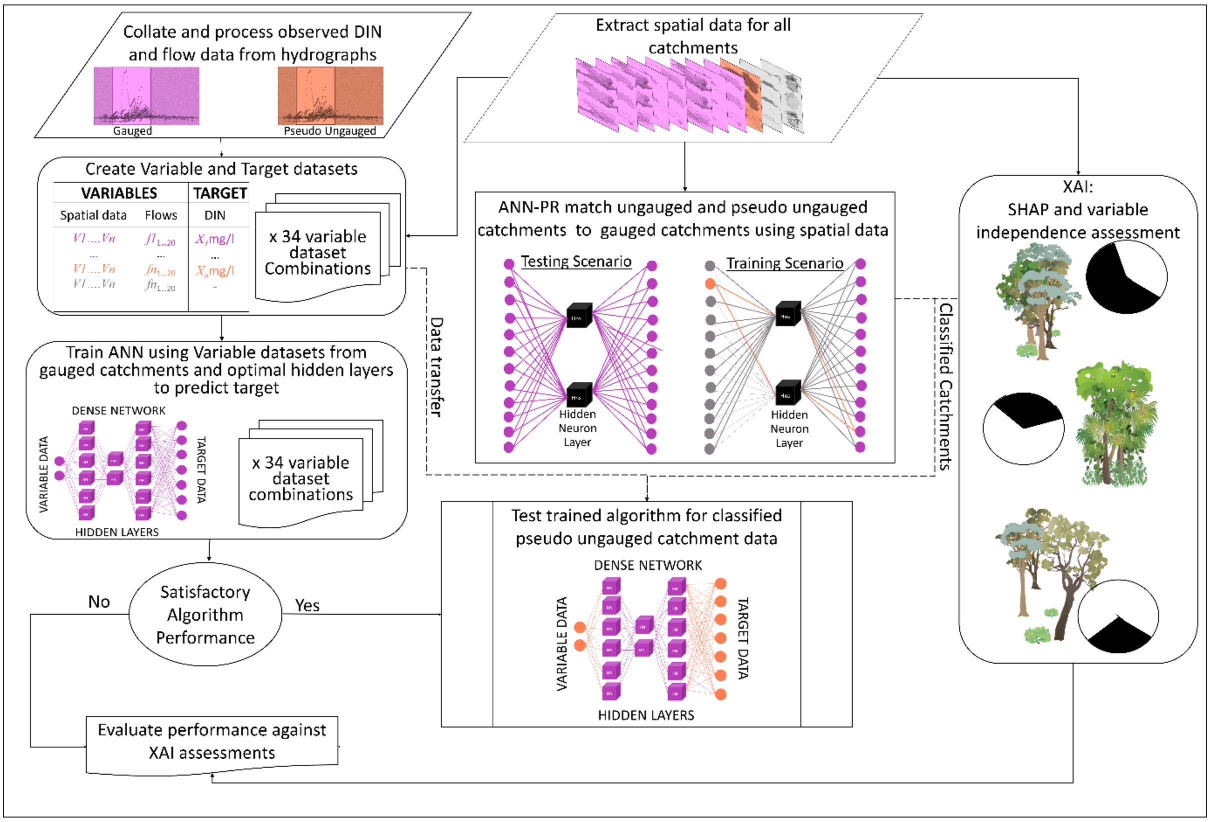


Figure 5.1 Workflow for Classifying Ungauged Catchments for DIN

5.2. Published journal paper

The paper accepted for publication for this chapter is provided below. The Supplementary material in support of the paper is provided in Appendix C.

scientific reports

Check for updates

OPEN

Explainable AI approach with original vegetation data classifies spatio-temporal nitrogen in flows from ungauged catchments to the Great Barrier Reef

Cherie M. O'Sullivan[1]✉, Ravinesh C. Deo[2,3] & Afshin Ghahramani[1,4]

Transfer of processed data and parameters to ungauged catchments from the most similar gauged counterpart is a common technique in water quality modelling. But catchment similarities for Dissolved Inorganic Nitrogen (DIN) are ill posed, which affects the predictive capability of models reliant on such methods for simulating DIN. Spatial data proxies to classify catchments for most similar DIN responses are a demonstrated solution, yet their applicability to ungauged catchments is unexplored. We adopted a neural network pattern recognition model (ANN-PR) and explainable artificial intelligence approach (SHAP-XAI) to match all ungauged catchments that flow to the Great Barrier Reef to gauged ones based on proxy spatial data. Catchment match suitability was verified using a neural network water quality (ANN-WQ) simulator trained on gauged catchment datasets, tested by simulating DIN for matched catchments in unsupervised learning scenarios. We show that discriminating training data to DIN regime benefits ANN-WQ simulation performance in unsupervised scenarios (p< 0.05). This phenomenon demonstrates that proxy spatial data is a useful tool to classify catchments with similar DIN regimes. Catchments lacking similarity with gauged ones are identified as priority monitoring areas to gain observed data for all DIN regimes in catchments that flow to the Great Barrier Reef, Australia.

Communicating catchment influences towards the ecology of the receiving environment is enhanced by water quality simulation tools. Customising water quality simulation models to the catchment they represent is essential for limiting uncertainty in results and maintaining trust in land use decisions they aim to inform[1,2]. Model design and development, referred in here as customisation, is achieved by using observed water quality data from gauging stations for design and verification of models[3]. However, many water catchments globally are ungauged, and a lesser proportion of those have corresponding water quality data to inform model customisation. Techniques to overcome such data voids in ungauged areas are necessary[4,5]. Methods to simulate flows in ungauged areas are well researched[6,7], however, refinement of methods that simulate nutrients in ungauged areas remained unresolved. This knowledge gap in water quality modelling needs addressing to best inform anthropogenic nitrogen management, and to demonstrate progress to the 2030 UN Nations Sustainable Development Goals commitment to reduce land-based nutrients that enter the oceans[8]. This has relevance for the Great Barrier Reef World Heritage Area where over ~ 20% of the terrestrial drainage area is ungauged, and nutrient balances are critical for the reef's health[9,10]. Logical explainability in nutrient models for ungauged areas can support communications and enable more responsive water quality improvement investments[11,12].

[1]University of Southern Queensland, Toowoomba, QLD 4350, Australia. [2]School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, QLD 4300, Australia. [3]Center for Applied Climate Sciences, University of Southern Queensland, Toowoomba, QLD 4350, Australia. [4]Present address: Department of Environment and Science, Queensland Government, Rockhampton, QLD 4700, Australia. ✉email: Cherie.O'Sullivan@unisq.edu.au

For both data-driven and process-based models that simulate water quality, observed water quality and quantity data, as well as a comprehensive understanding of catchment characteristics are required[13]. Data driven water quality models are useful to forecast water quality output, but water flows and water quality must be known a priori to develop covariates[14,15]. In contrast, process-based models use physical and empirical principles and can be established for catchments lacking observed water quality data. In ungauged areas, data is donated to ungauged catchments from the most similar gauged ones[16,17]. Alongside these traditional water quality modelling approaches, deep learning, particularly in the revised forms of Artificial Neural Networks, has been relatively successful in simulating water quality, including nitrogen, without the need for prior established principles[18,19]. As a subsector of Deep Learning, Artificial Neural Networks have demonstrated the ability to recognise patterns in input datasets, classify them, and establish algorithms to match target data. The merits of ANN are demonstrated to forecast and extend non-linear water quality data within respective catchment datasets[20], but their application to inform scenario simulation, and hence land management decisions, which is the benefit of process-based models, is lacking[21].

To exploit the benefits and overcome the drawbacks of each data driven vs process based model approach, the coupling of machine learning models such as ANN with process-based approaches can be performed to provide benefits of transfer learning[22,23]. However, machine learning models that incorporate process considerations for water quality modelling are disproportionally underrepresented in many research articles[22,23]. Additionally, where applied to ungauged areas, low landscape heterogeneity between drivers for the constituent being simulated is necessary[21,24]. While variations in patterns of nutrients are observed across gauged catchments that drain to the Great Barrier Reef[15,25,26], methods for classifying those catchments to the most similar ungauged catchments that drain to the Great Barrier Reef, based on similarity of nitrogen drivers, are unexplored.

In terrestrial landscapes, Dissolved Inorganic Nitrogen molecules are influenced by decomposers, vegetation uptake, nitrogen fixing bacteria etc., which change depending on a unique combination of physical and biological influences at each location[27,28]. The fluxing nature of these biotic processes mean catchment similarities for drivers of DIN differ from the abiotic drivers of flow have therefore been complicated to quantify[22,29]. Variability in drivers of DIN affect the consistency of water quality modelling of ungauged areas[25]. This disparity between biotic and abiotic influence on nitrogen drivers means classical classification approaches that only use physical similarities miss the influence of all biologically influenced differences that may exist between catchments. Spatio-temporal variability in nutrient drivers can be represented in catchment models by the natural physical drivers of geology, aspect, topography, climate etc., as well as land use to represent anthropogenic impacts, including standard fertiliser application rates which affect DIN[25]. Our earlier studies found Original Vegetation is a proxy dataset for the residual biotic responses to these, and any other unaccounted-for drivers that can parsimoniously classify catchments for DIN, and identify the classification drivers using explainable artificial intelligence, (XAI)[30,31].

Explainable artificial intelligence, (XAI) has outstanding capabilities to highlight the influential variables in machine learning algorithms, however, performance criteria for the corresponding ANN models are likely to vary unpredictably with changes to model architecture and scenarios[32]. Established process-based models instead can be customised to respective catchments using regionalised parameter data, enabling trials of different land management scenarios[18]. This technique has been effective for water quality constituents driven by abiotic processes, which result in consistent performance[6], and so is pragmatic for the purpose of informing land management decisions. Despite the acceptable track record for process-based models, the suitability of parameter transfer for constituents with biotic process drivers is still lacking, and studies regarding the spatio-temporal scales are necessary[33–35]. We found earlier that original vegetation can be a proxy for matching gauged catchments with dynamic DIN patterns[26]. However, no approach has yet been developed that matches ungauged to gauged catchments for DIN similarities, which would be beneficial for models that transfer data across catchments with similar processes.

This study extends our previous XAI-SHAP[30,36] approaches to match the currently ungauged to gauged catchments that flow to the Great Barrier Reef using mapped spatial data as a proxy for DIN. Mapped spatial data is useful because it provides data for all areas of the Great Barrier Reef catchments where water quality data is lacking[26]. In this study we verify the classification results by building and applying an ANN-WQ simulator to compare changes in simulation performance criteria for a case study catchment, under various dataset arrangement scenarios. Our earlier studies found dominant original vegetation data features may provide guidance to the part of the hydrograph that is relevant to consider for matched catchments and that it is a useful proxy to group gauged catchments that flow to the Great Barrier Reef to three DIN response categories[30]. In this study we evaluate whether our previous method is extendable to all ungauged catchments that flow to the Great Barrier Reef, and undertake a case study to verify its suitability as a proxy for DIN classification. Our verification case study aims to confirm catchments classified together based on original landscape variables also have transferrable water quality responses that can be exploited to simulate DIN.

The hypotheses investigated here are: (1) Original vegetation spatial features found to be a proxy for DIN discharge from gauged catchments in our previous studies[26,30] can be used to match gauged catchments to ungauged catchments that also flow to the Great Barrier Reef. (2) An ANN-WQ simulator trained using predictor variables of original vegetation, coupled with flow data characterised to match with DIN targets will achieve superior performance compared to an ANN-WQ simulator trained to simulate DIN using non-categorised flow data only.

(3) The trained ANN-WQ simulator can simulate DIN in an unsupervised scenario for a pseudo-ungauged case study catchment matched based on the spatial proxy data and achieve satisfactory performance criteria to verify the suitability of the catchment match approach. For this study, pseudo-ungauged means a gauged catchment, with the same data collection method as the other gauged catchments, but intentionally omitted from previous research that informed this study. DIN data for the pseudo ungauged catchment is used here for hypothesis validation purposes only.

The present study therefore aims to create an XAI approach for considering original vegetation data classification as the proxy for spatio-temporal nitrogen patterns in ungauged catchment flows for the specific case of the Great Barrier Reef in Australia.

Results

ANN-PR matches

Apart from the Mary Catchment, the results show that ungauged portions of gauged catchments do not necessarily classify together, and catchments do not necessarily classify with their nearest neighbours (Fig. 1). Catchment matches varied for each spatial dataset evaluated, and translation of those results to classify catchments based on corresponding DIN response Categories also varied (Table 1). While Category 2 matched catchments generally clustered together spatially, Category 3 matched catchments contrasted with distributions only north of Plane for the Original Vegetation (OV) dataset compared to further south where Land use (LU) variability was included independently or embedded within the Ecounit (EU) data. This indicates that the catchments in the different datasets show different spatial characteristics. For example, the catchments that matched Category 2 tended to be clustered together, while the catchments that matched Category 3 showed more variability when the LU dataset, which represents anthropogenic, in contrast to natural biotic response to environmental influences, was included.

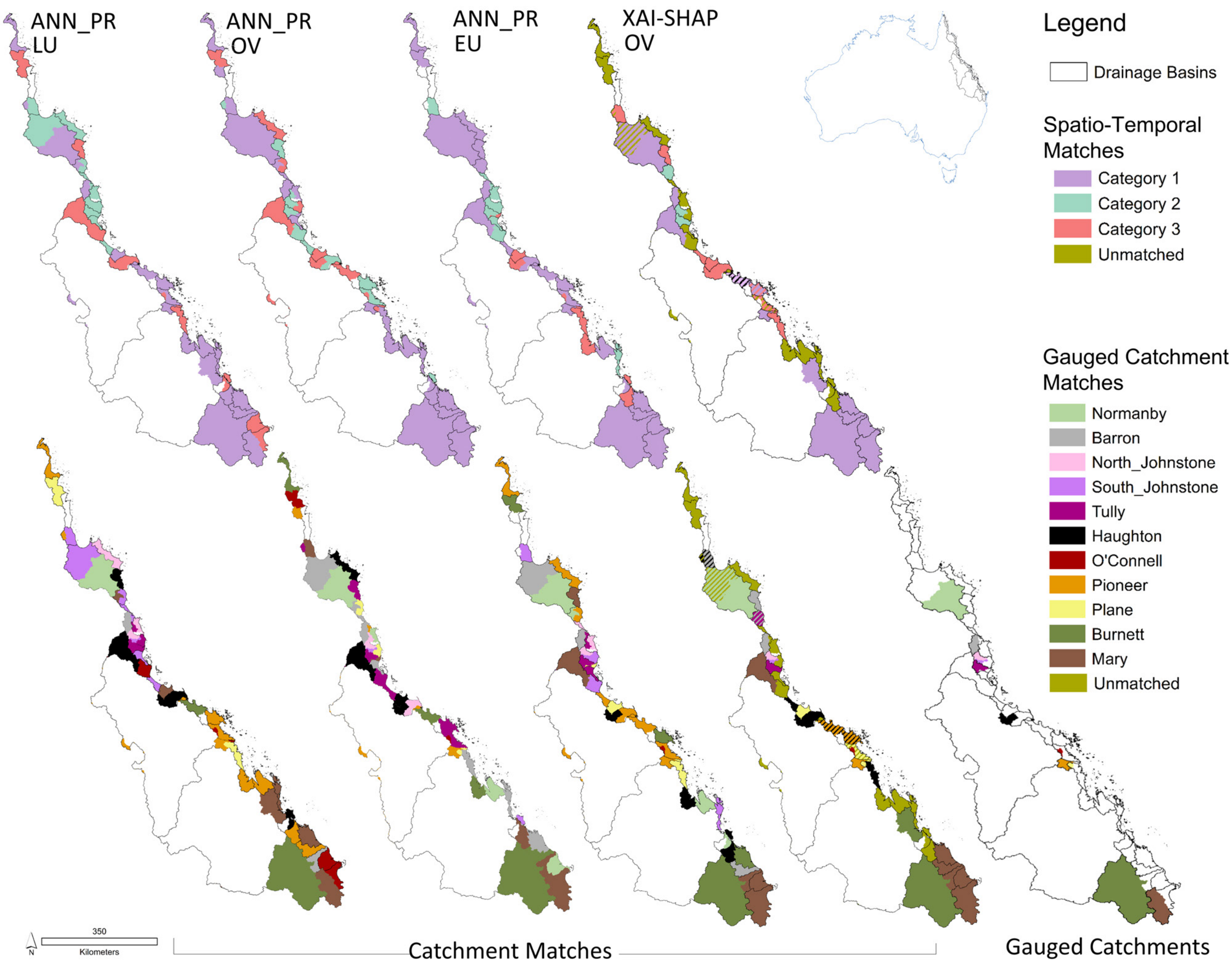


Figure 1. Catchment matches using ANN-PR and XAI-SHAP approach to identify ungauged catchment similarities to gauged catchments using spatial dataset (OV original vegetation, EU ecounit, LU land use). Top row shows the spatio-temporal category of the matched gauged catchment based on the gauged catchment allocation derived from our previous works[30], bottom row shows the matched catchment. Colours represent the gauged catchment as listed in the legend. Results show high variation between each dataset. Maps created by author using ArcMap 10.8.1, gauged catchments[10] supplied, Drainage Basins[37] licenced under a Creative Commons—Attribution 3.0 Australia licence (CC BY 3.0 AU). © State of Queensland (Department of Environment and Science) 2023.

61

Category	Scenario	Performance						Best hns					
		MSE	R2	NSE	d	RMSE	pde	MSE	R2	NSE	d	RMSE	pde
All	Control	0.004	0.390	0.130	0.336	0.060	-9.323	435	879	435	113	435	14
	EULUOV	0.002	0.845	0.634	0.809	0.048	32.780	920	976	920	989	920	2
	LU	0.005	0.361	0.116	0.279	0.073	81.593	893	840	893	926	893	14
	EU	0.006	0.324	0.093	0.219	0.074	12.930	607	848	607	449	607	3
	OV	0.005	0.340	0.111	0.268	0.074	6.920	911	911	911	485	911	27
Category 1	Control	0.008	***0.686***	-4.988	0.454	0.089	-5.208	653	170	653	170	653	363
	EULUOV	0.001	0.613	0.252	0.532	0.036	-10.058	621	988	621	948	621	3
	LU	0.001	0.512	0.247	0.568	0.036	***-0.703***	260	439	260	439	260	28
	EU	0.001	0.581	0.252	0.570	0.036	14.235	587	947	587	332	587	10
	OV	***0.001***	0.649	***0.375***	***0.637***	0.033	-1.830	975	867	975	827	975	16
Category 2	Control	***0.002***	***0.857***	***0.688***	***0.846***	***0.045***	12.451	990	990	990	990	990	8
	EULUOV	0.003	0.823	0.609	0.799	0.050	42.345	949	795	949	949	949	506
	LU	0.002	0.837	0.637	0.814	0.048	34.739	903	619	903	951	903	3
	EU	0.002	0.828	0.619	0.794	0.049	***12.045***	902	941	902	973	902	13
	OV	0.002	0.845	0.634	0.809	0.048	32.780	920	976	920	989	920	2
Category 3	Control	0.002	0.971	0.918	0.947	0.041	-0.250	426	946	426	426	426	380
	EULUOV	0.001	0.974	0.929	0.956	0.038	0.012	990	990	990	558	990	370
	LU	0.002	0.973	0.911	0.948	0.043	***-0.005***	908	614	908	908	908	457
	EU	***0.001***	***0.976***	***0.941***	0.955	***0.035***	0.053	269	269	269	269	269	452
	OV	0.001	0.975	0.938	***0.958***	0.036	-0.009	905	408	905	905	905	502

Table 1. Performance evaluation of ANN-WQ simulator for the Gauged Catchment scenarios. Shading intensity represents performance over all scenarios tested for each criteria. Scores styled **bold** are the best performing metric for each category, scores styled *italic* fail to meet minimum satisfactory performance criteria. Training datasets discriminated by category influence DIN simulation results.

Variable feature independence

In this study, matching variable deviations using the XAI-SHAP approach method[30] revealed that every catchment had a unique combination and weighting of deviated features. The same combinations of top XAI-SHAP 10% floristic structure variables did however match the most similar gauged catchment and group them to Categories based on the combination of deviated variables. It also revealed catchments that did not share the same combinations of deviated variables. Grouping the deviated variables by landform and vegetation descriptors in the Original Vegetation dataset allowed for 20 of the 37 pseudo/ungauged catchments to be matched to individual gauged ones, while 9 catchments were not matched to another ungauged or gauged catchment or spatio-temporal category. Of those, unable to match to gauged catchments XAI-SHAP results facilitated four closely matched groups to be identified.

Variable combinations only occurring in ungauged catchments and not in the gauged ones include: hilly alluvial with basalt, health land with sandplains and coast, or mangrove landform structures, as well as additional combinations of vineforest with woodland drainage, or open forests combined with grassland and open woodlands (Fig. 2).

ANN-WQ simulator performance

The most notable observation was that the combination of catchments included in the training datasets influenced the unsupervised performance of the ANN-WQ simulator (Table 1). When the ANN-WQ simulator was trained using data for individual catchments, simulations were only able to be generated in the unsupervised environment for the Wet Tropic catchments of Tully, and North and South Johnstone. Flatline simulations were observed in the unsupervised simulator environment for all other catchments, despite their adequate training performance (Supplementary Material Fig. SF1).

In contrast, training using data grouped from multiple catchments generated non flatline results for all scenarios. Satisfactory to very good performance for all metrics were achieved for all spatial dataset combinations grouped and discriminated in spatio-temporal Category 2 and 3. Except for their unsatisfactory Nash Sutcliffe Efficiency (NSE) performance, datasets grouped and discriminated to spatio-temporal Category 1 also achieved

Figure 2. Top 10% XAI-SHAP deviations for landform and flora sub-descriptors of the Original Vegetation Datasets. Catchment = subject catchment, Match = Catchment the subject catchment is deemed a closest match with, Category = spatio-temporal category of the gauged catchment as established by previous research[30]. Gauged catchments are shown in the top plates, the ungauged catchments are shown in the bottom 2 plates and arranged chronologically from north to south. Visualisation of this data shows that some catchments have combinations of similar feature deviations to gauged catchments, and others are unique.

satisfactory to very good performance (Table 1). Meanwhile, training datasets that grouped all gauged catchments together only met satisfactory performance criteria for the scenario that included all spatial data variables (i.e., Ecounit, Land use and Original Vegetation (EULUOV)) (Table 1).

Performance criteria for the control (i.e., flow only) scenario also varied where the dataset was first discriminated to spatio-temporal categories (Table 1). Simulation results for the control scenario trained on non-discriminated datasets failed performance criteria (NSE = 0.130), while the opposite was the case for training datasets discriminated by spatio-temporal category (NSE = 0.846 and 0.947 for Category 2 and 3 respectively). The NSE for Category 1 catchment Control worsened after discrimination, however, the R^2 value improved to 0.686 compared to 0.39 for the non-discriminated counterpart. Benefits of including spatial data in datasets were reduced after pre-discriminating to spatio-temporal regime (Table 1). Benefit losses include a lack of independence from the control scenario, as measured by Kruskal Wallis test for independence ($p = 0.483–0.981$), where spatial data was omitted (Supplementary Material Table ST1).

Grouping datasets by respective catchment categories, identified in our previous spatio-temporal study[30], prior to loading to the DIN simulator resulted in improved performance criteria ($R^2 = 0.984$ for Category 3, RMSE = 0.02382 for Category 1). Interestingly, for Category 2 flow datasets the control scenarios, which did not contain spatial variables achieved superior performance for MSE, R^2, NSE and Wilmott's d compared to the other Category 2 scenarios that did include information on spatial variables. In contrast, the Original Vegetation scenario discriminated to Category 3 records had the smallest pde score meaning that the inclusion of Original Vegetation variables improved the ability of the DIN simulator under extremes in the data for Category 3 catchments (Table 1).

Classifying ungauged to gauged catchments: variable independence vs ANN-PR

While the ANN-PR approach matched all ungauged catchments to a gauged counterpart, the XAI-SHAP variable independence approach using relative variable distributions was unable to match 17 catchments. Catchment matches using OV dataset for XAI-SHAP landform and floristic structure, most closely aligned to the ANN-PR catchment matches using the EU dataset. Matching using only the top 10% of deviated features using XAI-SHAP variable independence approach changed the catchment matches compared to the EU dataset using ANN-PR, where all variables are considered, but retained matches generally within the same category (Fig. 1).

Verification of catchment classification for DIN similarities

Both XAI-SHAP Variable Independence and ANN-PR techniques for catchment classification matched pseudo-ungauged Herbert to the Gauged Mary (Figs. 1 and 2) and identified it as a Category 1 catchment. Only Mary catchment training data scenarios achieved a satisfactory performance metric i.e. NSE > 0.5 (Supplementary Material Table ST2). The greatest performance criteria overall collectively clustered towards datasets discriminated to Mary and Category 1 flows only (Fig. 3 and Supplementary Material Table ST2).

Datasets first discriminated by the classified catchment resulted in the best overall performance, with further discrimination to flow regime improving results. Training datasets discriminated by spatio-temporal flows, also performed better where they were also discriminated to the flow regime. This is consistent with findings during the development of the ANN-WQ simulator where a significant difference was noted for training datasets discriminated by the category flow regime. Where catchment specific or catchment category classification was not included, performance improved the most for training datasets that included Ecounit spatial data compared to the control which did not include spatial data. The worst performing scenario was the control grouped to Category 1 catchments only, followed by the control for all gauged catchments that included no spatial data, but was discriminated by flow regime. In comparison to control scenarios, differences in the performance criteria for scenarios that include spatial data diminished for catchments trained only to Mary gauged data. This

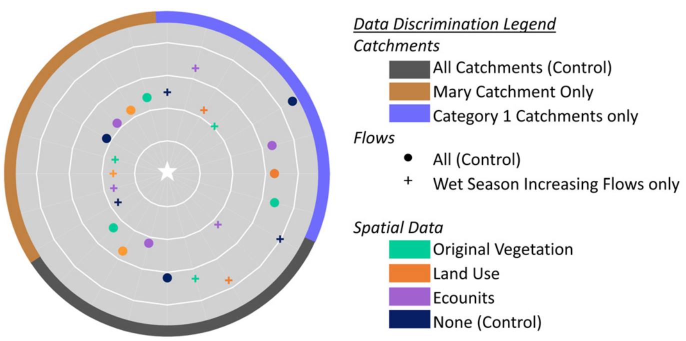


Figure 3. Additive Performance Criteria for pseudo-ungauged catchment DIN simulations. Dimensionless graph shows the additive scores for the best MSE, R^2, NSE, d, RMSE for each data discrimination scenario evaluated in the case study. Zero is the centrally located white star and represents the target/observed data. Discrimination of data to spatio-temporal regime (i.e., shown as crosses) and to the matched catchment improves results the most. Inclusion of Ecounit data improves simulation performance where training datasets are not discriminated to classified catchment or category.

suggests the benefit of adding spatial data reduced as the flow regime was refined to the catchment with the closest similarity to Herbert.

Training data discriminated to the individually matched catchment, Mary, and discriminated to wet season flows achieved the best performing DIN simulations ($R^2 = 0.80$, NSE = 0.62, d = 0.85 respectively). Visualisation of simulated vs true data demonstrates that these scenarios' pre-discriminated spatio-temporal flows result in simulations that include all the peaks in the observed dataset. On the other hand, training data discriminated to include all catchments in the corresponding Category 1, but using all flow and season records, with no spatial data failed to simulate half the peaks (Supplementary Material Fig. SF2). While simulated peaks were under estimated in all cases, a review of the raw data identified that the maximum nitrogen concentration in the dataset for Herbert Catchment was 1.8105 mg/L, which is the highest historical record, plus two additional peaks ranging between 1.320 mg/L and 1.694 mg/L. Maximum concentration for Mary was a smaller with a once off observed peak of 1.243 mg/L during unusual weather conditions of end 2012 start 2013[38,39] with remaining peaks in the dataset not exceeding 0.605 mg/L.

Discussion

Overview

Our research uniquely evaluates the classification potential for all ungauged catchments flowing to the Great Barrier Reef, based on proxy data for spatio-temporal drivers of Dissolved Inorganic Nitrogen (DIN). We adopt an explainable AI approach referred to as XAI-SHAP to provide a deeper understanding of the modelled classification results. In accordance with earlier research works, our satisfactory performance metrics show classification of the pseudo-ungauged area to the most similar gauged ones is validated and works well where data for proxy drivers of DIN are included because they facilitate grouping of catchments by the DIN regime. Evaluation of DIN simulation performances using transfer learning in an Artificial Neural network environment allowed us to demonstrate the variability in DIN patterns depending on the spatio-temporal regime of the ungauged catchments, as exposed by original vegetation data Additionally, the XAI-SHAP method allowed for ungauged catchments with insufficient similarity to the gauged ones to be identified, regardless of being classified by brute force using ANN-PR techniques.

Dataset complexity and consistency

Development and verification of the ANN-WQ simulator to establish DIN response similarities in datasets between pseudo-ungauged catchments with the gauged ones found dataset complexity and representative flow patterns were influential. This highlighted caution in direct application without prior understanding of the DIN to flow dynamics of the catchment. Flatline simulations that resulted in the unsupervised scenario are a known symptom of inadequate complexity in the dataset[40]. Likely explanations include hidden neuron complexity was low in the development trials and relationships between flow, spatial data and DIN response was not adequately formed to facilitate simulations in the unsupervised scenario. The contrasting ability of Wet Tropics catchment datasets to overcome possible lack of dataset complexity in the training dataset is explained by the different DIN and flow dynamics in wet tropics catchments compared to the others[27]. We previously demonstrated that DIN remains elevated in retreating flows for Wet Tropics catchments only[26]. One explanation for the contrast with Wet Tropics catchments could be a more consistent relationship between flows and DIN releases throughout the hydrograph which the ANN-WQ simulator was trained to simulate for[41,42]. This phenomenon demonstrates that consistency of DIN to flow relationships influence the performance of defined algorithm based models developed to use transferred data.

Training dataset influence

For catchments with inconsistent DIN to flow relationships, our results found training data arrangements that group catchments using prior knowledge of spatio-temporal similarities, i.e. either by prior discrimination (discriminated to Category 1, Category 2 and Category 3 as informed by Original Vegetation deviation using XAI-SHAP), or within the model training datasets (non-discriminated but including all EULUOV spatial variables which are identified as proxy drivers for DIN and used to inform XAI-SHAP) improved the performance. This approach to remove heteroskedasticity where seasonal differences for Nitrate are considered has already been shown to benefit model development[42,43]. The significant differences in performance criteria of DIN simulations ($p = 0.003$–0.045) depending on the data discrimination for catchment categories suggests that DIN dynamics differ between those categories. This finding of significant variation in nitrogen regimes through the Great Barrier Reef catchments, as demonstrated by the ANN-WQ simulator training dataset predictive performance, regardless of anthropogenic influence is consistent with proceeding research[34,44,45]. Our research shows variability in DIN regimes is an influential consideration for data transfer purposes in water quality models. The improved performance criteria where information on proxy drivers of DIN was considered supports our application of original vegetation spatial datasets used in this study to discriminate differences in DIN regimes for each catchment[30].

Spatio-temporal category differences for DIN simulation performance may be explained via the wide body of literature that demonstrate nitrogen is either flow or production limited[44,46,47], and also influenced by connectivity to stream network[48]. The superior performance of the control scenario for Category 2, compared to Category 2 scenarios that included spatial data, indicate this category is flow limited and abundant in DIN. It is demonstrated that soils higher in total organic carbon, consistent with rainforest soils, have higher supplies of nitrogen created by the residual soil biology[30,49,50]. The abundance of nitrogen generation in the soils, coupled with abundant flows, in the Wet Tropics catchments can result in consistent nitrogen to flow patterns and is a logical explanation for the ability of the ANN-WQ simulator to generate results in the development trials, where catchments from other categories flatlined. This supports our previous suggestions[30] that the timing of data

65

Note: Adjustment to this page has been made as underlined toalign with a post production correction request.

collection is important in Category 1 and 3 catchments, while Category 2 catchments could classify regardless of the season or flow phase. While our research is not designed to interrogate reasons for drivers of DIN in each category per-se, this is one of many possible explanations for how categorising datasets by vegetation removes noise associated with different combinations of biotic responses in each location[34,41,51]. The findings, therefore, support the second hypothesis that prior grouping of catchments by categories of Original Vegetation, as a proxy for the DIN to flow regime, is a necessary first step for identifying catchments that share similar DIN patterns.

Training dataset discrimination and variable combinations separately influenced the performance of the ANN-WQ simulations of DIN for the pseudo ungauged catchment. The ANN-WQ simulator achieved the best DIN simulation performance metrics for the pseudo-ungauged catchment when trained on data only from the classified catchment and therefore highlights that data transfer with the classified catchment achieved the best results. Concurrently, discriminating the dataset by the respective flow regime of wet season increasing flows had a greater influence on simulation performance than inclusion of spatial data variables. In contrast, training datasets using data from multiple catchments from the same flow regime, i.e. Category 1, achieved equivalent performance only where the training data was first discriminated to increasing above average flow regime, hence removing hetroskedacity of DIN in the retreating and below average flows. Both these findings are consistent with the ANN-WQ development phase and demonstrate that prior discrimination of the training dataset to flow regime reduces heteroscedasticity in DIN patterns to flow[41,52]. Once heteroskedasticity in the training dataset was removed, the influence of spatial variables as drivers to the DIN patterns became less relevant. Separately, the research also found that training the ANN-WQ simulator using data from all catchments improved where all EULUOV variables were included. This could be attributed to the ANN-WQ simulator discriminating datasets within the algorithms, as opposed to prior discrimination provided by classification, and further demonstrates the benefit of the spatial datasets to expose the drivers of the DIN patterns.

ANN-PR vs XAI-SHAP classification

Catchments matched using ANN-PR were not always the same as the catchments recommended to be matched by the XAI-SHAP deviation approach for variable independence. One reason could be that only 10% of the most influential variables were considered in the XAI-SHAP approach, in contrast the less deviated variables contributed to the ANN-PR matches. Catchment classification informed by the match options in both ANN-PR and XAI-SHAP approaches provide foundational guidance to rationalise catchments to evaluate in future data transfer investigations or models for DIN simulations to the Great Barrier Reef[53]. Varied performance for each scenario trialled in the ANN-WQ simulator development phase demonstrated that training data from catchments with the most similar proxy drivers of DIN dynamics, was more suitable for data transfer compared to training data from all catchments lumped together[54]. This demonstrates that rationalising training data to the most similar responding catchments reduces heteroskedasticity in the training dataset and benefits DIN simulation accuracy for the classified catchment. XAI-SHAP provided insight to identify catchments grouped by known DIN to flow proxy drivers. While classification using nearest neighbour catchments has historically been supported for their influence towards flow similarities[15,55], our finding demonstrates that catchments with the most similar drivers of DIN, in addition to flow, are not necessarily located as the neighbouring catchment, and are influential towards DIN simulation performance.

Practical application

This study established Original Vegetation as a suitable proxy for DIN dynamics for the benefit of water quality modelling. Therefore, the 20 ungauged catchments that matched to gauged ones, based off Original Vegetation similarity have justification to receive data from the corresponding gauged catchment. The remaining 21 ungauged catchments had combinations of original vegetation unique from the gauged catchments and therefore did not support the hypothesis that they share similar DIN drivers with gauged catchments. Consequently, this study found that only 20 of the 41 ungauged catchments were suitable to consider for data transfer with existing gauged catchments for satisfactory water quality modelling purposes.

Of the ungauged catchments that failed to match to gauged ones, 4 groups shared unique combinations of deviated spatial variables. Deviated original vegetation floristic structure and landform descriptors shown by the XAI-SHAP deviations, were grassland, heathland and mangrove. These are all coastal ecosystems and differ from vineforest, open woodlands and forest shown to be proxy indicators of DIN dynamics for gauged catchments[30] While data transfer from existing gauged catchments to the four coastal catchments is not supported by our study, our method can instead be used to inform where new water quality monitoring and gauging sites could have the greatest value to represent all DIN regimes[12,47,56–58]. New monitoring and gauging sites are recommended in each of the four coastal catchment groups to collect data representative of all DIN regimes, which could facilitate data transfer for modelled DIN predictions across all ungauged Great Barrier Reef catchments.

It is well known that performance of neural networks deteriorates when the unsupervised scenario includes extremes outside the range of the training dataset[59] and in our evaluation, models trained using Mary data were never exposed to high concentration peaks observed in the Herbert catchment. Limitation for simulating extremes not included in the training data could be addressed with differing model techniques, the ANN-WQ simulator was intended only as a coarse method to verify whether similarity in DIN drivers exists between catchments matched using the ANN-PR method, and this was demonstrated.

For the case study, the matched catchment was a Category 1 catchment. Collectively, Category 1 catchments showed the poorest performance in the ANN-WQ development phase. The fact the case study trial achieved satisfactory performance criteria for the poorest performing category in the development phase, it is expected that better results can be achieved for Category 2 and Category 3 catchments which have less heteroskedastic DIN to flow relationships. Further studies to refine the ANN-WQ simulator performance, along with a full comparison of all training dataset options, i.e. discrimination of data to Category 2 and 3 flows to evaluate difference in performance and facilitate year round classification is recommended. Regardless of these limitations, we encourage results from this study to be applied in established models that will benefit from data transfer from the most similar catchments for purpose of DIN modelling, and intentionally developed for superior performance[43].

Conclusions

This study matched all ungauged catchments that drain to the Great Barrier Reef to the gauged ones using ANN-PR coupled with Land use and Original Vegetation datasets. While ANN-PR enabled matching using proxy datasets for drivers of DIN, XAI-SHAP method explained similarities between catchments based on feature deviations as well as concurrently allowing grouping of catchments to known spatio-temporal categories. Prior knowledge of spatio-temporal DIN response categories within training datasets improved performance of the ANN-WQ simulator developed to verify catchment matches.

While all catchments matched to a gauged one using ANN-PR, consistent with hypothesis 1, the additional interrogation by XAI-SHAP deviations found 17 catchments did not share deviated feature similarity with a spatio-temporal category. The XAI-SHAP method instead provides justification to prioritise gauging and monitoring efforts in those unmatched catchments to better understand the spatial temporal dynamics of DIN in coastal areas that those unmatched catchments were located in. For the ungauged catchments that did match to gauged ones using the XAI-SHAP method, the subsequent ANN-WQ simulator development and case study to test the second hypothesis, found prior discrimination of data included in the training dataset, based on the spatio-temporal category of the ungauged catchment, improved performance of the ANN-WQ simulator in all scenarios tested. It was, however, an unexpected finding that, after the spatio-temporal discrimination by category was first applied, inclusion of Original Vegetation, Ecounit or Land Use variables had insignificant influence on results. Findings that emerged throughout this study therefore built nuance to our expected hypothesis 3 whereby although a trained ANN-WQ simulator successfully simulated DIN in the unsupervised scenario, it was the knowledge provided by original vegetation data to pre-process the training datasets into categories that mattered. Implications of these findings are that XAI-SHAP coupled with Original Vegetation data has demonstrated merit for customising catchment matching to the portion of water quality datasets most likely to share similar DIN to flow regimes between gauged and ungauged catchments.

Methods

Study area

This study includes all catchments that flow to the Great Barrier Reef, in north-eastern Australia. Each of those catchments is referred to herein as gauged, ungauged and pseudo-ungauged as shown in Fig. 4. Respective gauging allocation, sampling frequency for DIN, and flow data availability for each of the catchments are provided in Supplementary Material Table ST3.

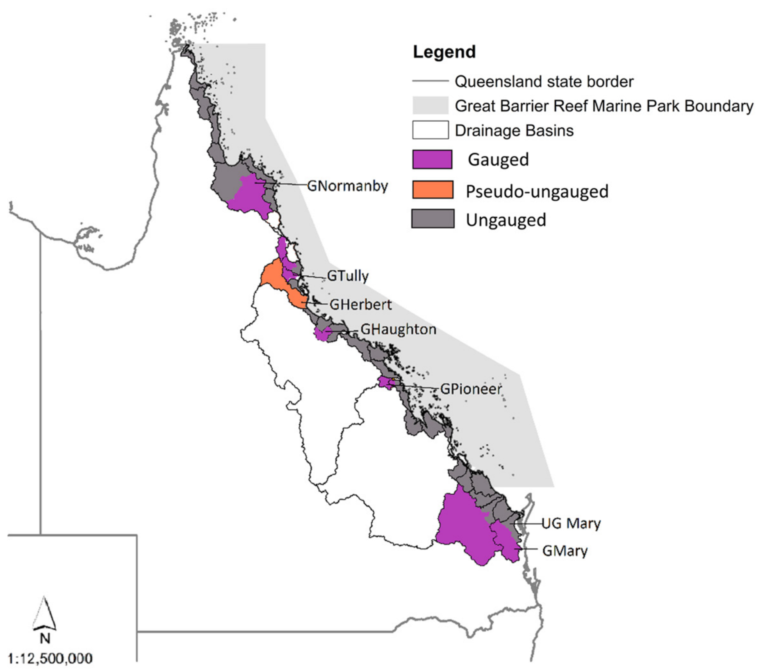


Figure 4. Study Area. G preceding catchment name infers a true gauged catchment. UG preceding catchment name infers ungauged catchment. Maps created by author using ArcMap 10.8.1, ungauged catchments[10] supplied, Drainage Basins[37] licenced under a Creative Commons—Attribution 3.0 Australia licence (CC BY 3.0 AU). © State of Queensland (Department of Environment and Science) 2023.

Study concept

The objective of this study is to establish whether patterns in the flow and spatial variable datasets contain sufficient information to simulate Dissolved Inorganic Nitrogen (DIN), and whether forecasting capabilities can extend to new catchments, referred to in this study as pseudo-ungauged. Because the influence of every variable input and their interrelationships to overall DIN response are unknown a priori, a dense fully connected Artificial Neural Network (ANN) algorithm was developed to trial the proof of concept approach. Algorithms were trained for a number of dataset arrangements and their performance metrics were compared to quantify the viability of the novel forecasting/data transfer concept within the Artificial Intelligence modelling environment. A workflow conceptualising the research approach is shown in Fig. 5 below.

Study dataset

Observed water quality data and flow records for those gauged and pseudo-ungauged catchments are from locations listed in Supplementary Material Table ST3. This data was sourced from Queensland State Government and was cleaned, transformed and flows arranged as detailed in our foundational research[26]. The spatial extent of gauged areas for catchments evaluated in this study are consistent with Khan et al.[13]. DIN records were collected at irregular frequencies depending on flows for each gauged catchment as detailed in Supplementary Material Table ST3. To overcome a large number of NaN values within a time series arrangement for the dataset, daily average stream and baseflows for 90 days preceding each DIN record were allocated as 90 separate column variables each on the same dataset row position as the corresponding DIN record as 1 day prior, 2 day prior….90 day prior. 90 days prior flows capture a full temperate climate season preceding each DIN record and were shown by cross correlations to be sufficient to capture residual information[60]. The water quality and flow datasets were duplicated then partitioned as outlined in O'Sullivan et al.[30] for wet season/increasing flows, dry season/retreating flows, and all flows/seasons to capture spatio-temporal influences.

Spatial data for all gauged and ungauged portions of catchments in the study area, were extracted from Queensland Government Q-Spatial mapping platform, as per the methods described in O'Sullivan et. al.[26]. The three separate spatial datasets were created as the proxy drivers of DIN. These included: Land use to represent

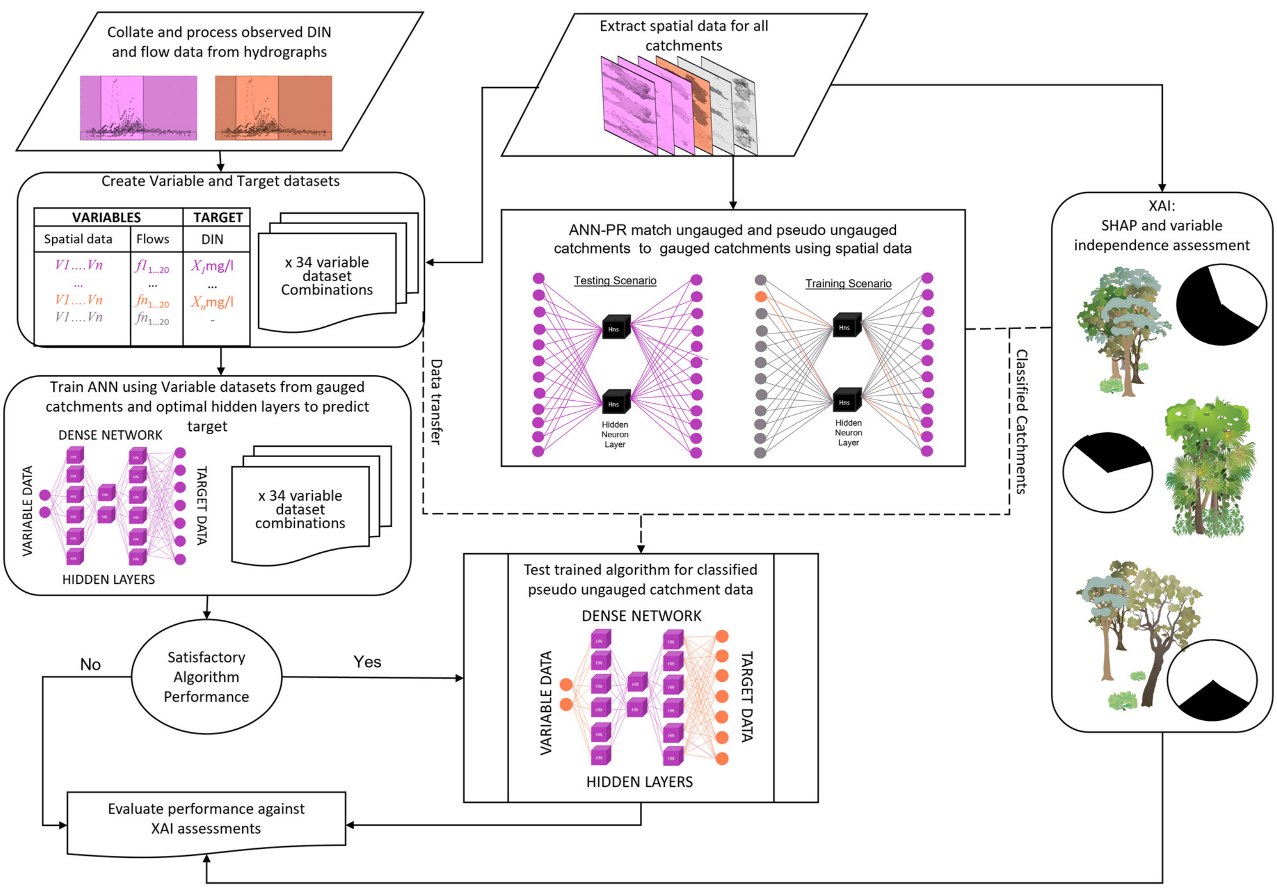


Figure 5. Conceptual framework of research. This framework shows data preparation for classification, process for simulating DIN for pseudo ungauged catchments, and relationship to XAI evaluation of results. Data relating to gauged catchments is represented in purple, orange represents pseudo-ungauged catchments, and ungauged catchments are represented by grey. Dashed lines show source and destination of data transfer for pseudo-ungauged catchment in the research. Vegetation images adapted from the Integration and Application Network licenced under Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

68

human biotic influence which included 6 variables[61], Original Vegetation consisting of 38 variables[31] intended as a parsimonious biotic response proxy for natural DIN responses across the catchment[26,62], and Ecounit which was a created via a combination of Land use and Original Vegetation and resulted in 179 variables. The area of variables for each catchment was established via clipping the spatial datasets to the catchment boundaries and extracting corresponding data tables from ArcGIS. The area of the gauged catchments extended only to the gauged monitoring point, the ungauged portion was created as a sub-catchment polygon for all areas that drain to the catchment's waterway downstream of the gauged monitoring point, or for fully ungauged catchments. For each catchment, the spatial dataset was duplicated to match the number of data rows to the same number of DIN records in each catchment dataset. For the pseudo-ungauged datasets, the number of spatial dataset rows were duplicated to match to the number of daily average flow records available.

A master dataset was created by joining the preceding flow and spatial dataset to create the training variables dataset, and the corresponding DIN data allocated as the target dataset. All data in each dataset was then normalised. Scenario datasets were then created by extracting subsets of data from the master dataset as detailed in Table 2.

Classifying gauged catchments to ungauged and pseudo-ungauged catchments

The novel aspect of this research is establishing whether pseudo-ungauged, and ungauged catchments share spatial data similarities suitable for classifying to gauged catchment classifiers, and for water quality classification data transfer purposes. Our previous studies used ANN-PR to classify only the gauged catchments together using the same spatial variables used in this study[26,30]. XAI evaluations of those datasets provide explainability to the

Dataset type	Purpose	Catchments included	Dataset reference	Variables (discriminated by:)				Target data	
				LU spatial variables	OV spatial variables	EU spatial variables	90 day prior flows	Corresponding DIN	Corresponding catchment
1. Spatial Classification Training Datasets	Training algorithms to match spatial data to corresponding gauged catchment	All gauged	1. GCAFALU	✓	–	–	–	–	✓
			1.GCAFAOV	–	✓	–	–	–	✓
			1.GCAFAEU	–	–	✓	–	–	✓
			1.GCAFAEULUOV	✓	✓	✓	–	–	✓
2. Spatial Classification Testing Datasets	Classify pseudo-ungauged and ungauged Catchments to gauged catchment	Individual ungauged/pseudo-ungauged	2.CAFALU	✓	–	–	–	–	Hidden
			2.CAFAOV	–	✓	–	–	–	Hidden
			2.CAFAEU	–	–	✓	–	–	Hidden
			2.CAFAEULUOV	✓	✓	✓	–	–	Hidden
3. ANN-WQ simulator development- training datasets	Establish whether recognisable patterns exist in the datasets to forecast DIN	All gauged together (non-discriminated)	3.GCAFALU	✓	–	–	✓	✓	–
			3.GCAFAOV	–	✓	–	✓	✓	–
			3.GCAFAEU	–	–	✓	✓	✓	–
			3.GCAFAEULUOV	✓	✓	✓	✓	✓	–
		All gauged—individual (Discriminated by catchment)	3.GC$_{i_1...i_n}$FALU	✓	–	–	✓	✓	–
			3.GC$_{i_1...i_n}$ FAOV	–	✓	–	✓	✓	–
			3.GC$_{i_1...i_n}$ FAEU	–	–	✓	✓	✓	–
			3.GC$_{i_1...i_n}$ FAEU-LUOV	✓	✓	✓	✓	✓	–
		Gauged Catchments grouped by Spatio-Temporal Category (Discriminated by Category 1, Category 2 or Category 3)	3.GC$_{1...3}$F$_{1...3}$LU	✓	–	–	✓	✓	–
			3.GC$_{1...3}$F$_{1...3}$OV	–	✓	–	✓	✓	–
			3.GC$_{1...3}$F$_{1...3}$EU	–	–	✓	✓	✓	–
			3.GC$_{1...3}$F$_{1...3}$EU-LUOV	✓	✓	✓	✓	✓	–
4. ANN-WQ Simulator-Trial Datasets	Evaluate suitability of matched dataset for data transfer to pseudo-ungauged catchment for DIN simulation purposes	Individual gauged and pseudo-ungauged catchment	4.C$_{i_1...i_n}$FALU	✓	–	–	✓	Hidden	–
			4.C$_{i_1...i_n}$ FAOV	–	✓	–	✓	Hidden	–
			4.C$_{i_1...i_n}$ FAEU	–	–	✓	✓	Hidden	–
			4.C$_{i_1...i_n}$ FAEULUOV	✓	✓	✓	✓	Hidden	–
		Gauged and pseudo-ungauged catchments grouped by categories	4.C$_{1...3}$F$_{1...3}$LU	–	✓	–	✓	Hidden	–
			4.C$_{1...3}$F$_{1...3}$OV	–	✓	–	✓	Hidden	–
			4.C$_{1...3}$F$_{1...3}$EU	–	✓	–	✓	Hidden	–
			4.GC$_{1...3}$F$_{1...3}$EU-LUOV	✓	✓	✓	✓	Hidden	

Table 2. Summary of data included in scenario datasets. Variations in each dataset intended to evaluate the influence of spatial data or flow variable toward DIN response. *G* = Gauged, *C* Catchment, *F* Flows, *A* All, *LU* LandUse, *OV* Original Vegetation, *EU* Ecounits $C_{i_1...i_n}$ reference to individual catchments.

corroborating ANN-PR results for both spatial data and water quality classification[30]. Here, we explore, for the first time, extending that classification approach beyond the gauged portion of the study area to classify catchments of the ungauged and pseudo-ungauged areas to gauged catchments. The method is therefore extending the ANN-PR approaches of our previous studies to now evaluate which gauged catchments the pseudo-ungauged and ungauged catchments classify to, and evaluate whether XAI explainability applies to the ANN-PR matches. To accomplish this, we apply a combination of the ANN-PR approach used in our previous studies and XAI explainability coupled with SHAP[36] (XAI-SHAP) to evaluate the similarities between the catchments. This allows classification of catchments that have not been gauged, based on the similarities between the gauged catchments, and provide a better understanding of the underlying similarities between the catchments. Importantly, inclusion of the XAI-SHAP method demonstrates whether the sufficient underlying similarity is likely to exist between the proxy drivers of DIN in the gauged and ungauged catchments for the purpose of data transfer.

Spatial classification using ANN-PR

This step used a similar approach explained in detail in our previous studies, however this time we trained the ANN-PR tool on all 11 gauged catchments, and introduced the spatial variable data for the ungauged catchments in the unsupervised environment, to force a match to one of the 11 gauged catchments. A 100-fold duplicate of each spatial variable in each gauged catchment was used to estimate the percentage match between the ungauged catchments and the gauged catchments. We then trained the ANN-PR classification tool in a supervised environment by applying the gauged catchment classification training datasets to standard codes extracted from "*MATLAB 2020a (The MathWorks Inc., 2020) Deep Learning toolbox* (Fig. 4*). The code used is a two-layer feed-forward network, with sigmoid transfer function in the hidden layer, and softmax transfer function in the output layer (The MathWorks Inc. 2020)"*[63]. For the spatial datasets, heuristics and previous knowledge for the gauged data spatial dataset meant that an architecture of 3 hidden neurons were used to set the classification training architecture for this model. Data were split within the coding architecture to 70% for network training, 15% network validation and 15% network testing. In the training phase, the network is designed to match spatial data variables for each row in the dataset to one of the 11 the gauged catchment categories the spatial data is sourced from. The network architecture is set such that training continues towards minimisation of cross entropy and stops once mean square error elevates above its minimum pivot point at which point the ANN-PR algorithm achieves optimal performance[64]. Optimal performance is for each of the 100 replicates of spatial data to allocate to the catchment category the data belonged to in the validation and testing phase.

Testing datasets were separately introduced in an unsupervised environment to the optimised classification algorithm trained to match spatial data to only one of the 11 gauged catchments. Spatial variables for each ungauged or pseudo-ungauged catchment were duplicated 100 times so that the catchment the ungauged or gauged spatial dataset was classified to was based on 100 replicates. The algorithm forces each of the 100 rows of spatial data variables the ungauged or pseudo-ungauged catchment to match to one of the 11 classifiers in the trained environment. This approach was repeated for the Land use, Original Vegetation and Ecounit spatial datasets for all 41 ungauged and pseudo-ungauged catchments. The gauged catchment with allocations of more than half the records for each gauged or ungauged catchment was deemed classified for the respective dataset.

Identifying variable feature independence in both gauged and ungauged catchments

The purpose of XAI, is to deduce the combination of variables most likely to have resulted in the classification between two catchments. To verify that the forced matches between gauged and ungauged catchments using ANN-PR were explainable, we therefore extend the additive deviation approach from previous work[30] to spatial variables for all catchments, shown in Eq. (1), and graphed the top 10%[30].

$$D_s = A_s - A_\forall \tag{1}$$

where: D: deviation of spatial dataset variable. A: proportional area of variable (A = area of variable /total catchment area), S: subject variable, ∀: all dataset variables excluding S.

Variables in the top 10%deviated from the mean were then graphed and visually compared for similarities between the deviated variables for gauged and ungauged catchments sharing similar combinations of deviations were categorised together. Because the Original Vegetation dataset had previously been shown to explain the ANN-PR matches between the gauged catchments[30] it was used directly in this study. Geology and landform has also been demonstrated as a fundamental driver of nitrate in hydrological processes[48], therefore in this study we also further scrutinised influences of the original vegetation dataset by breaking each variable down into its separate landform type and floristic structure descriptor as described by the data authority[31] to better visualise hydrological drivers of results.

Training ANN to forecast DIN in a supervised environment

An Artificial Neural Network water quality (ANN-WQ) simulator was developed to facilitate a rapid assessment of the similarity of matched catchments for DIN. The ANN-WQ simulator was intended for rapid comparison purposes only, and therefore method optimisation was outside the catchment classification scope intended for this research. Similarity between catchments was evaluated by the comparative accuracy of DIN simulations generated for catchments depending on the dataset scenario included in the ANN-WQ simulator training phase.

For each gauged catchment dataset, a Dense Deep Learning feed forward network was created in Matlab. The dense fully connected learning approach was selected to facilitate for all data relationships to be considered, to maximise the pattern recognition ability within the dataset, timesteps of data are still captured in variables as the corresponding time-date number. This architecture was resource intensive and therefore a ReLU hidden layer

Performance metric	Equation	Satisfactory criteria
Correlation Coefficient (R^2)	$R^2 = \left(\frac{\sum_{i=1}^{N} \left(Y_i^{obs} - {Y_{mean}}^{obs}\right)\left(Y_i^{sim} - {Y_{mean}}^{sim}\right)}{\sqrt{\sum_{i=1}^{N} \left(Y_i^{obs} - {Y_{mean}}^{obs}\right)^2} \sqrt{\sum_{i=1}^{N} \left(Y_i^{sim} - {Y_{mean}}^{sim}\right)^2}} \right)^2$ Equation 2[69]	> 0.5
Nash–Sutcliffe coefficient (NSE)	$NSE = 1 - \left[\frac{\sum_{i=1}^{N} \left(Y_i^{obs} - Y_i^{sim}\right)^2}{\sum_{i=1}^{N} \left(Y_i^{obs} - Y_i^{mean}\right)^2} \right]$ Equation 3[54,70]	> 0.5
Willmotts index (d)	$d = 1 - \left[\frac{\sum_{i=1}^{N} \left(Y_i^{obs} - Y_i^{sim}\right)^2}{\sum_{i=1}^{N} \left(\left\|Y_i^{sim} - {Y_{mean}}^{obs}\right\| + \left\|\left\|Y_i^{obs} - {Y_{mean}}^{sim}\right\|\right)^2} \right]$ Equation 4[71]	> 0.5
Root mean square error (RMSE)	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(Y_i^{sim} - Y_i^{obs}\right)^2}$ Equation 5[69,72]	Lowest
Peak percentage deviation (pde)	$100 \sum_{i=1}^{N} \frac{1 - Y_{max}^{sim}}{Y^{obs}_{max}}$ Equation 6[73]	< ±25
Mean absolute error	$MAE = \frac{1}{N} \sum_{i=1}^{N} \left\|Y_i^{sim} - Y_i^{obs}\right\|$ Equation 7[72]	Lowest

Table 3. Performance Metrics and nominated criteria for ANN_WQ simulation scenarios. Where: N = number, i = iteration, Y_i^{obs} = Observed data, Y_i^{sim} = target data from model simulation, ${Y^{mean}}_{sim}$ = mean of the simulation, ${Y^{mean}}_{obs}$ = mean of observed, ${Y^{max}}_{obs}$ = maximum of observed, ${Y^{max}}_{sim}$ = maximum of simulation.

activation was included due to its superior ability to deal with weights and bias over large intensity variations, as could be expected in the dataset[65,66].

Training datasets involved a data set split of 80% Training, 10% Verification and 10% Testing. Development of the Dense Deep Learning feed forward network began with a trial and error phase to scope for functionality at the default hidden neurons (< 10). To overcome inadequate complexity and dimensionality within datasets, trials of 1 to 1000 hidden layers were then undertaken for each dataset to identify the best performing hidden layer network suited to the training dataset[67]. Trialling up to 1000 hidden layers on big data creates heavy computing demands, therefore, Adam optimiser was selected for its minimal memory usage benefits whilst also addressing sparse gradients and non-stationary objectives[68]. The model performance metrics comprised of RMSE, MSE, Nash Sutcliffe Efficiency, Peak Deviation and Correlation as R^2 were recorded for each of the hidden layer trials, and the algorithm with the best performance metrics evaluated for the optimal hidden neuron and for pass or fail of satisfactory performance criteria. The performance metrics equations in Table 3 identify the corresponding satisfactory performance criteria for each. For this research, the ANN-WQ simulator was used to validate whether DIN patterns were detectable. Therefore, performance criteria that identified whether the results were satisfactory or not as nominated in Table 3 were selected to remain consistent with satisfactory performance criteria for water models published elsewhere[69–73].

To compare model DIN forecasts against observed DIN forecasts, the algorithm was rerun with the optimal number of hidden layers for every dataset. We normalized all data for graphing so we could compare forecasting potential across the different datasets.

DIN forecasting potential for classified pseudo-ungauged catchments

The trained algorithms that met minimum satisfactory performance criteria as well as demonstrating a simulation ability in the supervised environment were then used in an unsupervised environment to simulate DIN for their respective classified pseudo-ungauged catchment datasets based on flow inputs. For the data available, only the pseudo-ungauged Herbert was suitable for evaluation for the study and is evaluated as a case study within this article as proof of concept. For this study, scenario datasets evaluated included the ANN-WQ simulation results for the pseudo-ungauged catchment trained on the matched gauged catchment, ANN-WQ simulator trained using all gauged catchment data, and ANN-WQ simulator trained using data from the matching spatio-temporal category. Performance metrics for each scenario were then collated and visualised in a dart plot. To create the dart plot, performance metrics were adjusted using Eq. (8) to make zero the target score. This equation has not been scaled for the impact each performance metric has towards the accuracy of the model, but is developed here for rapid comparison of overall scenario performance.

$$\cup_{PC_k} = \sum_{\in PC_{1....n}}^{\in PC_1} k, \tag{8}$$

where: PC = Performance Criteria, K = unsupervised portion of ANN-WQ simulation scenario, 1….n = a performance criteria adjusted to make zero target i.e.{R^2_c, NSE_c, d_c, RMSE, MAE,Pdv_c}.

Where:

$R^2_c = 1 - R^2$.

$NSE_c = 1 - NSE$.

$pde_c = \left(pde \text{ if } pde>0\right| - pde \text{ if } Pdv<0\right) \times 0.01$

Data availability

The datasets analysed during the current study are available from the corresponding author on reasonable request, as well as in raw form from the following public sources: Observed water quality and flow records[37,74]—Queensland Government Water Monitoring Information Portal: https://water-monitoring.information.qld.gov.au/. Original Vegetation[31]—Pre-clearing broad vegetation groups—Queensland (v4): http://qldspatial.information.qld.gov.au/catalogue/custom/search.page?q=%22Pre-clearing broad vegetation groups - Queensland%22. Land Use[61]—Land use mapping—1999 to 2017—Queensland http://www.qld.gov.au/environment/land/vegetation/mapping/qlump/.

Received: 2 June 2023; Accepted: 17 October 2023
Published online: 24 October 2023

References

1. Creighton, C., Waterhouse, J., Day, J. C. & Brodie, J. Criteria for effective regional scale catchment to reef management: A case study of Australia's Great Barrier Reef. *Mar. Pollut. Bull.* **173**, 112882. https://doi.org/10.1016/j.marpolbul.2021.112882 (2021).
2. Gupta, A. & Govindaraju, R. S. Propagation of structural uncertainty in watershed hydrologic models. *J. Hydrol.* **575**, 66–81. https://doi.org/10.1016/j.jhydrol.2019.05.026 (2019).
3. Vu, D. T., Dang, T. D., Pianosi, F. & Galelli, S. Calibrating macro-scale hydrological models in poorly gauged and heavily regulated basins. Hydrol. Earth Syst. Sci. Discuss. https://doi.org/10.5194/hess-2023-35 (2023).
4. Garna, R. K., Fuka, D. R., Faulkner, J. W., Collick, A. S. & Easton, Z. M. Watershed model parameter estimation in low data environments. *J. Hydrol. Reg. Stud.* **45**, 101306. https://doi.org/10.1016/j.ejrh.2022.101306 (2023).
5. Tegegne, G. & Kim, Y. O. Modelling ungauged catchments using the catchment runoff response similarity. *J. Hydrol.* **564**, 452–466. https://doi.org/10.1016/j.jhydrol.2018.07.042 (2018).
6. Hrachowitz, M. *et al.* A decade of predictions in ungauged basins (PUB)—A review. *Hydrol. Sci. J.* **58**(6), 1198–1255. https://doi.org/10.1080/02626667.2013.803183 (2013).
7. Jahanshahi, A., Patil, S. D. & Goharian, E. Identifying most relevant controls on catchment hydrological similarity using model transferability–A comprehensive study in Iran. *J. Hydrol.* **612**, 128193. https://doi.org/10.1016/j.jhydrol.2022.128193 (2022).
8. UN General Assembly (2015), Transforming our world: the 2030 Agenda for Sustainable Development, 21 October 2015, A/RES/70/1, United Nations, New York. Available at: https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981 [Accessed 16 March 2021].
9. Ani, C. J., Smithers, S. G., Lewis, S., Baird, M. & Robson, B. eReefs modelling suggests Trichodesmium may be a major nitrogen source in the Great Barrier Reef. *Estuarine Coast. Shelf Sci.* https://doi.org/10.1016/j.ecss.2023.108306 (2023).
10. Wells, S. C., Cole, S. J., Moore, R. J., Khan, U., Hapuarachchi, P., Hasan, M. M., Gamage, N., Bari, M.A., & Tuteja, N. K. (2019). Distributed hydrological modelling for forecasting water discharges from the land area draining to the Great Barrier Reef coastline. In Geophysical Research Abstracts. Vol. 21.
11. Herr, A. & Kuhnert, P. M. Assessment of uncertainty in Great Barrier Reef catchment models. *Water Sci. Technol.* **56**(1), 181–188. https://doi.org/10.2166/wst.2007.450 (2007).
12. Kuhnert, P. M. *et al.* Making management decisions in the face of uncertainty: A case study using the Burdekin catchment in the Great Barrier Reef. *Mar. Freshw. Res.* **69**(8), 1187–1200. https://doi.org/10.1071/MF17237 (2018).
13. Montanari, A. *et al.* "Panta Rhei—Everything flows": Change in hydrology and society—the IAHS scientific decade 2013–2022. *Hydrol. Sci. J.* **58**(6), 1256–1275. https://doi.org/10.1080/02626667.2013.809088 (2013).
14. Khan, U. *et al.* Development of catchment water quality models within a realtime status and forecast system for the Great Barrier Reef. *Environ. Modell. Softw.* **132**, 104790. https://doi.org/10.1016/j.envsoft.2020.104790 (2020).
15. Liu, S. *et al.* Characterisation of spatial variability in water quality in the Great Barrier Reef catchments using multivariate statistical analysis. *Mar. Pollut. Bull.* **137**, 137–151. https://doi.org/10.1016/j.marpolbul.2018.10.019 (2018).
16. Oudin, L., Andréassian, V., Perrin, C., Michel, C. & Le Moine, N. Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resour. Res.* https://doi.org/10.1029/2007WR006240 (2008).
17. Guo, Y., Zhang, Y., Zhang, L. & Wang, Z. Regionalization of hydrological modelling for predicting streamflow in ungauged catchments: A comprehensive review. *Wiley Interdiscipl. Rev. Water* **8**(1), e1487. https://doi.org/10.1002/wat2.1487 (2021).
18. Yaseen, Z. M., Sulaiman, S. O., Deo, R. C. & Chau, K.-W. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *J. Hydrol.* **569**, 387–408. https://doi.org/10.1016/j.jhydrol.2018.11.069 (2019).
19. Merz, R., Tarasova, L. & Basso, S. Parameter's controls of distributed catchment models–How much information is in conventional catchment descriptors?. *Water Resour. Res.* https://doi.org/10.1029/2019WR026008 (2020).
20. Maestre, A., El-Sheikh, E., Williamson, D. & Ward, A. A machine learning tool for weighted regressions in time, discharge, and season. *Int. J. Adv. Comput. Sci. Appl.* https://doi.org/10.14569/IJACSA.2014.050314 (2014).
21. Noori, N., Kalin, L. & Isik, S. Water quality prediction using SWAT-ANN coupled approach. *J. Hydrol.* **590**, 125220 (2020).
22. Chen, S. *et al.* Physics-guided machine learning from simulated data with different physical parameters. *Knowl. Inform. Syst.* https://doi.org/10.1007/s10115-023-01864-z (2023).
23. Appling, A. P., Oliver, S. K., Read, J. S., Sadler, J. M. & Zwart, J. *Machine Learning for Understanding Inland Water Quantity, Quality, and Ecology* (Elsevier, 2022). https://doi.org/10.1016/B978-0-12-819166-8.00121-3.
24. Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **10**(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6 (1970).
25. Liu, S. *et al.* A multi-model approach to assessing the impacts of catchment characteristics on spatial water quality in the Great Barrier Reef catchments. *Environ. Pollut.* **288**, 117337. https://doi.org/10.1016/j.envpol.2021.117337 (2021).
26. O'Sullivan, C. M. *et al.* Classification of catchments for nitrogen using Artificial Neural Network Pattern Recognition and spatial data. *Sci. Total Environ.* **809**, 151139. https://doi.org/10.1016/j.scitotenv.2021.151139 (2022).
27. Calvert, J., McTaggart, A., Carvalhais, L. C., Drenth, A. & Shivas, R. Communities of tropical soil fungi differ between burned and unburned forest, with corresponding changes in plant community composition, litter and soil chemistry. *bioRxiv* https://doi.org/10.1101/2021.08.22.457293 (2021).
28. Sepp, S. K. *et al.* Global diversity and distribution of nitrogen-fixing bacteria in the soil. *Front. Plant Sci.* https://doi.org/10.3389/fpls.2023.1100235 (2023).
29. Lintern, A. *et al.* Key factors influencing differences in stream water quality across space. *Wiley Interdiscipl. Rev. Water* **5**(1), e1260. https://doi.org/10.1002/wat2.1260 (2018).
30. O'Sullivan, C. M., Ghahramani, A., Deo, R. C. & Pembleton, K. G. Pattern recognition describing spatio-temporal drivers of catchment classification for water quality. *Sci. Total Environ.* **861**, 160240. https://doi.org/10.1016/j.scitotenv.2022.160240 (2023).

31. Neldner, V. J., Niehus, R. E., Wilson, B. A., McDonald, W. J. F., Ford, A. J., & Accad, A. (2017). The vegetation of Queensland. Descriptions of broad vegetation groups. Version 3.0. Queensland herbarium, Department of Science. Information Technology and Innovation.
32. Wachter, S., Mittelstadt, B. & Russell, C. Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harv. JL Tech.* **31**, 841 (2017).
33. Wu, J. & Lu, J. Spatial scale effects of landscape metrics on stream water quality and their seasonal changes. *Water Res.* **191**, 116811. https://doi.org/10.1016/j.watres.2021.116811 (2021).
34. Liu, S. *et al.* A Bayesian approach to understanding the key factors influencing temporal variability in stream water quality–a case study in the Great Barrier Reef catchments. *Hydrol. Earth Syst. Sci.* **25**(5), 2663–2683. https://doi.org/10.5194/hess-25-2663-2021 (2021).
35. Wang, Y. N., Li, B. & Yang, G. Stream water quality optimized prediction based on human activity intensity and landscape metrics with regional heterogeneity in Taihu Basin, China. *Environ. Sci. Pollut. Res.* **30**(2), 4986–5004. https://doi.org/10.1007/s11356-022-22536-5 (2023).
36. Shapley, L. S. (1953). A value for n-person games. Contributions to the Theory of Games II, Annals of Mathematical Studies, 28 (1953)
37. State of Queensland (Department of Environment and Science) (2018) Great Barrier Reef catchment and river basins http://qldspatial.information.qld.gov.au/catalogue/custom/search.page?q=%22Great%20Barrier%20Reef%20catchment%20and%20river%20basins%22.
38. Australian Bureau of Meteorology. (2013). Special climate statement 43–extreme heat in January 2013.
39. Australian Bureau of Meteorology. (2013). Special climate statement 44–extreme rainfall and flooding in coastal Queensland and New South Wales.
40. Glorot, X., & Bengio, Y. (2010, March). Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics (pp. 249–256). JMLR Workshop and Conference Proceedings.
41. Hirsch, R. M. Large biases in regression-based constituent flux estimates: Causes and diagnostic tools. *JAWRA J. Am. Water Resour. Assoc.* **50**(6), 1401–1424. https://doi.org/10.1111/jawr.12195 (2014).
42. Rahmani, F., Shen, C., Oliver, S., Lawson, K. & Appling, A. Deep learning approaches for improving prediction of daily stream temperature in data-scarce, unmonitored, and dammed basins. *Hydrol. Processes* **35**(11), e14400. https://doi.org/10.1002/hyp.14400 (2021).
43. Saha, G., Rahmani, F., Shen, C., Li, L. & Raj, C. A deep learning-based novel approach to generate continuous daily stream nitrate concentration for nitrate data-sparse watersheds. *Sci. Total Environ.* https://doi.org/10.1016/j.scitotenv.2023.162930 (2023).
44. Souza, J. & Hooke, J. Influence of seasonal vegetation dynamics on hydrological connectivity in tropical drylands. *Hydrol. Processes* **35**(11), e14427. https://doi.org/10.1002/hyp.14427 (2021).
45. Wang, S., Wang, X. & Ouyang, Z. Effects of land use, climate, topography and soil properties on regional soil organic carbon and total nitrogen in the Upstream Watershed of Miyun Reservoir, North China. *J. Environ. Sci.* **24**(3), 387–395. https://doi.org/10.1016/S1001-0742(11)60789-4 (2012).
46. Burns, D. A. *et al.* Monitoring the riverine pulse: Applying high-frequency nitrate data to advance integrative understanding of biogeochemical and hydrological processes. *Wiley Interdiscipl. Rev. Water* **6**(4), e1348. https://doi.org/10.1002/wat2.1348 (2019).
47. Duncan, J. M., Band, L. E. & Groffman, P. M. Variable nitrate concentration–discharge relationships in a forested watershed. *Hydrol. Processes* **31**(9), 1817–1824. https://doi.org/10.1002/hyp.11136 (2017).
48. Pardo, L. H., Green, M. B., Bailey, S. W., McGuire, K. J. & McDowell, W. H. Identifying controls on nitrate sources and flowpaths in a forested catchment using a hydropedological framework. *J. Geophys. Res. Biogeosci.* **127**(2), e2020JG006140. https://doi.org/10.1029/2020JG006140 (2022).
49. Maggs, J. & Hewett, B. Organic C and nutrients in surface soils from some primary rainforests, derived grasslands and secondary rainforests on the Atherton Tableland in North East Queensland. *Soil Res.* **31**(3), 343–350. https://doi.org/10.1071/SR9930343 (1993).
50. Zheng, M., Xu, M., Li, D., Deng, Q. & Mo, J. Negative responses of terrestrial nitrogen fixation to nitrogen addition weaken across increased soil organic carbon levels. *Sci. Total Environ.* https://doi.org/10.1016/j.scitotenv.2023.162965 (2023).
51. Lintern, A. *et al.* The influence of climate on water chemistry states and dynamics in rivers across Australia. *Hydrol. Processes* **35**(12), e14423. https://doi.org/10.1002/hyp.14423 (2021).
52. Duncan, J. M., Welty, C., Kemper, J. T., Groffman, P. M. & Band, L. E. Dynamics of nitrate concentration-discharge patterns in an urban watershed. *Water Resour. Res.* **53**(8), 7349–7365. https://doi.org/10.1002/2017WR020500 (2017).
53. Todman, L. C., Bush, A. & Hood, A. S. 'Small Data'for big insights in ecology. *Trends Ecol. Evol.* https://doi.org/10.1016/j.tree.2023.01.015 (2023).
54. Wang, S., Peng, H. & Liang, S. Prediction of estuarine water quality using interpretable machine learning approach. *J. Hydrol.* **605**, 127320. https://doi.org/10.1016/j.jhydrol.2021.127320 (2022).
55. Karki, N. *et al.* Comparative performance of regionalization methods for model parameterization in ungauged Himalayan watersheds. *J. Hydrol. Reg. Stud.* **47**, 101359. https://doi.org/10.1016/j.ejrh.2023.101359 (2023).
56. Arora, B. *et al.* Differential CQ analysis: A new approach to inferring lateral transport and hydrologic transients within multiple reaches of a mountainous headwater catchment. *Front. Water* **2**, 24. https://doi.org/10.3389/frwa.2020.00024 (2020).
57. Lovett, G. M. *et al.* Who needs environmental monitoring?. *Front. Ecol. Environ.* **5**(5), 253–260. https://doi.org/10.1890/1540-9295(2007)5[253:WNEM]2.0.CO;2 (2007).
58. Kelleway, J. J. *et al.* Carbon storage in the coastal swamp oak forest wetlands of Australia. *Wetland Carbon Environ. Manag.* https://doi.org/10.1002/9781119639305.ch18 (2021).
59. Courtois, A., Morel, J. M. & Arias, P. Can neural networks extrapolate? Discussion of a theorem by Pedro Domingos. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A Matemáticas* **117**(2), 79. https://doi.org/10.1007/s13398-023-01411-z (2023).
60. Al-Musaylh, M. S., Deo, R. C., Adamowski, J. F. & Li, Y. Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF Reanalysis atmospheric predictors in southeast Queensland, Australia. *Renew. Sustain. Energy Rev.* **113**, 109293. https://doi.org/10.1016/j.rser.2019.109293 (2019).
61. ABARES. (2016). The Australian Land Use and Management Classification Version 8.
62. Lintern, A. *et al.* What are the key catchment characteristics affecting spatial differences in riverine water quality?. *Water Res. Res.* **54**(10), 7252–7272. https://doi.org/10.1029/2017WR022172 (2018).
63. The MathWorks Inc., 2020. MATLAB version 2020a [Deep Learning Toolbox]. The MathWorks Inc. www.mathworks.com.
64. Salari, M., Teymouri, E. & Nassaj, Z. Application of an artificial neural network model for estimating of water quality parameters in the Karun river, Iran. *J. Environ. Treat. Techn.* **9**(4), 720–727. https://doi.org/10.47277/JETT/9(4)727 (2021).
65. Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics (pp. 315–323). JMLR Workshop and Conference Proceedings.
66. Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10) (pp. 807–814).

67. Uzair, M., & Jamil, N. (2020, November). Effects of hidden layers on the efficiency of neural networks. In 2020 IEEE 23rd international multitopic conference (INMIC) (pp. 1–6). IEEE.
68. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv* https://doi.org/10.48550/arXiv.1412.6980 (2014).
69. Deo, R. C. & Şahin, M. Application of the artificial neural network model for prediction of monthly standardized precipitation and evapotranspiration index using hydrometeorological parameters and climate indices in eastern Australia. *Atmos. Res.* **161**, 65–81. https://doi.org/10.1016/j.atmosres.2015.03.018 (2015).
70. Moriasi, D. N. *et al.* Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **50**(3), 885–900. https://doi.org/10.13031/2013.23153 (2007).
71. Willmott, C. J. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **63**(11), 1309–1313. https://doi.org/10.1175/1520-0477(1982)063%3c1309:SCOTEO%3e2.0.CO;2 (1982).
72. Fox, D. G. Judging air quality model performance: A summary of the AMS workshop on dispersion model performance, woods hole, Mass., 8–11 September 1980. *Bull. Am. Meteorol. Soc.* **62**(5), 599–609. https://doi.org/10.1175/1520-0477(1981)062%3c0599:JAQMP%3e2.0.CO;2 (1981).
73. Deo, R. C., Tiwari, M. K., Adamowski, J. F. & Quilty, J. M. Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stochastic Environ. Res. Risk Assess.* **31**, 1211–1240. https://doi.org/10.1007/s00477-016-1265-z (2017).
74. State of Queensland Department of Natural Resources, Mines and Energy (2018) Surface Water Ambient Network (Water Quality) 2018–19, WMP014 version 6, June 2018.

Acknowledgements

Special thanks to Chris Davey at UniSQ's Advanced Data Analytics Research group for invaluable assistance with coding suited to infrastructure changes in High Performance Computing (HPC) Environments. Also Richard Young for ongoing support and operation of the UniSQ HPC facility. The authors are grateful to the University of Southern Queensland and the Australian Federal Government for the Research Training Program Scholarship and State of Queensland Government supported Queensland Water Quality Modelling Network initiative for top-up funding and motivation. Also, The Bureau of Meteorology for the provision of necessary data, Dr. Urooj Khan for support, and the numerous State of Queensland Government staff who passionately develop, ground truth, and maintain open access spatial datasets to provide unprecedented knowledge of nature.

Author contributions

C.M.O'S. and A.G. conceptualised the study and methodological approach. C.M.O'S. developed and applied the methodology, processed the data and simulations, analysed and interpreted the results, wrote the original manuscript. A.G. supervised the research. R.D. supervised writing the manuscript. All authors have reviewed, edited and approved submission of this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-45259-0.

Correspondence and requests for materials should be addressed to C.M.O.

5.3. Links and implications

This penultimate chapter applied the findings of the foundational proof of concept, and method refinement studies presented in paper 1 and 2, respectively, to classify ungauged catchments to gauged ones using only spatial data. Implications of this concluding study are that XAI-SHAP enhanced identifying the potential application and limitations for ANN-PR classification. The ANN-WQ simulator developed as a case study to identify data transferability between a pseudo ungauged catchment and its gauged classifier achieved satisfactory performance criteria and validated that transferring data between the classified catchments has merit.

The most novel finding of the study was differences between results generated by the ANN-WQ simulator trained on data from all catchments, compared to results where the ANN-WQ simulator was trained on data from catchments that had the closest Spatial Data deviation similarity, i.e. the classified catchment or classified Category. Variability between the performance of the ANN-WQ simulator under different combinations of training data suggests Original Vegetation has a useful purpose for identifying selection of training data to include in the ANN-WQ simulator.

The results demonstrated that training the ANN-WQ simulator using all Land Use, Original Vegetation and Ecounit data variables and flow for all catchments achieve more accurate simulations for unsupervised classification scenarios compared to using only one spatial dataset. However, exploiting knowledge embedded within the Original Vegetation dataset facilitated pre-grouping of training datasets to include catchments only of the same spatio-temporal pattern category and achieve the best overall ANN-WQ simulator performance.

In addition to classifying ungauged catchment to gauged ones, this final paper also demonstrates that any future endeavours to transfer data, particularly in an ANN-WQ simulator context, will be improved where the Original Vegetation data is used to inform the training dataset in some form. Original Vegetation data benefits the ANN-WQ simulator performance the most where it is first used as a proxy to identify the DIN regime category, and secondly it benefits the ANN-WQ simulator via inclusion as EULUOV variables where the DIN regime category is not first established. In conclusion, this paper demonstrated Original Vegetation data is a proxy for DIN patterns, is useful for classifying ungauged catchments, and is also useful for improving ANN-WQ simulations for DIN.

CHAPTER 6: DISCUSSION AND CONCLUSIONS

6. Foreword

This penultimate chapter synthesises the research findings by way of outlining the objectives delivered. It summarises the novel contributions of the study, and concludes with an evaluation of limitations and future opportunities associated with this doctoral thesis.

6.1. Synthesis of research findings

The findings of this doctoral thesis have contributed new knowledge in areas of water science and artificial intelligence methods that underpins water quality modelling for simulations of DIN in ungauged catchment areas. Catchment scale water quality modelling methods for simulating DIN in ungauged areas have been adapted from state-of-the-art methods designed to simulate flows and total suspended solids. These methods are reliant on catchment classification developed and designed from first principles for hydrological modelling of flows. While effective for simulating flows in ungauged areas, performance inconsistencies are reported where the methods are used for nutrient simulations in ungauged areas. This inconsistency in performance is also observed where land use, the identified driver of nutrients in CSWQMs, are homogenous. While flow and total suspended solids are driven by abiotic physical catchment attributes with linear relationships, the drivers of the Dissolved Inorganic Nitrogen cycle include a number of biotic drivers with non-linear relationships towards the catchment response for DIN and limits the suitability of existing classification and data transfer methods, for ungauged areas. This doctoral thesis identified datasets and evaluation methods that can overcome the non-linearity constraint associated with existing classification methods.

The methods and datasets explored include coupling Artificial Intelligence which has ability to overcome non-linearity, with datasets that represent biotic response to combinations of landscape drivers. In Queensland, Australia, Original Vegetation communities are mapped and are considered in this research as a parsimonious indicator of differing combinations of catchment features such as aspect, geology, water variability etc., that affect productivity at the different locations across the landscape, even after the original vegetation is removed. Prior to this doctoral thesis, the non-linear relationships between the heterogeneity in the landscape as shown by Original Vegetation data, and the catchment response for DIN have not been explored. Artificial Intelligence methods of Artificial Neural Networks and eXplainable artificial intelligence approaches were successfully exploited in this research to expose nonlinear and inconsistent patterns in datasets that corroborate with knowledge contained in Original Vegetation mapping and the catchment response for DIN.

6.1.1. Objective 1 – ANN-PR corroborates Inductive and Deductive Classification

In Chapter 3 as part of Objective 1, a new classification approach to identify proxy drivers of DIN patterns throughout the study area was developed. Novel aspects of the new classification method included development of an Artificial Neural Network Pattern Recognition method (ANN-PR) for deductive catchment classification for DIN. The ANN-PR method coupled inductive classification of gauged catchments sharing water quality patterns, with deductive classification of spatial mapping that is ubiquitously available across the entire study. Together the inductive and deductive classification results were corroborated to demonstrate the suitability of each of the three spatial datasets as a proxy catchment classifier for situations water quality data was unavailable. Spatial data evaluated for classification included similarities across catchments using Land Use mapping data, and separately classify catchments using similarities in Original Vegetation datasets. The results from the inductive classification of catchments using observed DIN patterns was then used to deduce whether or not the classification of catchments using Land Use, Original Vegetation or the combination referred to as Ecounits had merit. The results demonstrated while Land Use and Original Vegetation were both a suitably proxy dataset, Ecounits, achieved classification in areas that were unable to be classified using Land Use data on its own.

6.1.2. Objective 2 - Original Vegetation indicator for Spatio Temporal patterns

Development of the ANN-PR classification approach exposed a notable number of water quality records in each catchment that did not share water quality patterns with the classified catchment. While this finding exposed a potential source for inconsistent performance observed in CSWQMs that transfer data between classified catchments, the ANN-PR approach in isolation did not provide explainability for the results. To better understand the cause of inconsistency observed in the ANN-PR classification results, Chapter 4, Objective 2 evaluated the training datasets in new ways. First, for inductive classification, DIN training datasets were separated into records for above vs below average flows, as well as wet vs dry season. This separation was based on prior knowledge of the influence of water availability and seasonality on biotic process drivers that affect productivity, as well as flows and time being the only other variables used in the ANN-PR inductive classification training datasets. Applying separated datasets to the ANN-PR method demonstrated that DIN patterns observed in the gauged catchments of the Great Barrier Reef are not always consistent and instead allocate to 1 of 3 DIN regime category scenarios. Scenarios where inductively classified catchments share water quality record patterns matched by ANN-PR methods are: Category 1- increasing flows and wet season with no pattern similarity during other times, Category 2- DIN patterns detectable at all times, or Category 3- in retreating flows but the patterns were not shared other times. This finding confirms classification results were influenced by seasonal and flow drivers that affect biotic processes and demonstrates benefit of pre-applying process knowledge to training data composition in ANN-PR classification methods.

Secondly, Chapter 4, Objective 2 applied knowledge that game theory can provide explainability to results generated by Artificial Neural Networks. Using an adaption of the XAI SHAP, deviation of catchment features were able to be evaluated and corroborated to the deductive classification results. Interrogation of the deviations found that the three distinct categories of DIN patterns identified for inductive classification also shared a distinctively deviated Original Vegetation data variables. The results found catchments with the largest deviations of Original Vegetation variables for Woodlands and Open Woodlands corroborated with Category 1 inductive classification, where DIN patterns consistently match to the inductively classified catchment during the wet season, increasing flows respectively. Catchments allocated to Category 3 for DIN patterns matching to the inductively classified catchment only during dry season and retreating flows instead shared Original Vegetation dataset deviations for Open Forest variables. In contrast catchments deductively classified and sharing deviations in vineforest variables were allocated as Category 2 catchments, whereby water quality patterns were always shared with inductively classified catchments. This finding for Objective 2 is the first time that proxy drivers for catchments with varying water quality patterns throughout the spatio-temporal scale has been identified for Great Barrier Reef Catchments.

6.1.3. Objective 3 - Simulation of DIN in pseudo-ungauged areas is informed by spatial data classification

Finally, Chapter 5, Objective 3 applied the method developed in Objective 2 to deductively classify all ungauged to the gauged catchments that flow to the Great Barrier Reef. While ANN-PR was able to match all ungauged catchments to gauged counterparts, only approximately half of the ungauged catchments demonstrated Original Vegetation variable deviations consistent with Category 1, 2 or 3 catchments evaluated in Objective 2. This finding demonstrated that only catchments with spatial dataset deviations consistent with the Category 1, 2 or 3 DIN regime have an explanation for consideration as a proxy dataset for the driver of DIN.

The suitability of the classification approach towards informing data transfer from gauged to classified ungauged catchments was validated via a case study trial on a nominated pseudo-ungauged catchment. The pseudo-ungauged catchment water quality and spatial data was omitted from method development, however had sufficient data to validate the research and establish that that Original Vegetation variable deviations were consistent with a Category 1 catchment. The case study trial involved development of an ANN-WQ simulator trained using spatial data for the gauged catchments to predict DIN, and then tested in the unsupervised environment to predict DIN for the pseudo ungauged catchment, using corresponding spatial data only. Performance of the ANN-simulator improved where the training datasets were first segregated to only include water quality records from increasing flow consistent with the Category 1 allocation for the catchment. This finding demonstrates that while catchment classification for DIN in ungauged areas is possible, the shared data patterns are not consistent across all temporal or flow scales. The research demonstrates that the performance of water quality simulation models improve where the model is designed to recognise the temporal scale relevant for the classified catchment data, and therefore overcome heterogeneity in dataset patterns. This finding is consistent with other research that found neural network performance improves where training data is refined (Alshemali et al., 2020; Kavzoglu, 2009), and demonstrates the importance of the methods developed within this research to overcome non-linearity and heterogeneity in dataset patterns to improve simulation capacity for DIN.

Despite all catchments being matched to a gauged catchment using ANN-PR, coupling the classification results with the XAI-SHAP approach demonstrated only approximately half the ungauged catchments had Original Vegetation variable deviations consistent with gauged catchments. The other 20 catchments instead displayed deviations of coastal Original Vegetation variables that were not present in the gauged catchments. Because the gauged catchments do not contain notable deviations of coastal Original Vegetation variables, the relevant temporal scale of DIN patterns for catchments with dominant deviations of coastal original vegetation were unable to be established in this PhD doctoral thesis. For the purpose of informing investment prioritisation, the findings of Objective 3 therefore support for provision of new water quality modelling and gauging station investments in catchments that do not share deviated Original Vegetation variables with any gauged catchments. Investment to build knowledge on DIN patterns in the ungauged catchments dominated by coastal Original Vegetation variable types can then establish the relevant temporal scale of water quality records to use in any data transfer studies for water quality simulations in those areas.

6.2. Novel contributions of the study

In summary the novel contributions to knowledge from this PhD research are:

1. The first contribution is the finding that catchments suitable to classify for flow pattern similarities do not necessarily also share DIN pattern similarities. ANN-

PR using a Leave One Out method demonstrated a significant difference between catchment records matched together for flow vs DIN patterns and vindicated the hypothesis that the suitability of catchment classification differs for flow compared to DIN (Hypothesis 1, Objective 1);

2. The second contribution was that ANN-PR (Artificial Neural Network Pattern Recognition) methods facilitated for comparison of classification results using continuous water quality data used for inductive classification with the static spatial datasets used for deductive classification. This approach found that Original Vegetation spatial datasets coupled with Land Use datasets were a suitable proxy for deductively classifying catchments together for sharing Dissolved Inorganic Nitrogen patterns, and provides an approach to classify ungauged catchments to gauged counterparts (Hypothesis 1, Objective 1);

3. the third novel contribution of this doctoral thesis research was that splitting ANN-PR training datasets based on knowledge of biological changes over differing temporal scale or flows informs the most suitable catchments to classify together in response to those changes (Hypothesis 2, Objective 2).

4. Fourth novel contribution was that SHAP-XAI evaluation of Original Vegetation data revealed vegetation types that are a proxy indicator for the spatio-temporal period of DIN data most suitable for classification (Hypothesis 2, Objective 2).

5. Fifth novel contribution was that catchments with similar XAI-SHAP deviations in vineforest share consistent water quality patterns, catchments sharing deviations in Open Forest vegetation type share water quality patterns during dry season and retreating flows only, catchments sharing deviations in woodlands or open woodlands share data during wet season and increasing flows only (Hypothesis 2, Objective 2).

6. The sixth novel contribution is that development of an ANN-WQ simulator trained on spatial and flow data to predict DIN for classified catchments found that the predictive performance of the ANN-WQ simulator improved where the training datasets were first separated to include data records from the temporal

scale that matched to the deviated Original Vegetation variable category for the catchment (Objective 3).

7. The penultimate novel contribution of this research is that the 11 gauged catchments evaluated in this research and flowing to the Great Barrier Reef contain sufficient data to justify catchment classification with over 50% of the ungauged catchments that also share similar deviations in Original Vegetation variables (Hypothesis 3, Objective 3).

Overall, this research demonstrated that Original Vegetation datasets contain knowledge that can be a proxy dataset for classification in CSWQMs. Additionally, the knowledge within the Original Vegetation dataset provides insight to the spatio temporal suitability of the data transfer where observed water quality data is unavailable. Use of the Original Vegetation dataset coupled with XAI-SHAP exposes deviations of vegetation composition in the catchment to confirm whether classification recommended by ANN-PR is justified.

6.3. Limitations and corresponding future opportunities

Overall, this PhD doctoral research developed methods that inform classification of ungauged areas to inform data transfer for DIN simulations in CSWQMs. However, it is prudent that limitations of the approach are duly considered prior to direct application of these findings to CSWQMs that inform land management decisions. Opportunities to use limitations of this research, as a foundation for further knowledge development are:

1. While the method evaluated in this study was specifically designed for the influence of biotic drivers in the landscape towards nutrient patterns, the relationships was only explored for DIN. Further opportunity remains to repeat methodology established herein to explore whether the classification approach is also applicable to other forms of nitrogen and phosphorous to more completely inform CSWQMs.

2. This research applied Artificial Neural Network (ANN) methods to identify whether spatial datasets are a suitable proxy for classifying catchments that also shared DIN patterns and simulate those patterns. While the method achieved the objective of this doctoral research, methods that overcome inconsistencies of water quality patterns over all temporal scales are needed to increase the certainty of classification at all times. The research found water quality patterns in catchments with Original Vegetation variables deviated for woodlands and open forest were not established outside the relevant DIN regime. For these catchments, alternative deep learning methods may have potential to further optimise pattern detection for the temporal scales that the ANN-PR and ANN-WQ simulator didn't.

3. While this research found that splitting ANN-PR training datasets based on knowledge of biological changes over differing temporal scale or flows informs the most suitable catchments to classify, those results have only been verified using the ANN-WQ simulator for Category 1 catchments with dominant deviations in Woodland and Open Woodland original vegetation types. Additional case studies are recommended to a) apply the classification results to inform parameters transfer in established process based or data driven CSWQMs and evaluate and performance differences; and b) further explore the applicability of the method for Category 2 and 3 catchments.

4. XAI-SHAP method was applied manually to facilitate timely interrogation of each spatial dataset variable and corroborate with the water quality records. A number of integrated XAI approaches are now emerging for simultaneous XAI evaluations of Artificial Intelligence and integration of those approaches into the application of this work to other areas could facilitate for the drivers of the ANN-PR classification to be identified automatically, and simultaneously expose any hyperparameter influence towards the ANN-PR result (Hedström et al., 2023; Pahde et al., 2023).

5. By virtue of historical data availability, the methodology was limited to exploring relationships between non-continuous samples of DIN and spatial mapping in the gauged catchments only. Inclusion of XAI to the classification method

demonstrated that the explainability for classification results was not extendable to all of the ungauged areas. The reason was approximately half the ungauged areas had original vegetation deviations that were not detected in the gauged catchments. Opportunities therefore exist to prioritise investment for new gauging stations and monitoring points at locations where the temporal patterns for DIN as well as other water quality constituents can be established for the ungauged areas with original vegetation types dominated by Coastal ecosystems.

REFERENCES

ABARES (2016), The Australian land use and Management Classification Version 8, *Australian Bureau of Agricultural and Resource Economics and Sciences*, Canberra. CC BY 3.0.

Aerts, J. P., Hoch, J. M., Coxon, G., van de Giesen, N. C., & Hut, R. W. (2023). On the importance of observation uncertainty when evaluating and comparing models: a hydrological example. *EGUsphere*, 2023, 1-26. https://doi.org/10.5194/egusphere-2023-1156

Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J., & Capell, R. (2012). A comparison of similarity indices for catchment classification using a cross-regional dataset. *Advances in Water Resources*, 40, 11-22. https://doi.org/10.1016/j.advwatres.2012.01.008

Alnahit, A. O., Mishra, A. K., & Khan, A. A. (2022). Stream water quality prediction using boosted regression tree and random forest models. *Stochastic Environmental Research and Risk Assessment*, 36(9), 2661-2680. https://doi.org/10.1007/s00477-021-02152-4

Andréassian, V., Le Moine, N., Perrin, C., Ramos, M. H., Oudin, L., Mathevet, T., Berthet, L. (2012). All that glitters is not gold: the case of calibrating hydrological models. *Hydrological Processes*, 26(14), 2206-2210. https://doi.org/10.1002/hyp.9264

Anonymous. (2003). Reef Water Quality Protection Plan; for catchments adjacent to the Great Barrier Reef World Heritage Area. *Queensland Department of Premier and Cabinet*, Brisbane

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, 58, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Alshemali, B., & Kalita, J. (2020). Improving the reliability of deep neural networks in NLP: *A review. Knowledge-Based Systems*, 191, 105210. https://doi.org/10.1016/j.knosys.2019.105210

Australian and Queensland Governments, 2022, Catchment loads monitoring methods, Reef Water Quality Report Card 2020, *State of Queensland*, Brisbane.

Ayana, E. K., Worqlul, A. W., & Steenhuis, T. S. (2015). Evaluation of stream water quality data generated from MODIS images in modeling total suspended solid emission to a freshwater lake. *Science of the Total Environment*, 523, 170-177. doi:10.1016/j.scitotenv.2015.03.132 https://doi.org/10.1016/j.scitotenv.2015.03.132

Baird, M. E., Mongin, M., Skerratt, J., Margvelashvili, N., Tickell, S., Steven, A. D., Robillot, C., Ellis, R., Waters, D., Kaaniewska, P., & Brodie, J. (2021). Impact of catchment-derived nutrients and sediments on marine water quality on the Great Barrier Reef: An application of the eReefs marine modelling system. *Marine Pollution Bulletin*, 167, 112297. https://doi.org/10.1016/j.marpolbul.2021.112297

Baker, J. (2003). A Report On The Study Of Landsourced Pollutants and Their Impacts On Water Quality In And Adjacent to The Great Barrier Reef. Retrieved from Brisbane: https://www.reefplan.qld.gov.au/__data/assets/pdf_file/0026/46169/report-impact-of-land-pollutants-on-gbr.pdf

Bartley, R., Waters, D., Turner, R., Kroon, F., Wilkinson, S., Garzon-Garcia, A., Waterhouse, J. (2017). *Scientific Consensus Statement 2017: A synthesis of the science of land-based water quality impacts on the Great Barrier Reef, Chapter 2: Sources of sediment, nutrients, pesticides and other pollutants to the Great Barrier Reef.* Report Section. State of Queensland

Bell, V., Kay, A., Jones, R., & Moore, R. (2007). Development of a high resolution grid-based river flow model for use with regional climate model output*. Hydrology and Earth System Sciences Discussions*, 11(1), 532-549. https://doi.org/10.5194/hess-11-532-2007

Booker, D. J., & Woods, R. A. (2014). Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. *Journal of Hydrology*, 508, 227-239. https://doi.org/10.1016/j.jhydrol.2013.11.007

Brassington, G. B., Tuteja, N., Colberg, F., Sandery, P., Sakov, P., Sakova, I., Bari, M. (2017). eReefs: An integrated catchment and coastal forecasting system for the Great Barrier Reef and Queensland coast. Australasian Coasts & Ports 2017: *Working with Nature*, 166.

Buzacott, A. J., Tran, B., van Ogtrop, F. F., & Vervoort, R. W. (2019). Conceptual models and calibration performance—investigating catchment bias. *Water*, 11(11), 2424. https://doi.org/10.3390/w11112424

Carter, E. & Thulstrup, H., 2022. *Report on the joint world heritage centre/iucn reactive monitoring mission to the Great Barrier Reef (Australia)*, UNESCO. Switzerland. Retrieved from https://policycommons.net/artifacts/3178903/australia-gbr-rm-mission-whc-iucn-30mar22/3977422/ on 09 Oct 2023. CID: 20.500.12592/frzbpp.

Chen, Y., Minchin, S. A., S. Seaton , K.D. Joehnk , Robson, B. J., & Bai, Q. (2011). eReefs – a new perspective on the Great Barrier Reef Paper presented at the *19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011* Perth Australia.

Cheng, P., Meng, F., Wang, Y., Zhang, L., Yang, Q., & Jiang, M. (2018). The impacts of land use patterns on water quality in a trans-boundary river basin in northeast China based on eco-functional regionalization. *International Journal of Environmental Research and Public Health*, 15(9), 1872. https://doi.org/10.3390/ijerph15091872

Cohen, S., Dror, G., & Ruppin, E. (2007). Feature selection via coalitional game theory. *Neural Computation*, 19(7), 1939-1961. https://doi.org/10.1162/neco.2007.19.7.1939

Cole, S., Moore, R., Bell, V., & Jones, D. (2006). Issues in flood forecasting: ungauged basins, extreme floods and uncertainty. Paper presented at the *Frontiers in Flood Forecasting, 8th Kovacs Colloquium*, UNESCO, Paris.

Creighton, C., Waterhouse, J., Day, J. C., & Brodie, J. (2021). Criteria for effective regional scale catchment to reef management: A case study of Australia's Great Barrier Reef. *Marine Pollution Bulletin*, 173, 112882. https://doi.org/10.1016/j.marpolbul.2021.112882

Dadson, S. J., Hirpa, F., Thomson, P., & Konar, M. (2019). Monitoring and Modelling Hydrological Processes. *Water Science, Policy, and Management: A Global Challenge*, 117-137. https://doi.org/10.1002/9781119520627.ch7

DeLancey, E. R., Simms, J. F., Mahdianpari, M., Brisco, B., Mahoney, C., & Kariyeva, J. (2020). Comparing Deep Learning and Willow Learning for Large-Scale Wetland Classification in Alberta, Canada. *Remote Sensing*, 12(1), 2. https://doi.org/10.3390/rs12010002

Di Baldassarre, G., Sivapalan, M., Rusca, M., Cudennec, C., Garcia, M., Kreibich, H., Konar, M., Mondino, E., Mard, J., Pande, S., Sanderson, M.R., Tian, F., Viglione, A., Wei, J., Yu, D., Srinivasan, V., Blöschl, G. (2019). Sociohydrology: Scientific challenges in addressing the sustainable development goals. *Water Resources Research*, 55(8), 6327-6355. https://doi.org/10.1029/2018WR023901

Di Prinzio, M., Castellarin, A., & Toth, E. (2011). Data-driven catchment classification: application to the pub problem. *Hydrology and Earth System Sciences*, 15(6), 1921-1935. https://doi.org/10.5194/hess-15-1921-2011

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018). The Global Streamflow Indices and Metadata Archive (GSIM)–Part 1: The production of a daily streamflow archive and metadata. *Earth System Science Data*, 10(2), 765-785. https://doi.org/10.5194/essd-10-765-2018

Dupas, R., Curie, F., Gascuel-Odoux, C., Moatar, F., Delmas, M., Parnaudeau, V., & Durand, P. (2013). Assessing N emissions in surface water at the national level: Comparison of country-wide vs. regionalized models. *Science of the Total Environment*, 443, 152-162. doi:10.1016/j.scitotenv.2012.10.011

Ebeling, P., Dupas, R., Abbott, B., Kumar, R., Ehrhardt, S., Fleckenstein, J. H., & Musolff, A. (2021). Long-Term Nitrate Trajectories Vary by Season in Western European Catchments. *Global Biogeochemical Cycles*, 35(9). https://doi.org/10.1029/2021GB007050

Fatichi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., Tarboton, D. (2016). An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, 537, 45-60. doi:https://doi.org/10.1016/j.jhydrol.2016.03.026

Furnas, M. (2003). Catchments and corals : terrestrial runoff to the Great Barrier Reef, Science Communication, *Australian Institute of Marine Science*. Townsville.

Fu, B., Merritt, W. S., Croke, B. F., Weber, T., & Jakeman, A. J. (2019). A review of catchment-scale water quality and erosion models and a synthesis of future prospects. *Environmental modelling & software*, 114, 75-97. https://doi.org/10.1016/j.envsoft.2018.12.008

Gao, W., Fan, C., Zhang, W., Li, N., Liu, H., & Chen, M. (2023). Heterotrophic nitrification of organic nitrogen in soils: process, regulation, and ecological significance. *Biology and Fertility of Soils*, 59(3), 261-274. https://doi.org/10.1007/s00374-023-01707-7

Gazzaz, N. M., Yusoff, M. K., Ramli, M. F., Juahir, H., & Aris, A. Z. (2015). Artificial neural network modeling of the water quality index using land use areas as predictors. *Water Environment Research*, 87(3), 99-112. doi:10.2175/106143014X14062131179276

GBRMPA. (2001). Authority, Great Barrier Reef Marine Park. "Great Barrier Reef catchment water quality action plan: a report to Ministerial Council on targets for pollutant loads." Townsville: *Great Barrier Reef Marine Park Authority* (2001). Retrieved from: https://elibrary.gbrmpa.gov.au/jspui/handle/11017/357

Giesbrecht, I. J., Tank, S. E., Frazer, G. W., Hood, E., Gonzalez Arriola, S. G., Butman, D. E., D'Aore, D. V., Hutchinson, D., Bidlack, A., & Lertzman, K. P. (2022). Watershed classification predicts streamflow regime and organic carbon dynamics in the Northeast Pacific Coastal Temperate Rainforest. *Global Biogeochemical Cycles*, 36(2), e2021GB007047. https://doi.org/10.1029/2021GB007047

Goodwell, A. E., Jiang, P., Ruddell, B. L., & Kumar, P. (2020). Debates—Does Information Theory Provide a New Paradigm for Earth Science? Causality, Interaction, and Feedback. *Water Resources Research*, 56(2), https://doi.org/10.1029/2019WR024940.

Goodarzi, M. S., & Navardi, S. (2019). A Brief Survey of Regionalization Modeling; A Bonus from Gauged to Ungauged Basins. *Adv Civil Eng Tech. 3(4).* http://dx.doi.org/10.31031/ACET.2019.03.000566

Hallouin, T., Bruen, M., & O'Loughlin, F. E. (2020). Calibration of hydrological models for ecologically relevant streamflow predictions: a trade-off between fitting well to data and estimating consistent parameter sets?. *Hydrology and Earth System Sciences*, 24(3), 1031-1054. https://doi.org/10.5194/hess-24-1031-2020

Hameed, M., Sharqi, S. S., Yaseen, Z. M., Afan, H. A., Hussain, A., & Elshafie, A. (2017). Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Computing and Applications*, 28, 893-905. doi:10.1007/s00521-016-2404-7

Harris, G. P. (2012). Introduction to the special issue:‘Achieving ecological outcomes’. Why is translational ecology so difficult? *Freshwater Biology*, 57, 1-6. https://doi.org/10.1111/j.1365-2427.2012.02773.x

Hedström, A., Bommer, P., Wickstrøm, K. K., Samek, W., Lapuschkin, S., & Höhne, M. M. C. (2023). The Meta-Evaluation Problem in Explainable AI: *Identifying Reliable Estimators with MetaQuantus*. arXiv preprint arXiv:2302.07265.

Hrachowitz, M., Savenije, H., Blöschl, G., McDonnell, J., Sivapalan, M., Pomeroy, J., Ehret, U. (2013). A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal*, 58(6), 1198-1255. https://doi.org/10.1080/02626667.2013.803183

Husic, A., Fox, J. F., Clare, E., Mahoney, T., & Zarnaghsh, A. (2023). Nitrate hysteresis as a tool for revealing storm-event dynamics and improving water quality model performance. *Water Resources Research*, https://doi.org/10.1029/2022WR033180.

Huang, H., Ji, X., Xia, F., Huang, S., Shang, X., Chen, H., Zhang, M., Dahlgren R.A., Mei, K. & Mei, K. (2020). Multivariate adaptive regression splines for estimating riverine constituent concentrations. *Hydrological Processes*, 34(5), 1213-1227. https://doi.org/10.1002/hyp.13669

Ighalo, J. O., Adeniyi, A. G., & Marques, G. (2021). Artificial intelligence for surface water quality monitoring and assessment: A systematic literature analysis. *Modeling Earth Systems and Environment*, 7(2), 669-681. https://doi.org/10.1007/s40808-020-01041-z

Jay, V. A., & Neumann, M. (2021, December). Site productivity indices for native forests in southeast Queensland. In *IOP Conference Series: Earth and Environmental Science* (Vol. 932, No. 1, p. 012006). IOP Publishing.

Jehn, F. U., Bestian, K., Breuer, L., Kraft, P., & Houska, T. (2020). Using hydrological and climatic catchment clusters to explore drivers of catchment behavior. *Hydrology and Earth System Sciences*, 24(3), 1081-1100. https://doi.org/10.5194/hess-24-1081-2020

Jin, T., Cai, S., Jiang, D., & Liu, J. (2019). A data-driven model for real-time water quality prediction and early warning by an integration method. *Environmental Science and Pollution Research*. doi:10.1007/s11356-019-06049-2

Jones, M. W., Peters, G. P., Gasser, T., Andrew, R. M., Schwingshackl, C., Gütschow, J., Houghton, R. A., Friedlingstein, P., Pongratz, J. & Le Quéré, C. (2023). National contributions to climate change due to historical emissions of carbon dioxide, methane, and nitrous oxide since 1850. Scientific Data, 10(1), 155.

Kavzoglu, T. (2009). Increasing the accuracy of neural network classification using refined training data. *Environmental Modelling & Software*, 24(7), 850-858. https://doi.org/10.1016/j.envsoft.2008.11.012

Khalil, B., & Adamowski, J. (2014). Comparison of OLS, ANN, KTRL, KTRL2, RLOC, and MOVE as record-extension techniques for water quality variables. *Water, Air, and Soil Pollution*, 225(6). doi:10.1007/s11270-014-1966-1

Khalil, B., Adamowski, J., Abdin, A., & Elsaadi, A. (2019). A statistical approach for the estimation of water quality characteristics of ungauged streams/watersheds under stationary conditions. *Journal of Hydrology*, 569, 106-116. doi:10.1016/j.jhydrol.2018.11.041

Khalil, B., Ouarda, T. B., & St-Hilaire, A. (2011). Estimation of water quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis. *Journal of Hydrology*, 405(3-4), 277-287. https://doi.org/10.1016/j.jhydrol.2011.05.024

Khan, U., Cook, F. J., Laugesen, R., Hasan, M. M., Plastow, K., Amirthanathan, G. E., Bari, N. K., & Tuteja, N. K. (2020). Development of catchment water quality models within a realtime status and forecast system for the Great Barrier Reef. *Environmental Modelling & Software*, 132, 104790. https://doi.org/10.1016/j.envsoft.2020.104790

Khan, U., Wells, S.C., Hapuarachchi, P., Cole, S.J., Hasan, M.M., Laugesen, R., Moore, R.J., Bari, M.A. and Tuteja, N.K. (2019) Streamflow simulations for gauged and ungauged catchments of the Great Barrier Reef region using a distributed hydrological model, 23rd International Congress on Modelling and Simulation (*MODSIM2019*), Canberra, Australia, 1-6 December 2019

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts *Big Data & Society*, April-June pp1-12. https://doi.org/10.1177/2053951714528481

Kreibich, H., Schröter, K., Di Baldassarre, G., Van Loon, A., Mazzoleni, M., Abeshu, G. W., Agafonova, S., AghaKouchak, A., Aksoy, H., Alvarez-Garreton, C., Aznar, B., Balkhi, L., Barendrecht, M. H., Biancamaria, S., Bos-Burgering, L., Bradley, C., Budiyono, Y., Buytaert, W., Capewell, L., ... Ward, P. J. (2023). Panta Rhei benchmark dataset: socio-hydrological data of paired events of floods and droughts. *Earth System Science Data Discussions*, 2023, 1-27. https://doi.org/10.5194/essd-15-2009-2023

Kroon, F. J., Kuhnert, P. M., Henderson, B. L., Wilkinson, S. N., Kinsey-Henderson, A., Abbott, B., Brodie, J. E., & Turner, R. D. (2012). River loads of suspended solids, nitrogen, phosphorus and herbicides delivered to the Great Barrier Reef lagoon. *Marine pollution bulletin*, 65(4-9), 167-181. https://doi.org/10.1016/j.marpolbul.2011.10.018

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S., & Nearing, G. S. (2019). Toward improved predictions in ungauged basins: Exploiting the power of machine learning. *Water Resources Research*, 55(12), 11344-11354. https://doi.org/10.1029/2019WR026065

Kuentz, A., Arheimer, B., Hundecha, Y., & Wagener, T. (2017). Understanding hydrologic variability across Europe through catchment classification. *Hydrology and Earth System Sciences*, 21(6), 2863-2879. https://doi.org/10.5194/hess-21-2863-2017

Li, H., Zhao, B., Wang, D., Zhang, K., Tan, X., & Zhang, Q. (2023). Effect of multiple spatial scale characterization of land use on water quality. *Environmental Science and Pollution Research*, 30(3), 7106-7120. https://doi.org/10.1007/s11356-022-22720-7

Lintern, A., Liu, S., Minaudo, C., Dupas, R., Guo, D., Zhang, K., Bende-Michl, U., & Duvert, C. (2021). The influence of climate on water chemistry states and dynamics in rivers across Australia. *Hydrological Processes*, 35(12). https://doi.org/10.1002/hyp.14423

Lintern, A., Webb, J. A., Ryu, D., Liu, S., Bende-Michl, U., Waters, D., Leahy, P., Wilson, A.W., & Western, A. W. (2018). Key factors influencing differences in stream water quality across space. Wiley Interdisciplinary Reviews: *Water*, 5(1), https://doi.org/10.1002/wat2.1260.

Liu, S., Ryu, D., Webb, J. A., Lintern, A., Guo, D., Waters, D., & Western, A. W. (2021). A multi-model approach to assessing the impacts of catchment characteristics on spatial water quality in the Great Barrier Reef catchments. *Environmental Pollution*, 288, 117337. https://doi.org/10.1016/j.envpol.2021.117337

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30. In 31st annual conference on neural information processing systems (NIPS 2017), Long Beach, California, USA.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature machine intelligence*, 2(1), 56-67. https://doi.org/10.48550/arXiv.1905.04610

Lu, Y., Wang, L., Zhu, C., Zou, L., Zhang, M., Feng, L., & Cao, Q. (2023). Predicting surface solar radiation using a hybrid radiative Transfer–Machine learning model. *Renewable and Sustainable Energy Reviews*, 173, 113105. https://doi.org/10.1016/j.rser.2022.113105

McCloskey, G. L., Baheerathan, R., Dougall, C., Ellis, R., Bennett, F. R., Waters, D., Darr, S., Fentie, B., Hateley L.R & Askildsen, M. (2021). Modelled estimates of dissolved inorganic nitrogen exported to the Great Barrier Reef lagoon. *Marine Pollution Bulletin*, 171, 112655. https://doi.org/10.1016/j.marpolbul.2021.112655

Merz, R., & Blöschl, G. (2004). Regionalisation of catchment model parameters. *Journal of Hydrology*, 287(1-4), 95-123. https://doi.org/10.1016/j.jhydrol.2003.09.028

Merz, R., Tarasova, L., & Basso, S. (2020). Parameter's controls of distributed catchment models–How much information is in conventional catchment descriptors? *Water Resources Research.* https://doi.org/10.1029/2019WR026008.

McCloskey, G. L., Baheerathan, R., Dougall, C., Ellis, R., Bennett, F. R., Waters, D., Darr, S., Fentie, B., Hateley, L. R., & Askildsen, M. (2021). Modelled estimates of fine sediment and particulate nutrients delivered from the Great Barrier Reef catchments. *Marine pollution bulletin*, 165, 112163. https://doi.org/10.1016/j.marpolbul.2021.112163

McCloskey, G. L., & Waters, D., (2017) Validation and calibration of Source water quality models in the Great Barrier Reef catchments 22nd International Congress on Modelling and Simulation, Hobart, Tasmania, Australia, 3 to 8 December 2017 mssanz.org.au/modsim2017

Moliere, D. R., Lowry, J. B., & Humphrey, C. L. (2009). Classifying the flow regime of data-limited streams in the wet-dry tropical region of Australia. *Journal of Hydrology*, 367(1-2), 1-13. https://doi.org/10.1016/j.jhydrol.2008.12.015

Narbondo, S., Gorgoglione, A., Crisci, M., & Chreties, C. (2020). Enhancing Physical Similarity Approach to Predict Runoff in Ungauged Watersheds in Sub-Tropical Regions. *Water*, 12(2), 528. https://doi.org/10.3390/w12020528

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282-290. https://doi.org/10.1016/0022-1694(70)90255-6

Neldner, V. J., Niehus, R. E., Wilson, B. A., McDonald, W. J. F., Ford, A. J., & Accad, A. (2017). The vegetation of Queensland. Descriptions of broad vegetation groups. Version 3.0. *Queensland herbarium*, Department of Science. Information Technology and Innovation.

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., Prieto, C. & Gupta, H. V. (2021). What role does hydrological science play in the age of machine learning?. *Water Resources Research*, 57(3). https://doi.org/10.1029/2020WR028091.

Newall, P., & Tiller, D. (2002). Derivation of nutrient guidelines for streams in Victoria, Australia. *Environmental Monitoring and Assessment*, 74(1), 85-103. https://doi.org/10.1023/A:1013806805798

Niroula, S., Wallington, K., & Cai, X. (2023). Addressing data challenges in riverine nutrient load modeling of an intensively managed agro-industrial watershed. JAWRA *Journal of the American Water Resources Association*, 59(2), 213-225. https://doi.org/10.1111/1752-1688.13097

Oehler, F., & Elliott, A. H. (2011). Predicting stream N and P concentrations from loads and catchment characteristics at regional scale: A concentration ratio method. *Science of the Total Environment*, 409(24), 5392-5402. doi:10.1016/j.scitotenv.2011.08.025

Olden, J.D., Kennard, M.J. and Pusey, B.J. (2012), A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. *Ecohydrol*., 5: 503-518. https://doi.org/10.1002/eco.251

Oudin, L., Kay, A., Andréassian, V., & Perrin, C. (2010). Are seemingly physically similar catchments truly hydrologically similar?. *Water Resources Research*, 46(11). https://doi.org/10.1029/2009WR008887

Pagliero, L., Bouraoui, F., Diels, J., Willems, P., & McIntyre, N. (2019). Investigating regionalization techniques for large-scale hydrological modelling. *Journal of hydrology*, 570, 220-235. https://doi.org/10.1016/j.jhydrol.2018.12.071

Pahde, F., Yolcu, G. Ü., Binder, A., Samek, W., & Lapuschkin, S. (2023). Optimizing Explanations by Network Canonization and Hyperparameter Search.

In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3818-3827).

Parrott, L. (2010). Measuring ecological complexity. *Ecological Indicators*, 10(6), 1069-1076. https://doi.org/10.1016/j.ecolind.2010.03.014

Park, S. R., & Lee, S. W. (2020). Spatially varying and scale-dependent relationships of land use types with stream water quality*. International Journal of Environmental Research and Public Health*, 17(5), 1673. https://doi.org/10.3390/ijerph17051673

Rafiei, V., Nejadhashemi, A. P., Mushtaq, S., Bailey, R. T., & An-Vo, D. A. (2022). An improved calibration technique to address high dimensionality and non-linearity in integrated groundwater and surface water models. *Environmental Modelling & Software*, 149, 105312. https://doi.org/10.1016/j.envsoft.2022.105312

Saadi, M., Oudin, L., & Ribstein, P. (2019). Random Forest Ability in Regionalizing Hourly Hydrological Model Parameters. Water, 11(8), 1540.
Sharifi, A., Yen, H., Boomer, K. M. B., Kalin, L., Li, X., & Weller, D. E. (2017). Using multiple watershed models to assess the water quality impacts of alternate land development scenarios for a small community. *Catena*, 150, 87-99. doi:10.1016/j.catena.2016.11.009

Singh, S. K., Pahlow, M., Booker, D. J., Shankar, U., & Chamorro, A. (2019). Towards baseflow index characterisation at national scale in New Zealand. *Journal of Hydrology*, 568, 646-657. https://doi.org/10.1016/j.jhydrol.2018.11.025

Sivakumar, B., Singh, V. P., Berndtsson, R., & Khan, S. K. (2015). Catchment classification framework in hydrology: challenges and directions. *Journal of Hydrologic Engineering*, 20(1), A4014002. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000837

Sivapalan, M. (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological processes*, 17(15), 3163-3170. https://doi.org/10.1002/hyp.5155

Sivapalan, M. (2018). From engineering hydrology to Earth system science: milestones in the transformation of hydrologic science. *Hydrology and Earth System Sciences*, 22(3), 1665-1693. https://doi.org/10.5194/hess-22-1665-2018

Snelder, T. H., Larned, S. T., & McDowell, R. W. (2018). Anthropogenic increases of catchment nitrogen and phosphorus loads in New Zealand*. New Zealand Journal of Marine and Freshwater Research*, 52(3), 336-361. https://doi.org/10.1080/00288330.2017.1393758

Soltani-Gerdefaramarzi, S., Gheisouri, M., Saberi, A., & Yarami, N. (2021). The effect of land use change on surface water quality under the wet and dry years in a semi-arid catchment (case study: the Godarkhosh catchment). *Environment, Development and Sustainability*, 23(4), 5371-5385. https://doi.org/10.1007/s10668-020-00820-z

State of Queensland. (2011). Reef Water Quality Protection Plan - First Report 2009 - Baseline. Retrieved from 100 George St, Brisbane Qld 4000: https://www.reefplan.qld.gov.au/__data/assets/pdf_file/0017/46106/methods.pdf

State of Queensland. (2018). Reef 2050 Water Quality Improvement Plan 2017-2020. Retrieved from https://www.reefplan.qld.gov.au/resources/key-documents#2009

State of Queensland Department of Environment and Resource Management (2012) State Surfacewater Ambient Water Quality Network WMP014 version 2.

Steven, A. D., Baird, M. E., Brinkman, R., Car, N. J., Cox, S. J., Herzfeld, M., Margvelashvili, N. (2019). eReefs: An operational information system for managing the Great Barrier Reef. *Journal of Operational Oceanography*, 1-17. https://doi.org/10.1080/1755876X.2019.1650589

Swain, J. B., & Patra, K. C. (2019). Impact of catchment classification on streamflow regionalization in ungauged catchments. *SN Applied Scien*ces, 1, 1-14. https://doi.org/10.1007/s42452-019-0476-6

Tabari, H., & Talaee, P. H. (2015). Reconstruction of river water quality missing data using artificial neural networks. *Water Quality Research Journal of Canada*, 50(4), 326-335. doi:10.2166/wqrjc.2015.044

Tao, H., Bobaker, A. M., Ramal, M. M., Yaseen, Z. M., Hossain, M. S., & Shahid, S. (2019). Determination of biochemical oxygen demand and dissolved oxygen for semi-arid river environment: application of soft computing models. *Environmental Science and Pollution Research*, 26(1), 923-937. doi:10.1007/s11356-018-3663-x

Teutschbein, C., Grabs, T., Laudon, H., Karlsen, R. H., & Bishop, K. (2018). Simulating streamflow in ungauged basins under a changing climate: The importance of landscape characteristics. *Journal of Hydrology*, 561, 160-178. https://doi.org/10.1016/j.jhydrol.2018.03.060

Tocchetti, A., & Brambilla, M. (2022). The role of human knowledge in explainable AI. *Data*, 7(7), 93. https://doi.org/10.3390/data7070093

Toth, E. (2013). Catchment classification based on characterisation of streamflow and precipitation time series. *Hydrology and Earth System Sciences*, 17(3), 1149-1159. https://doi.org/10.5194/hess-17-1149-2013

Trancoso, R., Larsen, J. R., McAlpine, C., McVicar, T. R., & Phinn, S. (2016). Linking the Budyko framework and the Dunne diagram. *Journal of Hydrology*, 535, 581-597. https://doi.org/10.1016/j.jhydrol.2016.02.017

Tung, T. M., & Yaseen, Z. M. (2020). A survey on river water quality modelling using artificial intelligence models: 2000-2020. *Journal of Hydrology*, 124670. https://doi.org/10.1016/j.jhydrol.2020.124670

Turak, E., Harrison, I., Dudgeon, D., Abell, R., Bush, A., Darwall, W., Hermoso, V. (2017). Essential biodiversity variables for measuring change in global freshwater biodiversity. *Biological Conservation*, 213, 272-279. https://doi.org/10.1016/j.biocon.2016.09.005

Tyralis, H., Papacharalampous, G., & Langousis, A. (2019). A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. *Water*, 11(5), 910. doi:ARTN 91010.3390/w11050910

Valizadeh, N., Mirzaei, M., Allawi, M. F., Afan, H. A., Mohd, N. S., Hussain, A., & El-Shafie, A. (2017). Artificial intelligence and geo-statistical models for stream-flow forecasting in ungauged stations: state of the art. *Natural Hazards*, 86(3), 1377-1392. doi:10.1007/s11069-017-2740-7

Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion*, 76, 89-106. https://doi.org/10.1016/j.inffus.2021.05.009

Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment classification and hydrologic similarity. *Geography compass*, 1(4), 901-931. https://doi.org/10.1111/j.1749-8198.2007.00039.x

Wang, S., Peng, H., & Liang, S. (2022). Prediction of estuarine water quality using interpretable machine learning approach. *Journal of Hydrology*, 605. https://doi.org/10.1016/j.jhydrol.2021.127320

Waterhouse, J., Brodie, J., Tracey, D., Smith, R., Vandergragt, M., Collier, C., Baird, M. (2017). Chapter 5: Overview of key findings, management implications and knowledge gaps. Retrieved from Chapter 5
https://www.reefplan.qld.gov.au/__data/assets/pdf_file/0034/45997/2017-scientific-consensus-statement-summary-chap05.pdf

Wells, S. C., Cole, S. J., Moore, R. J., Black, K. B., Khan, U., Hapuarachchi, P., Bari, M. (2017). Forecasting the water flows draining to the Great Barrier Reef using the G2G distributed hydrological model [Draft].

Worland, S. C., Farmer, W. H., & Kiang, J. E. (2018). Improving predictions of hydrological low-flow indices in ungaged basins using machine learning. *Environmental modelling & software*, 101, 169-182. https://doi.org/10.1016/j.envsoft.2017.12.021

Yaseen, Z. M., Ramal, M. M., Diop, L., Jaafar, O., Demir, V., & Kisi, O. (2018). Hybrid Adaptive Neuro-Fuzzy Models for Water Quality Index Estimation. *Water Resources Management*, 32(7), 2227-2245. doi:10.1007/s11269-018-1915-7

Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., & Chau, K.-W. (2019). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction*. Journal of Hydrology*, 569, 387-408. https://doi.org/10.1016/j.jhydrol.2018.11.069

Zhang, Q., Bostic, J. T., & Sabo, R. D. (2022). Regional patterns and drivers of total nitrogen trends in the Chesapeake Bay watershed: Insights from machine learning approaches and management implications. *Water Research*, 218. https://doi.org/10.1016/j.watres.2022.118443.

Zhi, W., & Li, L. (2020). The shallow and deep hypothesis: Subsurface vertical chemical contrasts shape nitrate export patterns from different land uses. *Environmental science & technology*, 54(19), 11915-11928. https://doi.org/10.1021/acs.est.0c01340

APPENDIX A: Journal Paper 1 – Supplementary Material

Figure S1: Spatial location of deductive datasets that represent gauged catchments, nested within parent catchment.
* =Location of gauging station. Colour coding standard from ABARES (2016) for Land use, and Neldner (2019) for original ecosystem.

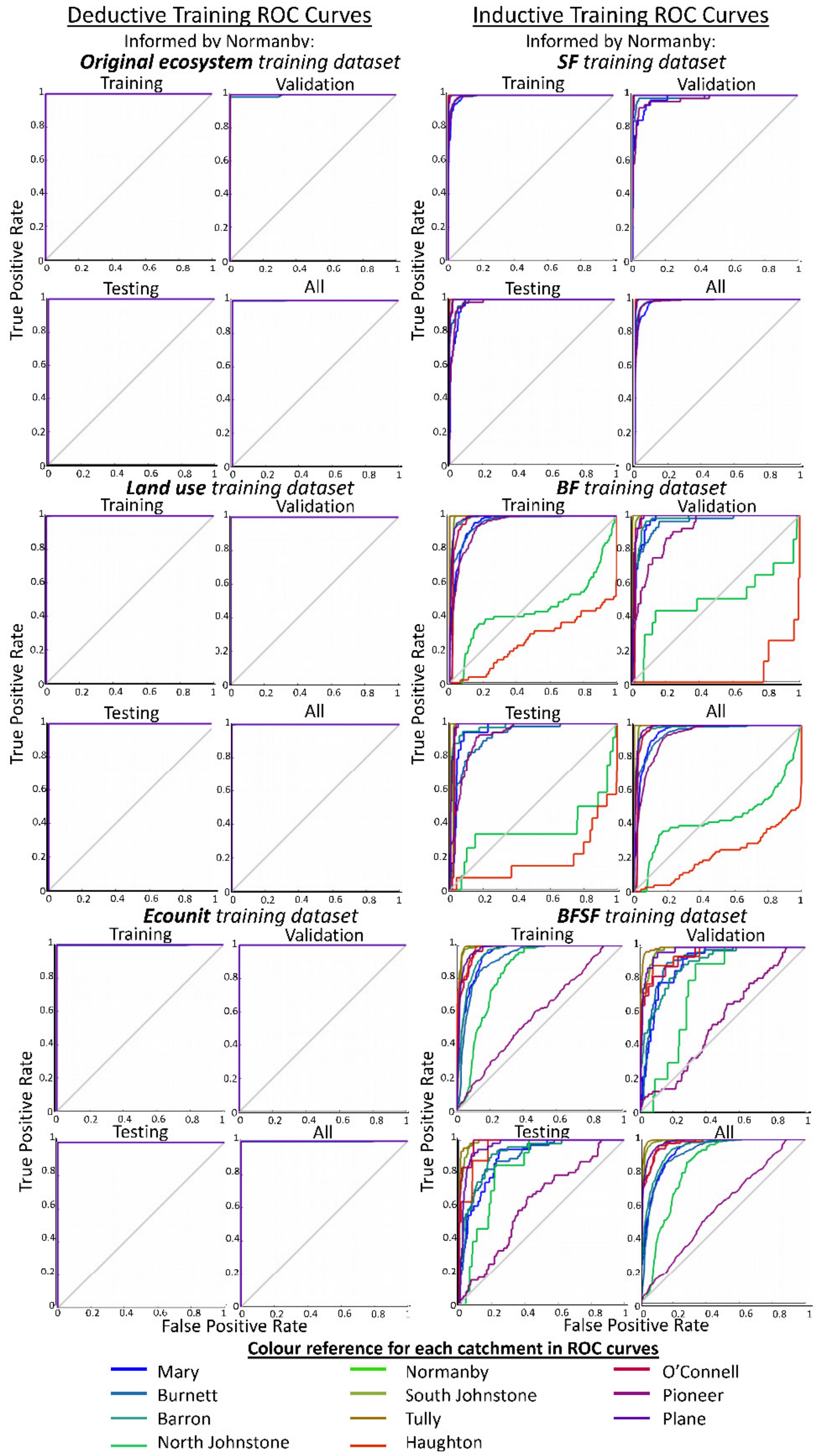
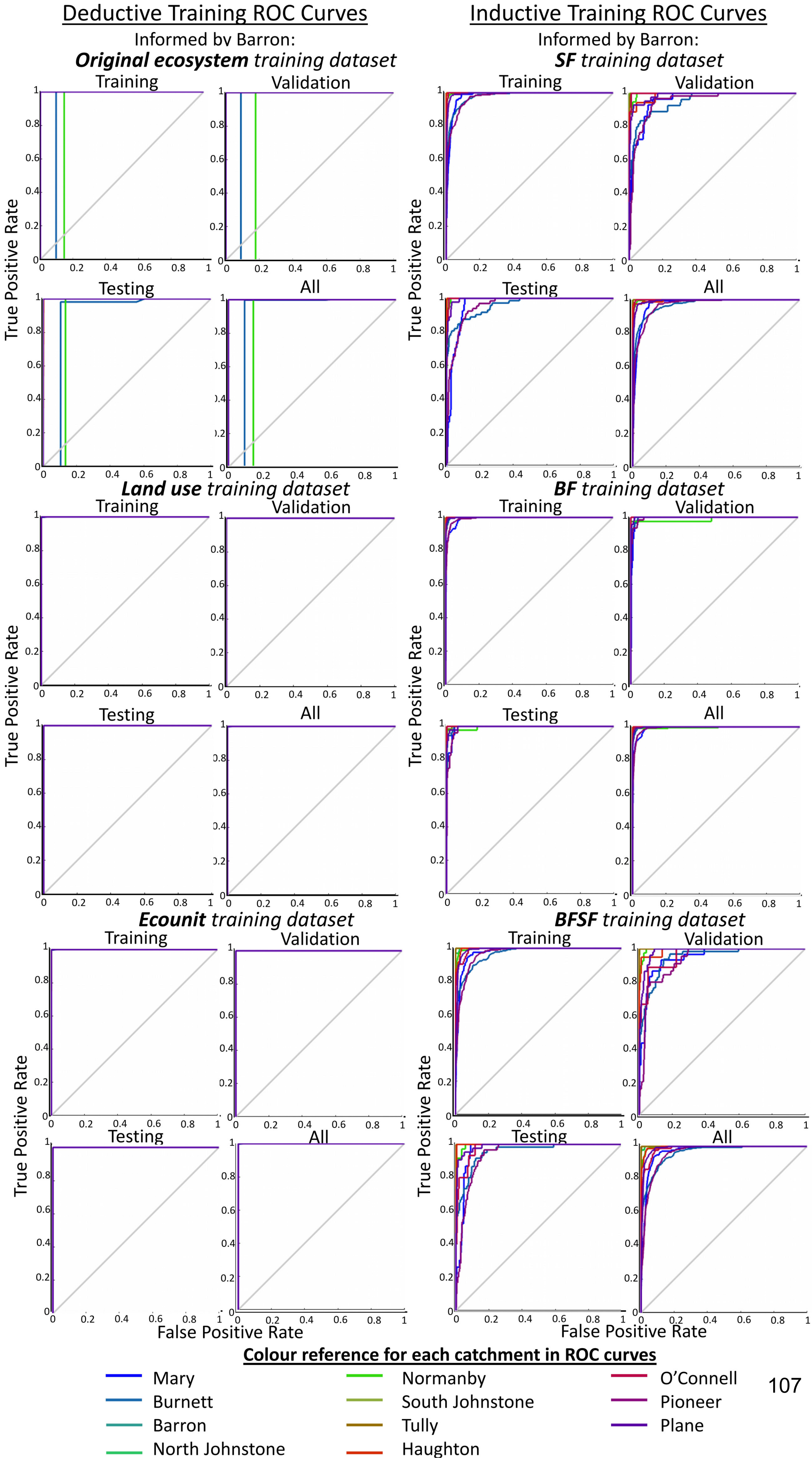
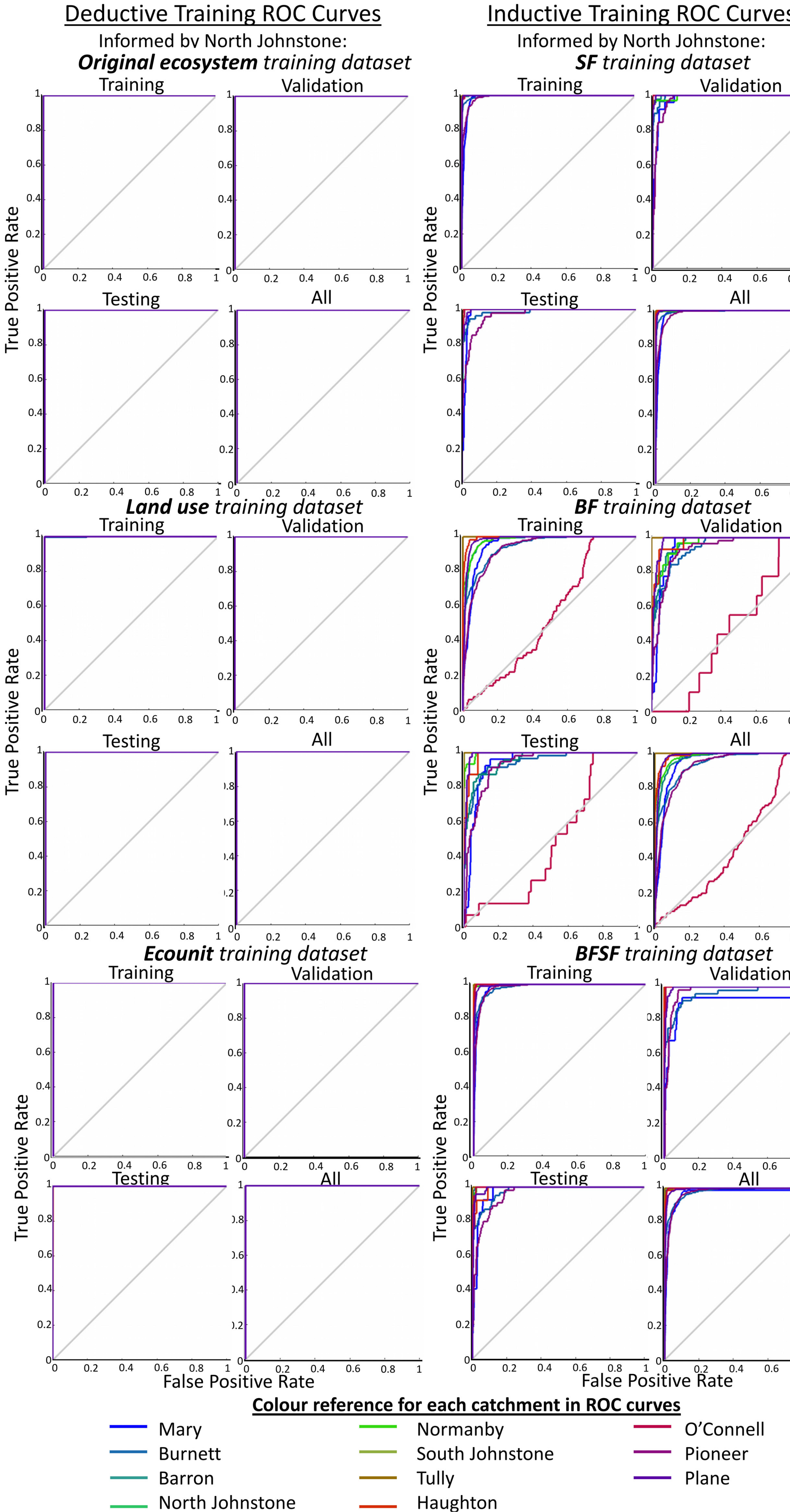
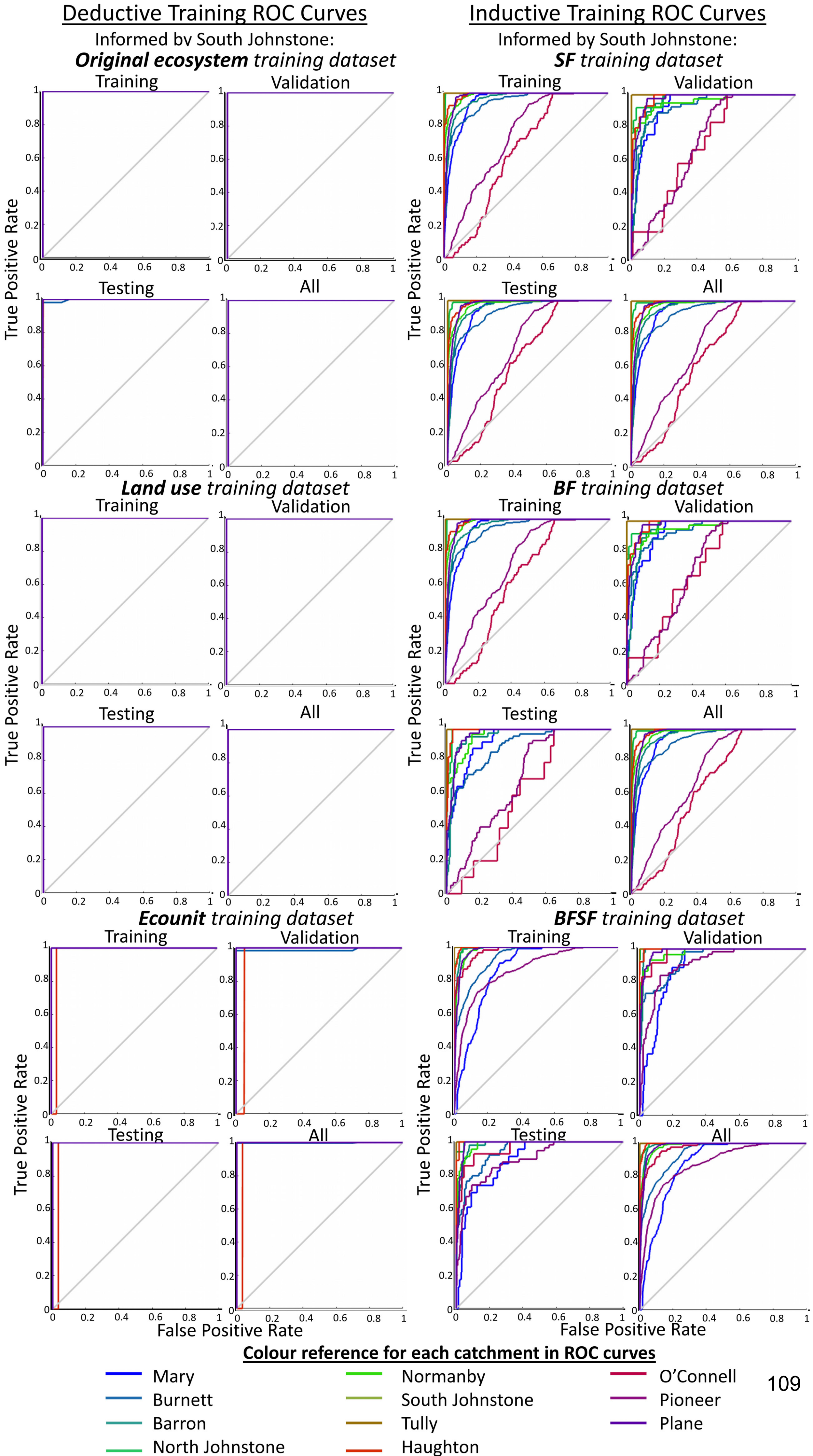
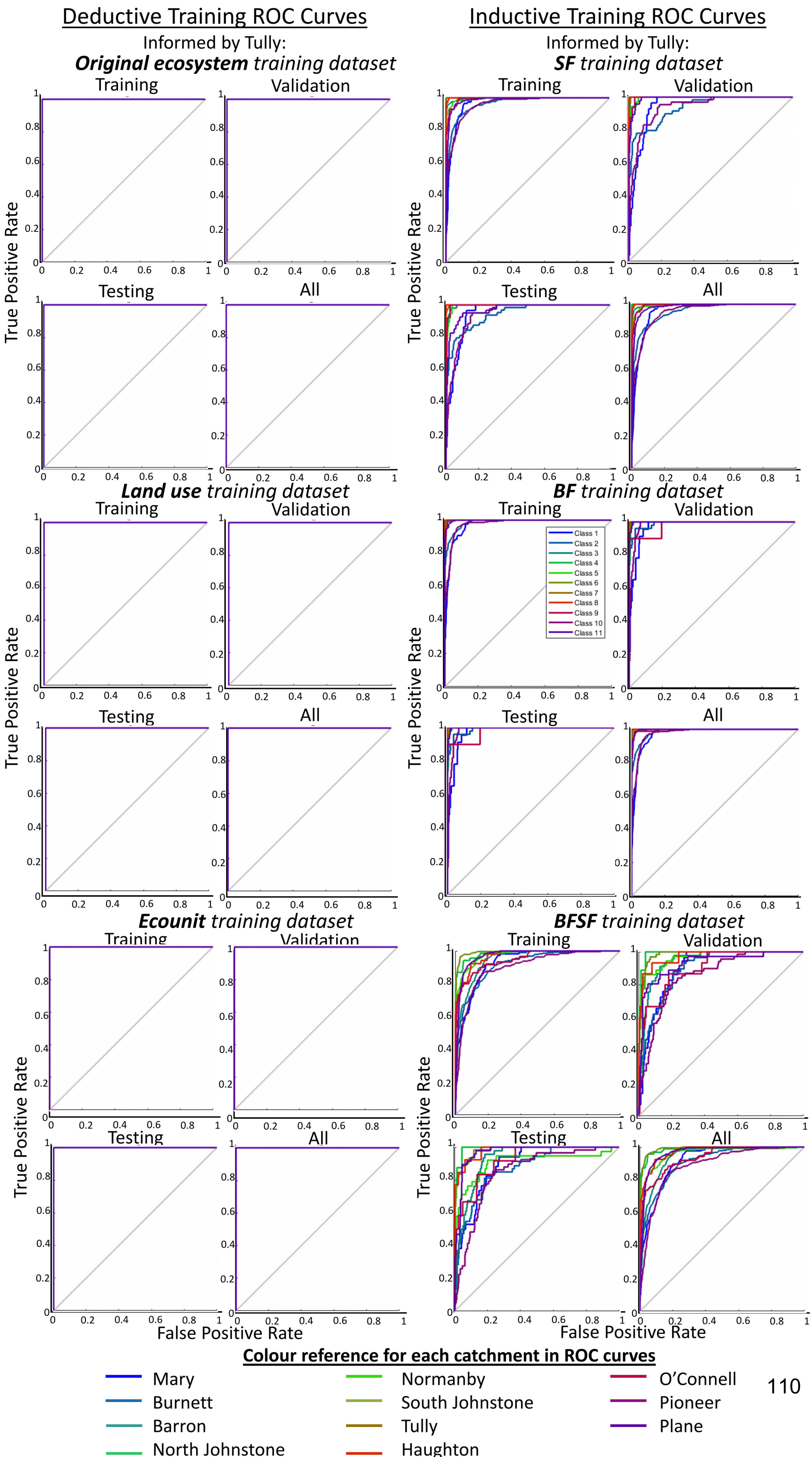


Fig. S2. ROC curves for ANN-PR training performance of classification of deductive (left column) and inductive (right column) datasets for each catchment, as stated. Each panel includes top left to right, then bottom left to right: ROC curves for initial training (70% of data), Validation (15% of data), Testing (15% of data) and the overall performance in All ROC. True positive is shown on a scale of 0-1 on the y axis, and false -positive is shown as 0-1 on the x-axis. Lines show the ability of ANN-PR to train the dataset for each respective catchment.

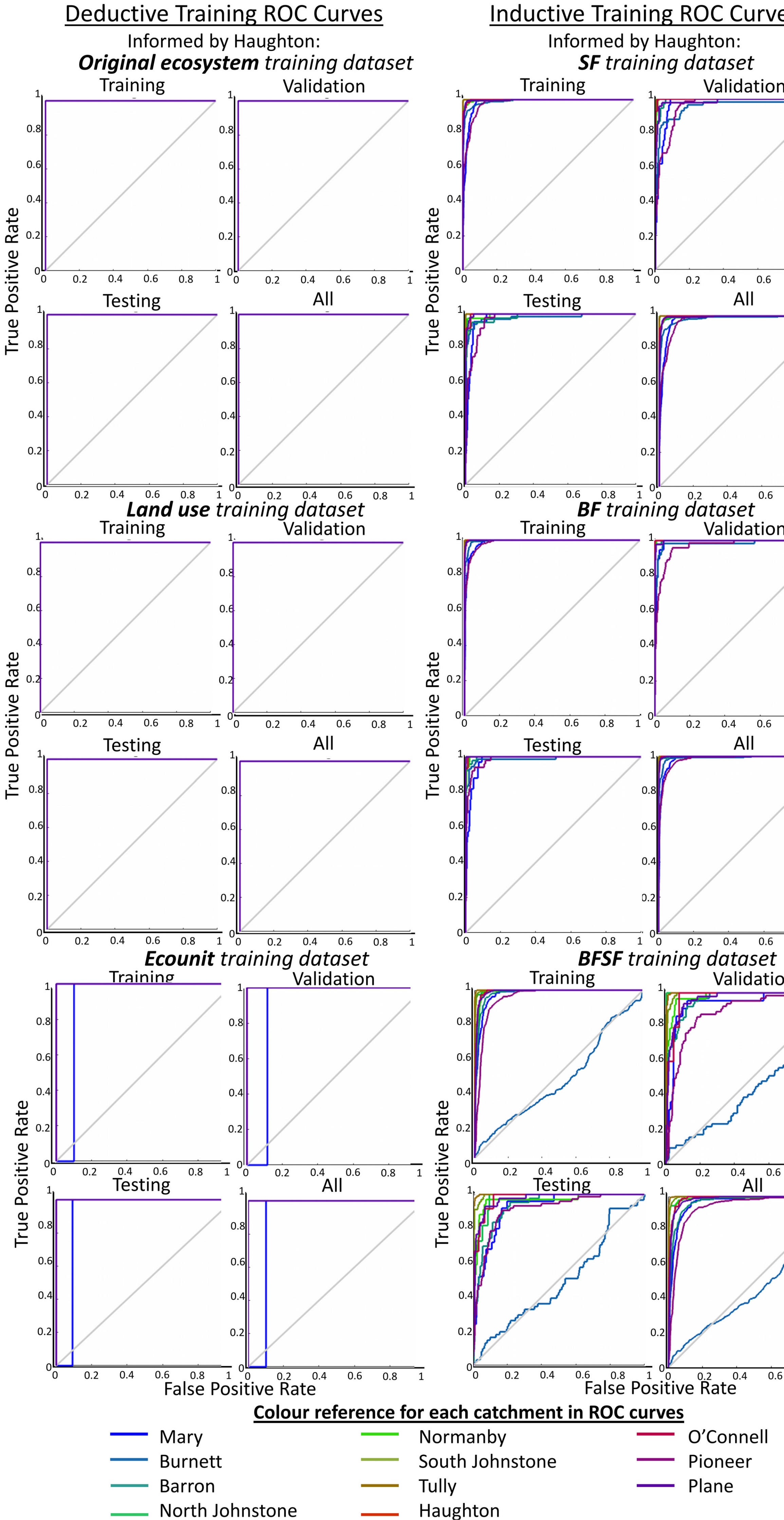

Deductive Training ROC Curves
Informed by Barron:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by Barron:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane


Deductive Training ROC Curves
Informed by North Johnstone:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by North Johnstone:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane

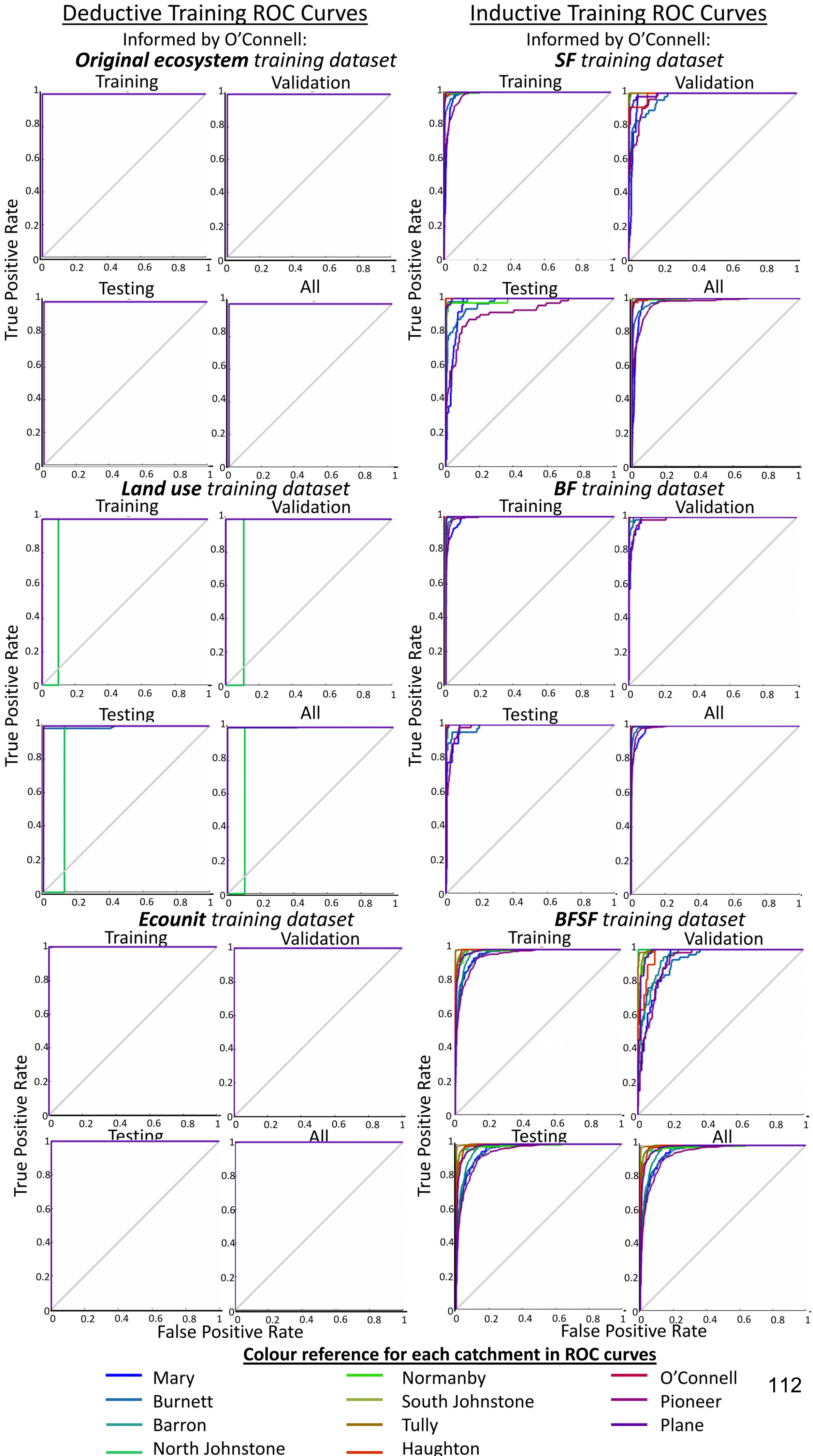

Deductive Training ROC Curves
Informed by South Johnstone:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by South Johnstone:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane



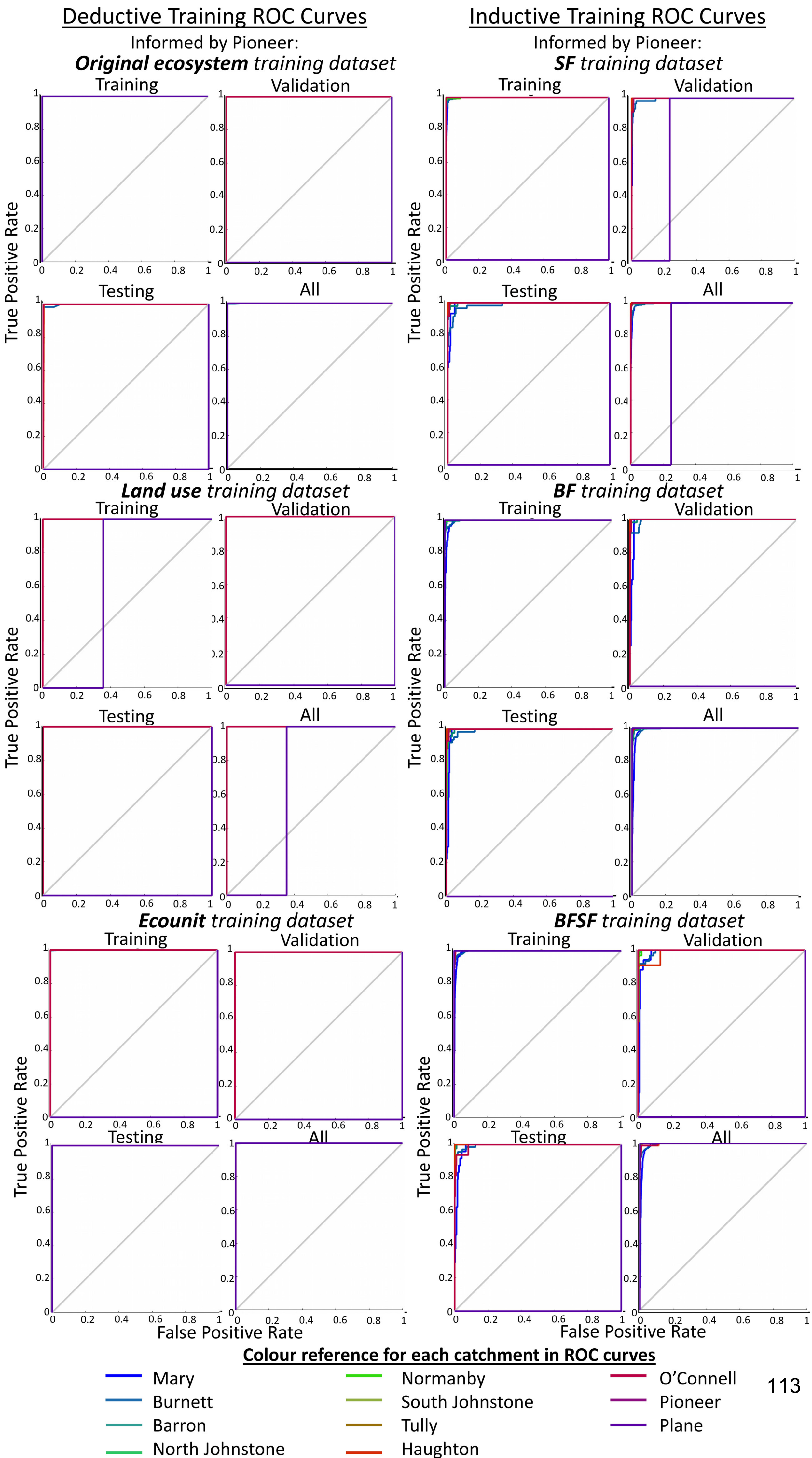
Deductive Training ROC Curves
Informed by Tully:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by Tully:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Class 1
Class 2
Class 3
Class 4
Class 5
Class 6
Class 7
Class 8
Class 9
Class 10
Class 11
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane

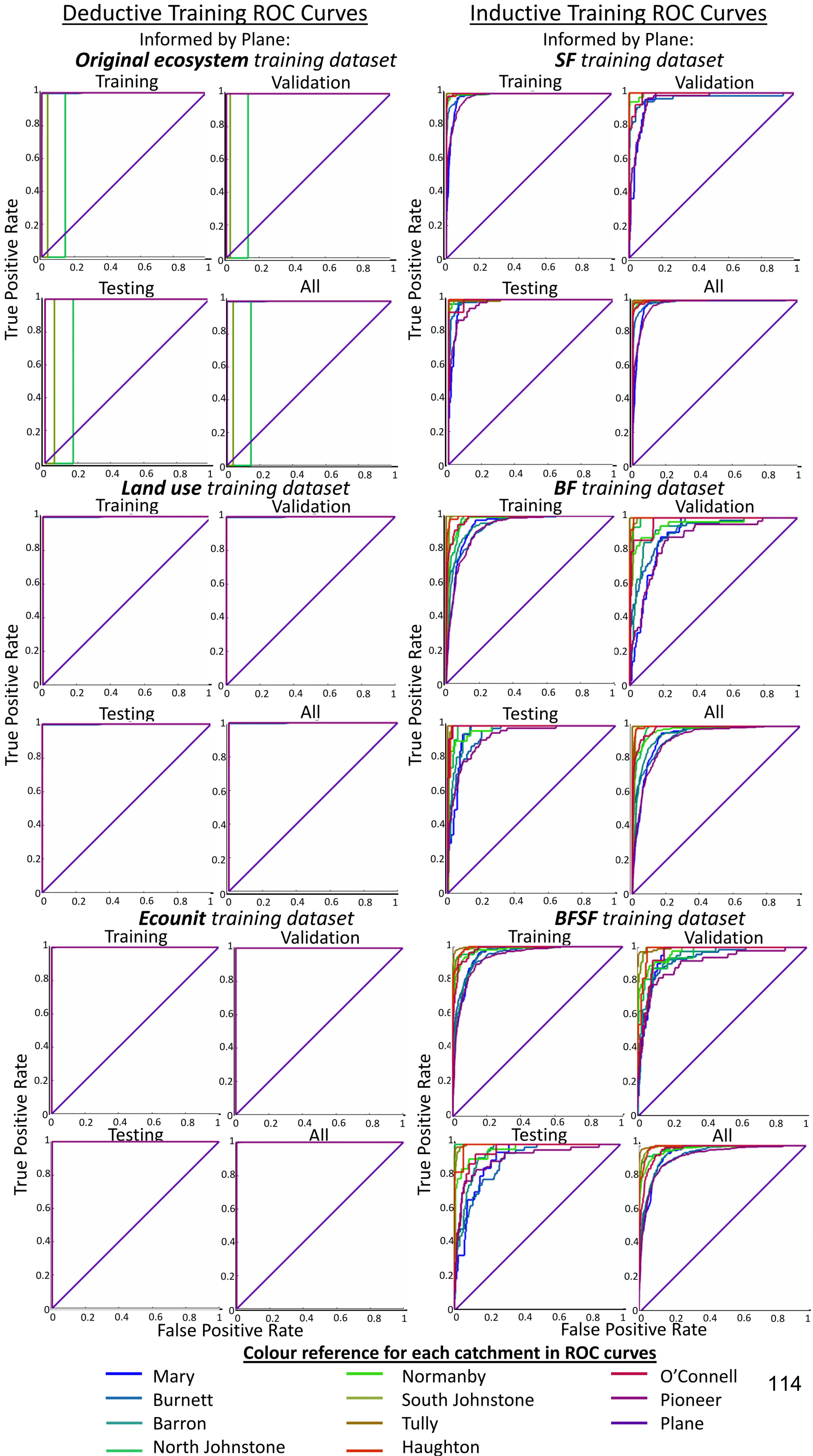


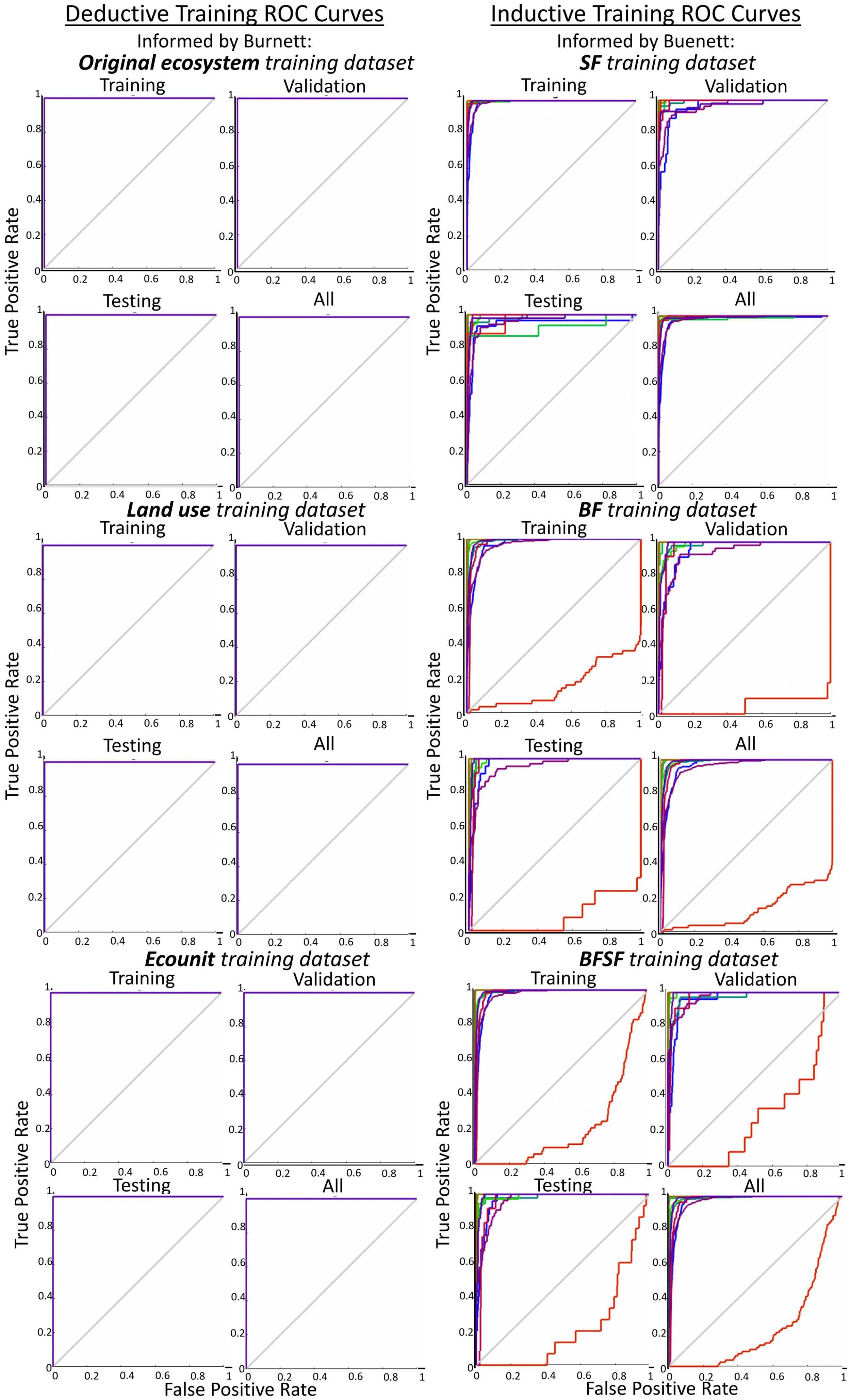
Deductive Training ROC Curves
Informed by Haughton:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by Haughton:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane



Deductive Training ROC Curves
Informed by O'Connell:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by O'Connell:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
0
0.2
0.4
0.6
0.8
1
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane

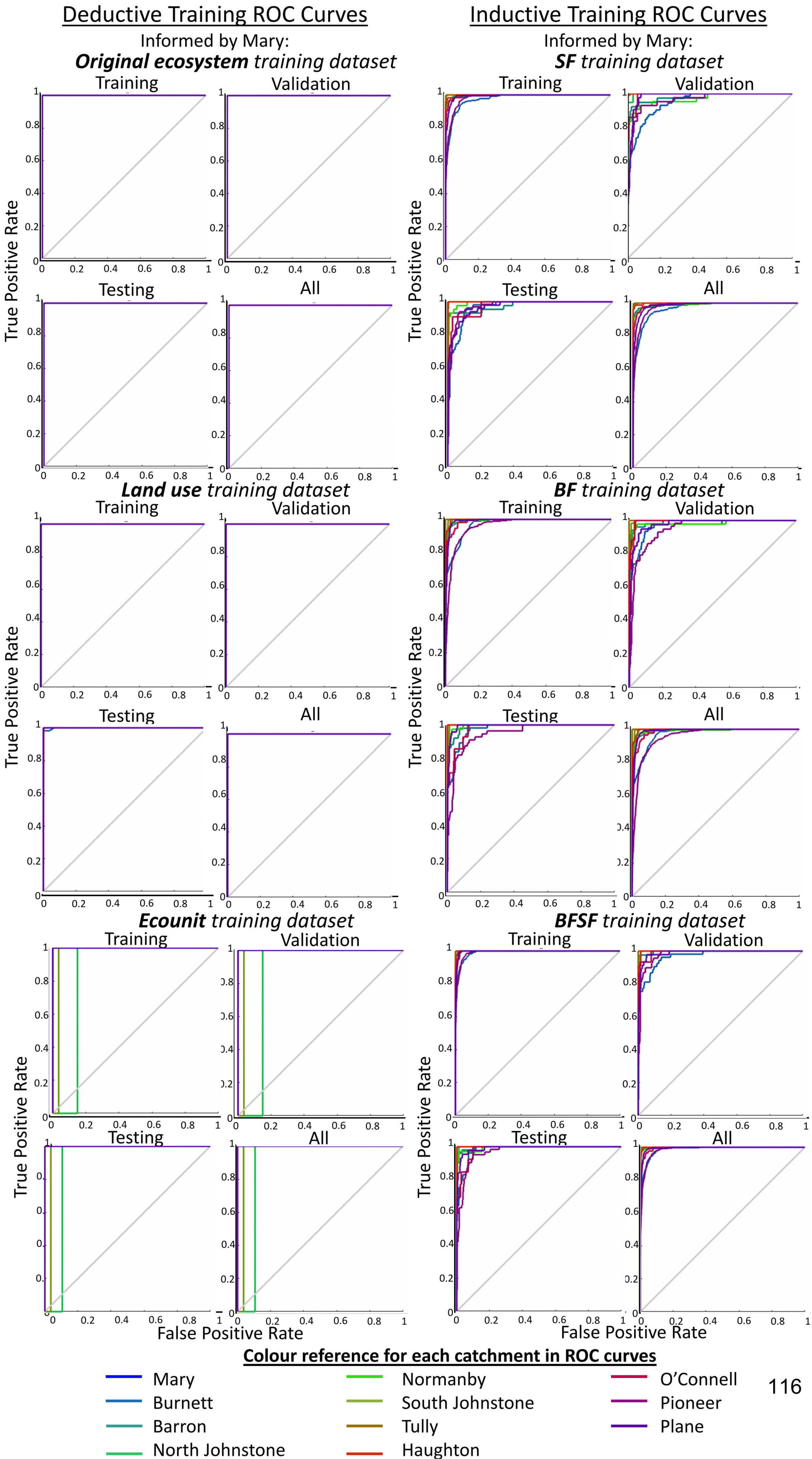

Deductive Training ROC Curves
Informed by Pioneer:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by Pioneer:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane


Deductive Training ROC Curves
Informed by Plane:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by Plane:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane

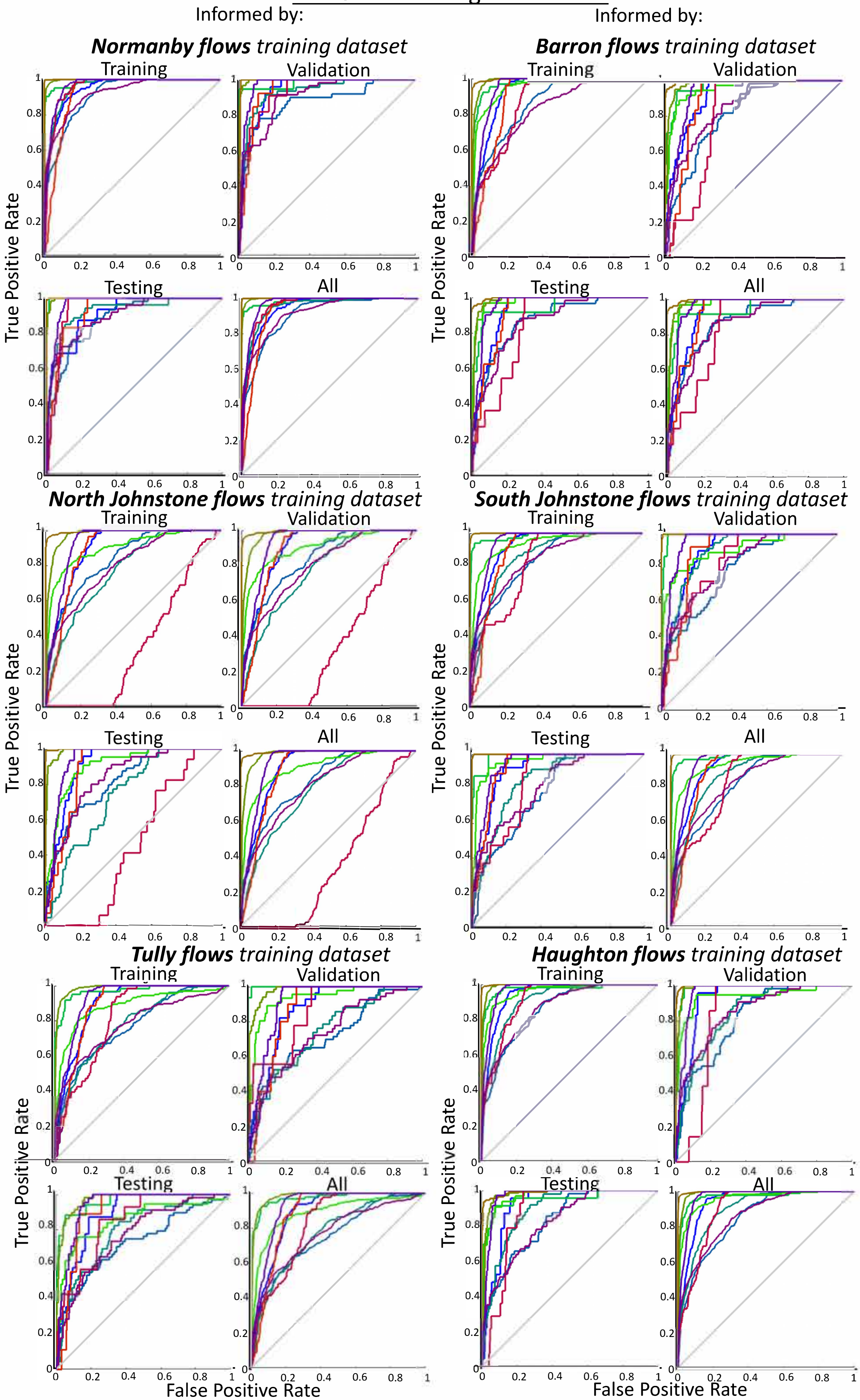


Colour reference for each catchment in ROC curves

Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane

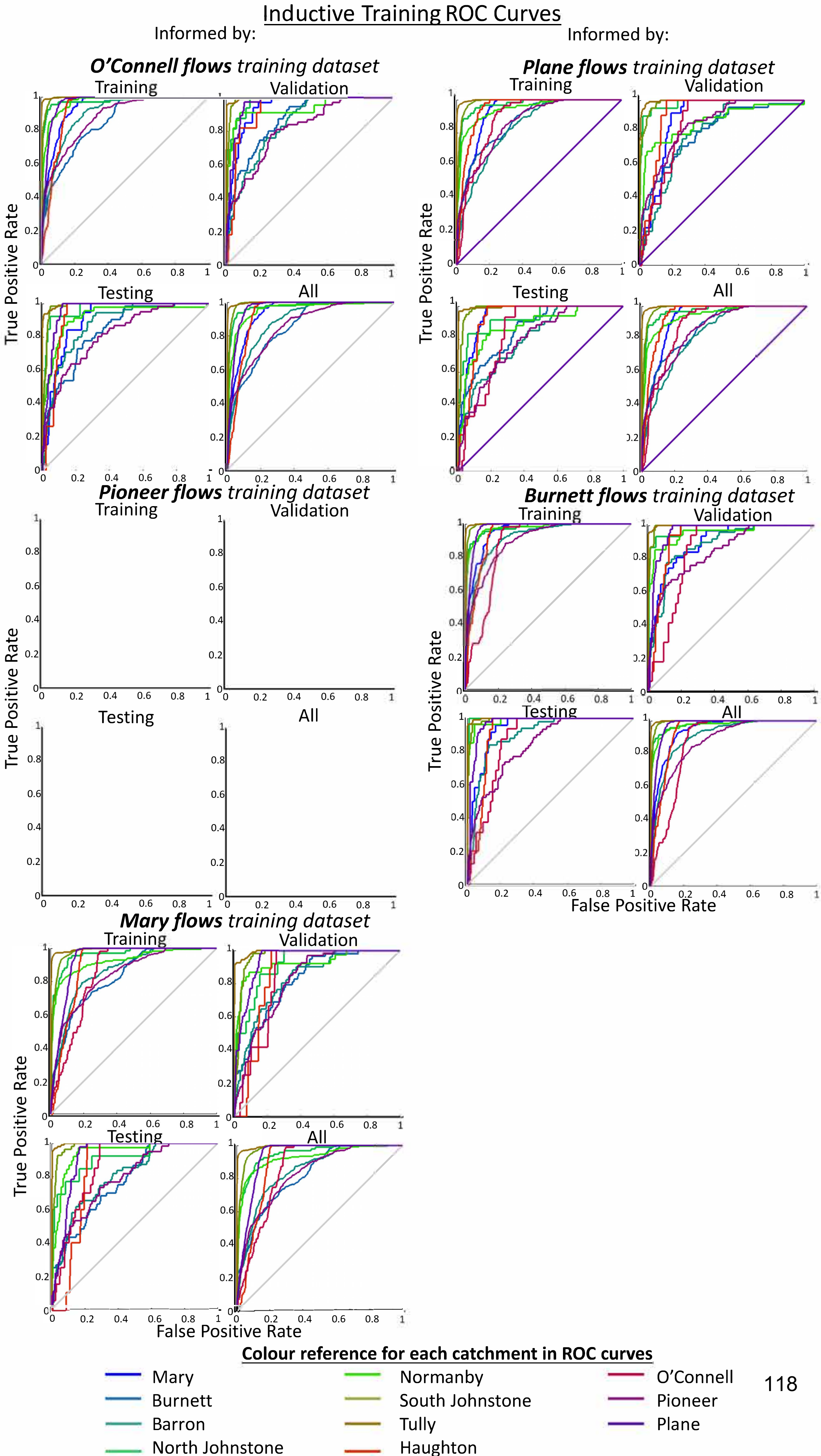

Deductive Training ROC Curves
Informed by Mary:
Original ecosystem training dataset
Land use training dataset
Ecounit training dataset
Inductive Training ROC Curves
Informed by Mary:
SF training dataset
BF training dataset
BFSF training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane

Inductive Training ROC Curves



Colour reference for each catchment in ROC curves

Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane



Inductive Training ROC Curves
Informed by:
O'Connell flows training dataset
Training
Validation
Testing
All
True Positive Rate
Informed by:
Plane flows training dataset
Training
Validation
Testing
All
True Positive Rate
Pioneer flows training dataset
Training
Validation
Testing
All
True Positive Rate
Burnett flows training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Mary flows training dataset
Training
Validation
Testing
All
True Positive Rate
False Positive Rate
Colour reference for each catchment in ROC curves
Mary
Burnett
Barron
North Johnstone
Normanby
South Johnstone
Tully
Haughton
O'Connell
Pioneer
Plane

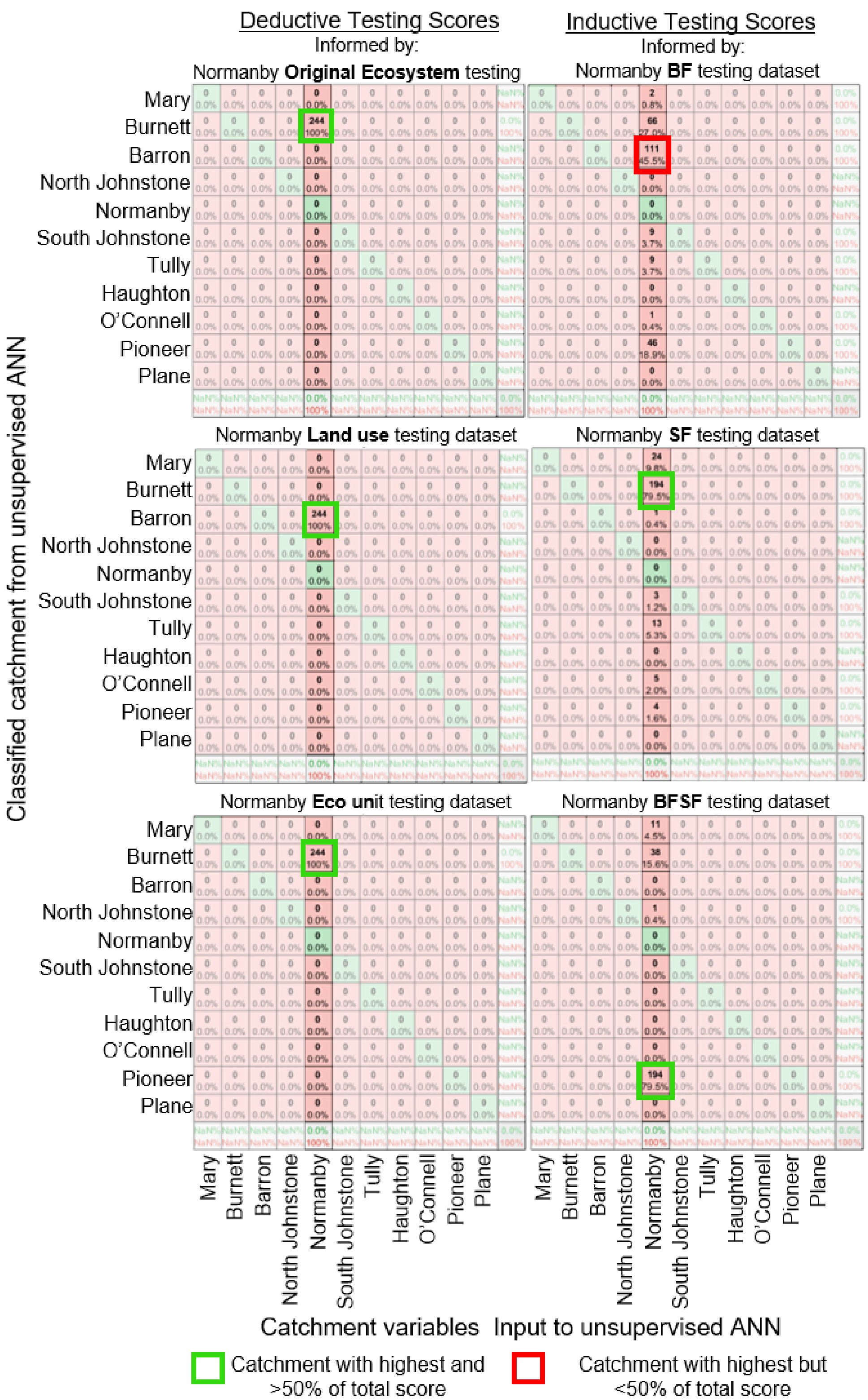


Fig. S3. Confusion matrix for application of the trained ANN-PR networks to deductive (left) and inductive (right) datasets for Normanby. The classification scores are the number of times the trained network allocated each variable record for Normanby to the catchment the trained ANN-PR network deemed had the most similar data pattern. Catchments with the highest score are the catchments classified for the dataset.

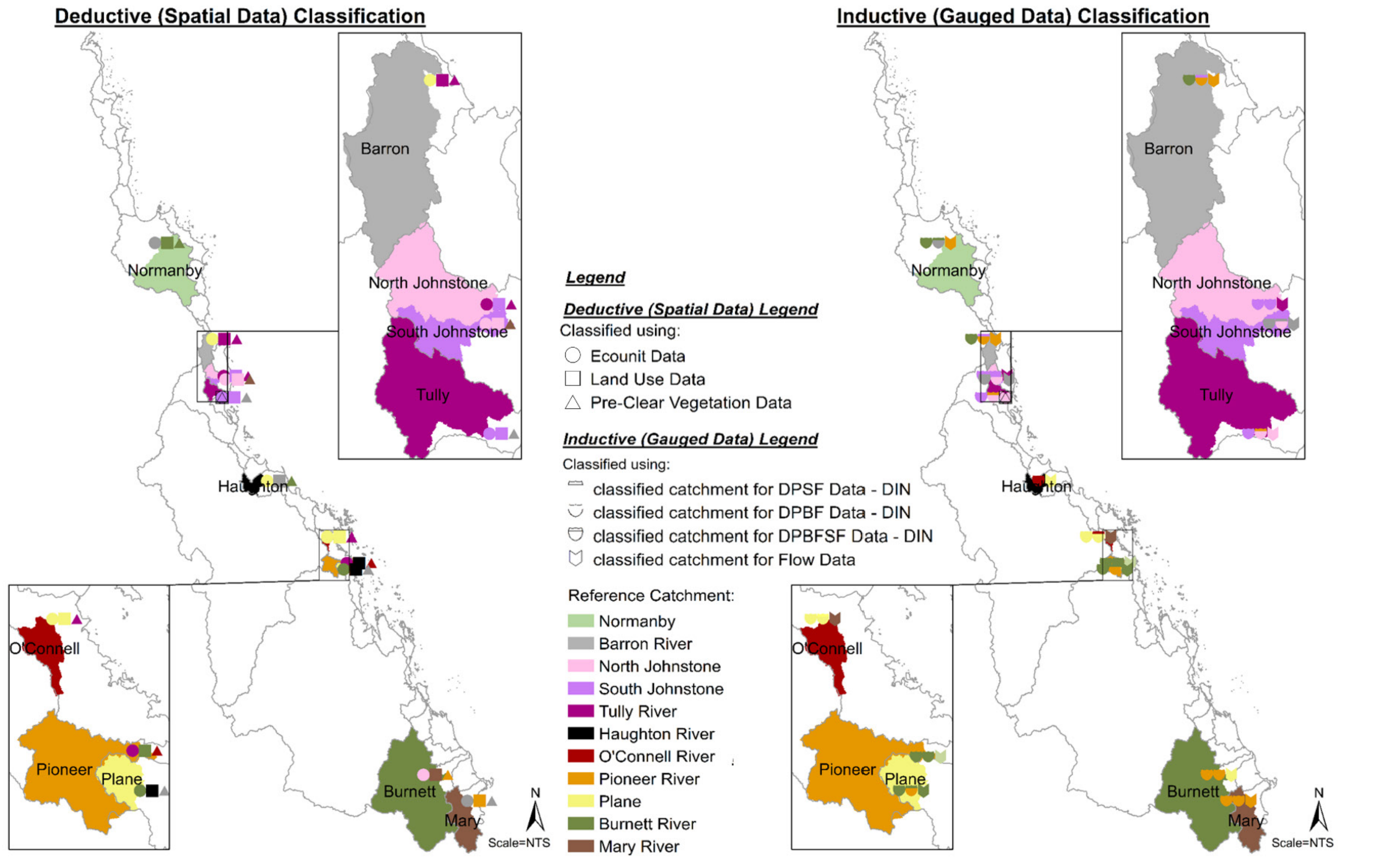


Fig. S4. Spatial representation of deductive and inductive pairing results. Map shows the catchments in their reference colours along with symbols coloured to reference the paired catchment in the unsupervised ANN environment. The class that the reference catchments variables were clustered to are represented by the colour of the respective symbol. Shapes of the symbol represent the dataset the pairing clustered to as follows: = SFBF, = BF, = SF, = flows. ○ = Ecounits, □ = land use and Δ = Original ecosystem.

Table S1: Catchment characteristics and data summary

Catchment Name	Normanby	Barron	North Johnstone	South Johnstone	Tully	Haughton	O'Connell	Pioneer	Plane	Burnett	Mary
Gauging station ID for observed data	105107A	110001D	112004A	112101B	113006A	119003A	124001B	125013A	126001A	136007A	138014A
Gauged catchment area (km^2)	12,828	1,950	926	399	1,386	1,807	336	1,464	327	30,724	6,863
Natural Resource Management Region	Cape York	Wet tropics	Wet tropics	Wet tropics	Wet tropics	Burdekin	Mackay Whitsunday	Mackay Whitsunday	Mackay Whitsunday	Burnett Mary	Burnett Mary
Gauged Catchment Centroid Latitude (decimal °)	-15.46	-17.05	-17.5	-17.66	-17.87	-19.72	-20.77	-21.23	-21.24	-25.73	-26.19
Gauged Catchment Centroid Longitude (decimal °)	144.56	145.51	145.69	145.77	145.72	146.81	148.56	148.74	148.94	151.28	152.49
DIN Record Period(range of years)	3/10/2006-25/08/2017	19/01/06-15/09/17	30/01/2006-15/09/2017	24/02/2006-15/09/2017	13/01/2006-19/04/2018	20/12/2012-25/09/2017	25/01/2007-24/08/2017	18/10/2006-13/09/2017	4/09/2009-26/08/2017	23/10/20006-15/09/2017	25/09/2013-29/06/2018
DIN sampling frequency	Frequently^ Jan-March	Regular^ monthly, Frequently^ Jan-March	Infrequent^ half yearly, Frequently^ Jan-March	Regular^ monthly, Frequently^ Jan-March	Frequent^ monthly, Frequently^ Jan-March	Regular^ monthly, Frequently^ Jan-March	Irregularly^, Frequently^ Jan-March	Frequent^ monthly, Frequently^ Jan-March	Regular^ monthly, Frequently^ Jan-March	Frequent^ monthly, Frequently^ Jan-March	Frequent^ monthly, Frequently^ Jan-March
Number of records in DIN record period	244	318	94	414	723	80	87	402	302	400	176
Max DIN (mg/L)	1.704	0.634	0.372	0.365	1.876	0.331	0.831	3.557	3.865	4.659	1.293
Min DIN (mg/L)	1.50E-03	0.0015	0.002	0.004	0.006	0.004	0.004	0.0015	0.0015	0.0015	0.0015
Mean DIN (mg/L)	5.50E-02	1.17E-01	1.47E-01	1.26E-01	2.37E-01	6.64E-02	1.09E-01	2.31E-01	4.24E-01	1.61E-01	2.01E-01
Median DIN (mg/L)	3.30E-02	1.01E-01	1.38E-01	1.28E-01	2.06E-01	2.55E-02	6.20E-02	1.70E-01	2.08E-01	1.02E-01	0.1615
Standard Deviation DIN (mg/L)	0.12821	0.090790697	0.072591	0.064849	0.153683	0.088368	0.139678	0.27511	0.533129	0.317718	0.203688
Max Corresponding Streamflow (m^3/s)	1873.7	2148.7	1680.3	1049.9	1030.8	482.97	489.21	3447.7	1494.9	16422	2494.4
Min Corresponding Streamflow (m^3/s)	1.00E-99	0.54	7.132	3.56	10.638	0.009	1E-99	1E-99	0.021	0.0168	1E-99
Mean Corresponding Streamflow (m^3/s)	2.85E+02	1.24E+02	2.10E+02	8.11E+01	2.64E+02	3.20E+01	3.07E+01	2.20E+02	5.48E+01	4.74E+02	1.68E+02
Median Corresponding Streamflow (m^3/s)	1.48E+02	2.48E+01	1.54E+02	4.83E+01	1.97E+02	2.46E+00	6.25E+00	7.15E+01	6.66E+00	3.33E+01	26.652
Standard Deviation Corresponding Streamflow (m^3/s)	375.3701	264.2324438	252.8503	123.8116	217.0935	86.8781	63.60238	426.8036	138.3731	1575.409	333.1158
Max Corresponding Baseflow (m^3/s)	390.4533	102.7543327	127.7338	57.96873	275.6916	5.703567	15.07863	95.88294	17.34001	616.5559	34.94684227
Min Corresponding Baseflow (m^3/s)	1.00E-99	0.409397922	6.120425	3.17202	7.79123	0.013372	1E-99	1E-99	0.014856	0.002498	1.66992E-05
Mean Corresponding Baseflow (m^3/s)	5.74E+01	1.28E+01	4.48E+01	2.08E+01	8.17E+01	1.00E+00	2.85E+00	2.00E+01	2.31E+00	5.28E+01	7.93E+00
Median Corresponding Baseflow (m^3/s)	3.71E+01	5.30E+00	3.98E+01	1.91E+01	7.62E+01	5.60E-01	1.25E+00	9.92E+00	6.98E-01	8.28E+00	4.733613293
Standard Deviation Corresponding Baseflow (m^3/s)	64.61651	19.36324239	30.21545	13.47825	48.82631	1.188735	3.925653	23.88492	3.532949	123.627	8.573484165

*Publically available data from the State Government of Queensland, sourced from Kahn et al. (2020)

^Frequent means mostly >1, regular = mostly 1, infrequent mostly <1

Table S2**:** Datasets input to ANN-PR to identify classified catchments. Note that the data for the catchment to be classified is the only data included in the testing dataset, and the only data omitted from the training datasets. Inductive and deductive datasets are input separately. Inductive datasets are the empirical water quality and flow data, deductive datasets are open-source spatial datasets sourced from publicly available government mapping.

Catchment to be classified	Catchment data for classifiers	Classifiers included in Training datasets: Inductive datasets	Classifiers included in Training datasets: Deductive datasets	Records in Training dataset	Records in Testing dataset: Inductive datasets	Records in Testing dataset: Deductive datasets
Normanby	Normanby	-	-	2996	244	244
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
Barron	Normanby	✓	✓	2922	-	-
	Barron	-	-		318	318
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
North Johnstone	Normanby	✓	✓	3146	-	-
	Barron	✓	✓		-	-
	North Johnstone	-	-		94	94
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
South Johnstone	Normanby	✓	✓	2826	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	-	-		414	414
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
Tully	Normanby	✓	✓	2517	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	-	-		723	723
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-

Catchment to be classified	Catchment data for classifiers	Classifiers included in Training datasets		Records in Training dataset	Records in Testing dataset	
		Inductive datasets	Deductive datasets		Inductive datasets	Deductive datasets
Haughton	Normanby	✓	✓	3160	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	-	-		80	80
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
O'Connell	Normanby	✓	✓	3153	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	-	-		87	87
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
Pioneer	Normanby	✓	✓	2838	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	-	-		402	402
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
Plane	Normanby	✓	✓	2938	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	-	-		302	302
	Burnett	✓	✓		-	-
	Mary	✓	✓		-	-
Burnett	Normanby	✓	✓	2840	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	-	-		400	400
	Mary	✓	✓		-	-
Mary	Normanby	✓	✓	3064	-	-
	Barron	✓	✓		-	-
	North Johnstone	✓	✓		-	-
	South Johnstone	✓	✓		-	-
	Tully	✓	✓		-	-
	Haughton	✓	✓		-	-
	O'Connell	✓	✓		-	-
	Pioneer	✓	✓		-	-
	Plane	✓	✓		-	-
	Burnett	✓	✓		-	-
	Mary	-	-		176	176

Table S3: ROC training performance and best hidden neuron summary. Well trained algorithm indicated by (✓) and defined as more than 7 classifiers (lines) closest to the top right corner of the ROC graph, and exceeding 0.8 on both axes for both true and false positive. Those without a (✓) are weakly trained. Unacceptable classifiers in each training environment are identified by name. Numerals are the optimal number of hidden neurons applied to the training dataset.

Training Dataset	**Deductive Datasets (Spatial data as a proxy for DIN)**			**Inductive Datasets (Observed gauging station data)**			
	Original ecosystem	**land use**	**Ecounit**	**BF**	**SF**	**BFSF**	**Flow**
Normanby ROC performance	✓	✓	✓	N.Johnston Haughton	N.Johnston	✓	
best hn	4	2	4	943	271	859	56
Barron ROC performance	✓	✓ Burnett Normanby	✓	✓	✓		
best hn	2	4	4	983	812	569	35
North Johnstone ROC performance	✓	✓	✓	O'Connell	✓		
best hn	2	3	4	334	673	978	348
South Johnstone ROC performance	✓	✓	✓ Haughton	O'Connell Pioneer			
best hn	3	3	4	776	476	938	309
Tully ROC performance	✓	✓	✓	✓	✓ Haughton		✓
best hn	2	3	3	562	477	809	63
Haughton ROC performance	✓	✓	✓ Mary	✓	✓	Burnett	
best hn	4	6	3	645	149	645	81
O'Connell ROC performance	✓	✓ Normanby	✓	✓	✓		
best hn	2	3	3	510	389	525	44
Pioneer ROC performance	✓	✓ Plane	✓ Plane	✓	✓ Plane	✓	
best hn	6	4	7	490	847	94	45
Plane ROC performance	✓ N.Johnstone S.Johnstone	✓	✓				
best hn	2	3	5	377	180	566	85
Burnett ROC performance	✓	✓	✓	✓ Haughton	✓	✓ Haughton	
best hn	2	3	3	559	525	216	53
Mary ROC performance	✓	✓	N.Johnstone S.Johnstone	✓	✓	✓	
best hn	2	4	4	478	326	669	53

Table S4: Kruskal-Wallis Test for Independent analysis. Results here show that classification scores for Flow datasets are independent of the deductive dataset classification scores, whereas DIN datasets are not.

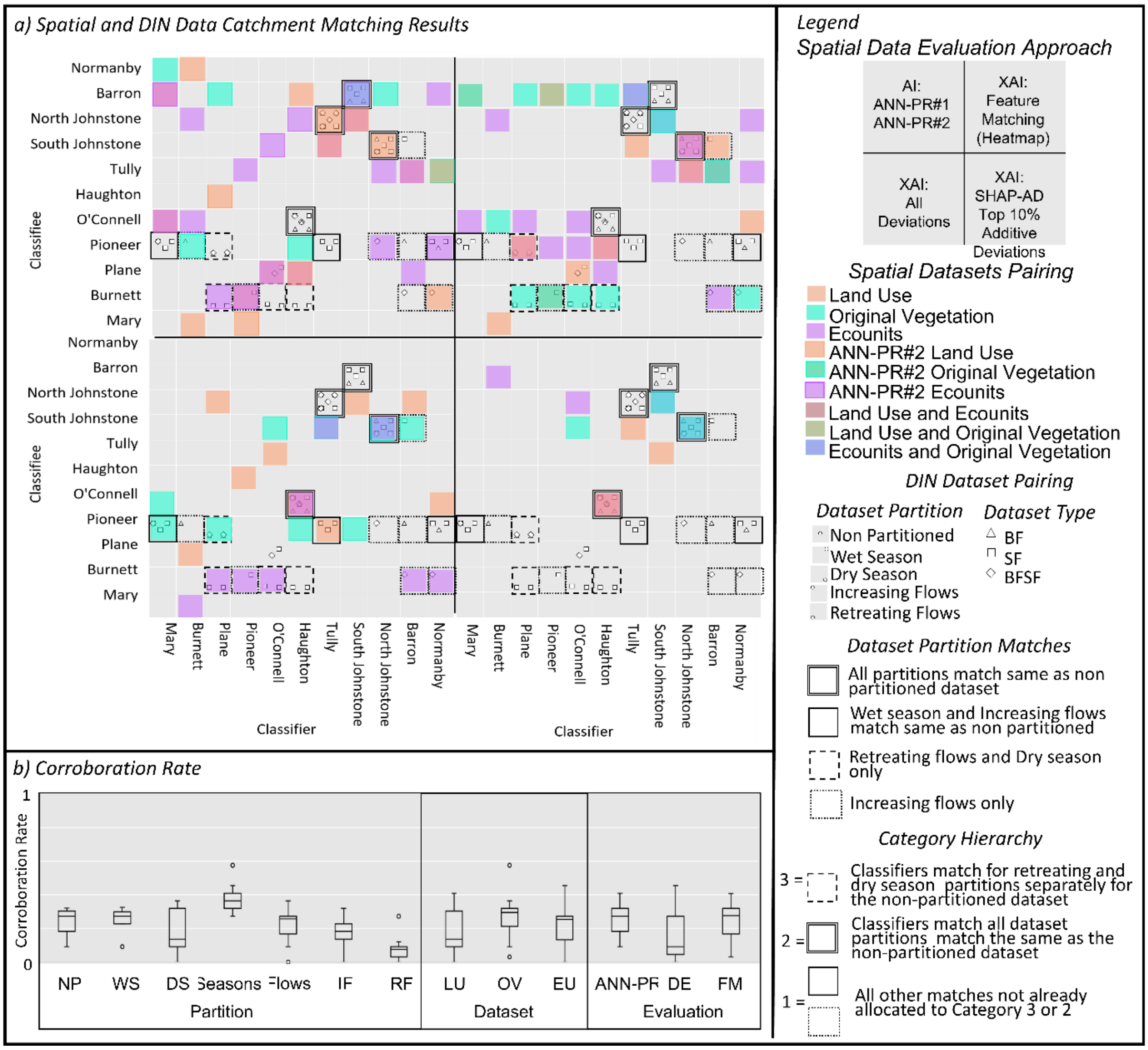
Null Hypothesis	**Statistical Significance**	**Decision**
The distribution of classification scores for "Flows" is the same across deductive classification categories of original ecosystem, Land-use, Ecounit and others for the 11 catchments	0.09	Retain the null hypothesis
The distribution of classification scores for "BF" is the same across deductive classification categories of original ecosystem, Land-use, Ecounit and others for the 11 catchments	0.01	Reject the null hypothesis
The distribution of classification scores for "SF" is the same across deductive classification categories of original ecosystem, Land-use, Ecounit and others for the 11 catchments	0.02	Reject the null hypothesis
The distribution of classification scores for "SFBF" is the same across deductive classification categories of original ecosystem, Land-use, Ecounit and Others for the 11 catchments	0.02	Reject the null hypothesis

Table S5: Classification rates for different classification categories. Classification Category 1=majority (>50%) of records from input catchment paired to the same classifier as a deductive dataset, Category 2=highest number of records from input catchment paired to the same classifier as a deductive dataset, 3=Deductive dataset classifiers paired with any amount of records regardless of highest or lowest in the dataset.

Corroborated classification rates

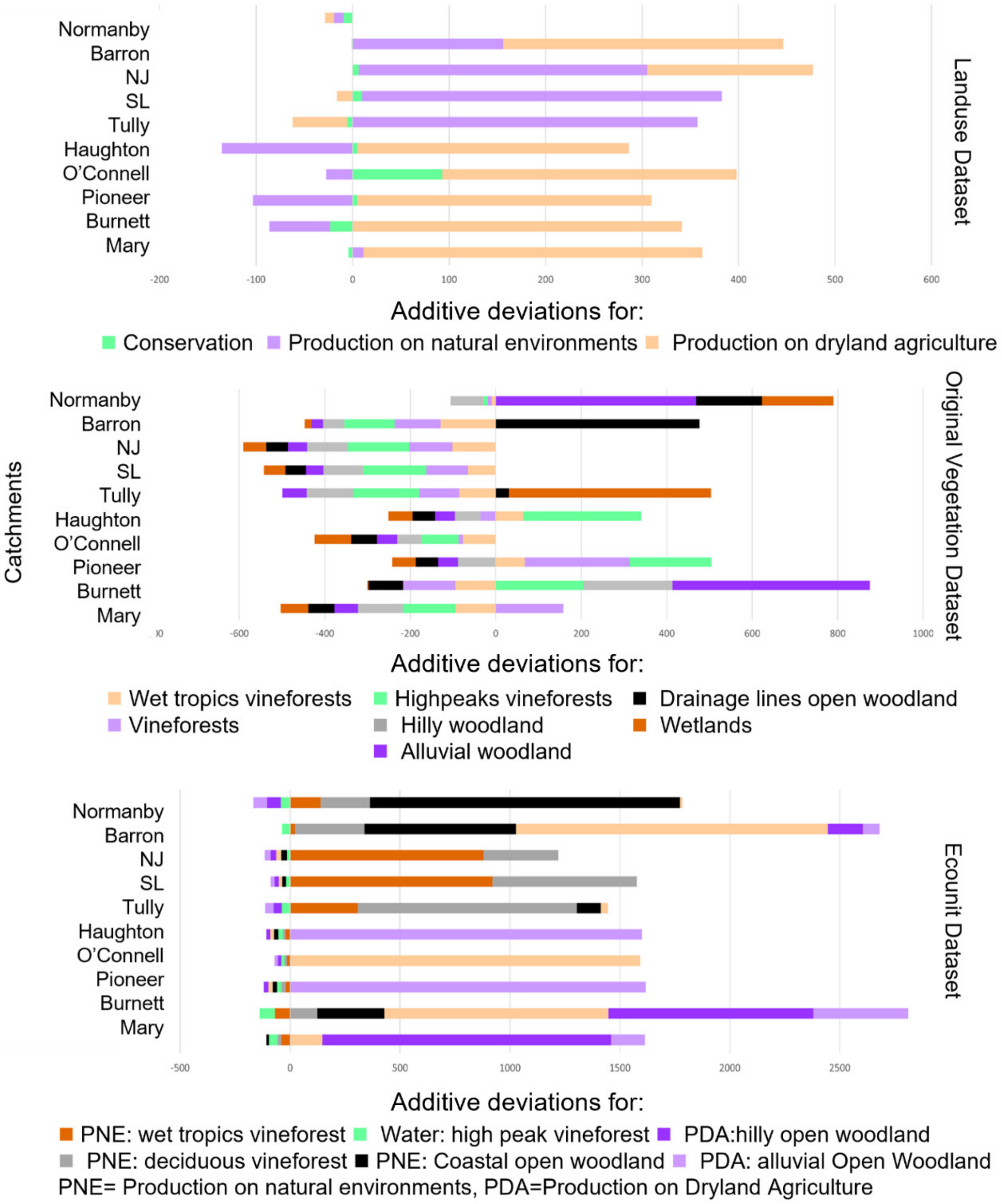
Classification category			Flows	BFSF	SF	BF	Land use	Original ecosystem	Ecounit
	1	By dataset	0.27	0.45	0.55	0.36			
		DIN dataset		0.64			0.45	0.27	0.18
		Flow dataset	0.27				0.00	0.09	0.27
	2	By dataset	0.36	0.73	0.64	0.64			
		DIN dataset		0.82			0.63	0.18	0.45
		Flow dataset	0.36				0.18	0.18	0.27
	3	By dataset	1.0	1.0	1.0	1.0			
		DIN dataset		1.0			0.81	0.81	1.0
		Flow dataset	1.0				0.90	0.81	0.9

APPENDIX B: Journal Paper 2 – Supplementary Material



Supplementary Material Figure SF1: Corroboration of Spatial and DIN Data Catchment Matching Results. (a) Catchment matches for partitioned DIN and spatial data pairing using different evaluation approaches. Legend describing elements of (a). (b) Corroboration rate (Eq 5) for matching DIN dataset pairing with spatial dataset pairing for DIN dataset partitioning, spatial dataset, and evaluation technique indices. NP: Non-Partitioned, WS: Wet Season, DS: Dry Season, IF: Increasing Flows, RF: Retreating Flows,) dataset type (LU: Land Use, OV: Original Vegetation, EU: Ecounits), and evaluation techniques (ANN-PR: Artificial Neural Network Pattern Recognition, DE: XAI Deviation Evaluations from SHAP-AD principles, FM: XAI Feature Matching. Results show rate of corroboration for the DIN dataset pairing with spatial data pairing is the best for seasonal partitioning across all datasets and evaluation techniques combined because it achieved the highest rate of corroboration with lowest variability.

Supplementary Material Figure SF2: Barcode heatmap of spatial features that deviate from the dataset mean across all catchments. This figure facilitates for manual evaluation of catchments that share the same portion of catchment features that deviate from all others in the dataset. The darker the indicator, the more the area of each feature deviates from all others in the dataset. Each increment represents spatial data features as listed.



Supplementary Material Figure SF3: SHAP-AD results, i.e. the Top 10% of the additive deviation from the for all spatial data X-axis is a unitless ratio of additive deviation of spatial variables from all other variables in the dataset. Catchments are arranged north to south along the Y-axis. Number references for each colour are the spatial data reference and are listed in the spatial data index. All three graphs show catchments North of Tully have different dominating spatial features compared to catchments from Haughton south.

Supplementary Material S1: Inductive Classification Method

S.1.1. Inductive Classification Datasets

Dataset records suitable for inductive classification for stream flow and Dissolved Inorganic Nitrogen are open source and collected by the State Government of Queensland using standardised methods (State of Queensland Department of Environment and Resource Management, 2012; State of Queensland Department of Natural Resources, Mines and Energy, 2018). These data include all available streamflow and dissolved inorganic nitrogen records over the time period from 2006 until 2018. The frequency and period of collection for each record are outlined in Table 1 of the main manuscript. Quality assurance checks were performed on observed data as described by Khan et al., (2020). For this study, baseflow was derived using the approach described in Nathan and McMahon (1990). This involved subtracting the recursive digital filter at the sampling instant as the index of baseflow from the original streamflow (Nathan and McMahon, 1990). Consistent with the methods in O'Sullivan et al., (2022), the dataset was then expanded by transformations detailed in Table S1 to standardise and sharpen integration of the data.

Records from the DIN and flow observations dataset were then partitioned as detailed in Table 2 of the manuscript to represent repeating flow events as well as temporal periods. There was a variation in the time water quality records were collected in each catchment, therefore, the number of records for each partitioned dataset varied as detailed in Table 1 of the main manuscript.

S.1.2. Partitioned Inductive Classification

Partitioned inductive classification refers to catchments that have the most closely matched water quality patterns within the dataset partition zones used for this study. The ANN-PR method for inductive classification (O'Sullivan et al., 2022) was used to establish classification scores for each respective dataset. The ANN-PR method facilitates catchments being matched together based on patterns in their datasets. This is achieved by nominating the data in the catchment seeking to be matched to another in a pseudo ungauged scenario as the classifier catchment, and all other catchments with data as classifee catchments. The ANN-PR trains algorithms using sigmoid transfer function and scaled conjugate gradient backpropagation to identify signals in the data in a supervised scenario to correctly match the patterns in variable datasets to their respective classifier catchments. Untrained variable data from classifier catchments is then introduced as the only variable data and forced to match each record in the variable dataset to the trained signal patterns for the classifee catchments in the output layer. Because the amount of data included within each partitioned dataset varied, classification scores were normalised using Eq S1 to represent the ratio of pattern matches for each classifee catchment to the number of input records for the classifier catchment dataset.

$$S_{i,F,P} = \left(\frac{k_{i,F,P}}{K_{I,F,P}}\right) \qquad \text{Eq S1}$$

where:

S= Classifee ratio score for i,F,P

i : Classifee catchment name

F: Datasetfor Flow state (BFSF, SF, or BF) with associated DIN

P: Dataset partition for F (Wet Season, Dry Season, Increasing flows, Retreating flows)

$k =$ number of records matched between classifee (i)and Classifier (I)

K=Total Number of input records in dataset for classifier catchment (I) (See Table S1)

I: Classifier catchment name

Catchments were classified together where the majority of records in the variable input dataset for the classifier catchment were allocated via the ANN-PR trained algorithms to one classifee catchment(k*i*). Criteria for majority records included 50% of records for the classifier (K) dataset, consistent with O'Sullivan et al., (2022), plus the upper 95% confidence interval to overcome any bias in this study for different dataset sizes ($k_{i,F,P}$) resulting from partitioning of the dataset results.

The Confidence Interval was calculated for the Poisson distribution of the number of records for the classifier catchment (K_I), relative to the fixed number of 10 classifee catchments. Addition of the upper confidence interval increases the precision needed for a particular classifee being pattern matched with the majority of the classifier dataset records ($K_{I,F,P}$) (Schoenberg, 1983). The majority dataset is hence established using Eq S2.

$$M_{IFP} = \left\{ \frac{0.5K_{I,F,P} + \left(\mathrm{P}(0.5K_{I,F,P}{}^{k_{\mathbb{N}}} \exp\left(-0.5K_{I,F,P}\right)/k_{\mathbb{N}}!)\right)}{K_{I,F,P}} \right\} \qquad \text{Eq S2}$$

where:

M: proportional majority of the classifier dataset

K : total number of records in the classifier dataset in ANN-PR input layer

i : Classifee catchment name

F: Dataset for flow state (BFSF, SF, or BF) with associated DIN

P: Dataset partition for F (Wet Season, Dry Season, Increasing flows, Retreating flows).

$k_{\mathbb{N}}$: total number of classifee groups in ANN-PR output layer (i.e., number of groups for data to match to)

P : upper 95% confidence interval calculated for equal distribution of scores across all classifier categories and calculated from a fitted Poisson distribution

In this study $\mathrm{P}(0.5K_{I,F,P}{}^{k_{\mathbb{N}}} \exp\left(-0.5K_{I,F,P}\right) / k_{\mathbb{N}}!$ is the upper confidence interval for the Poisson Probability function and was established using the Poisson distribution function in Matlab (Evans et al., 1993).

Classification scores ($\mathrm{S}_{i,F,P}$) that exceeded the proportional majority for the classifier dataset (M_{IFP}) were then considered a pattern match for further evaluation. Eq S3.

$$\mathrm{ICC}_{I,F,P} \textit{ if } \mathrm{S}_{i,F,P} > M_{IFP} \qquad \text{Eq S3}$$

where:

ICC: Inductively Classified Catchment

$S_{i,F,P}$: Classifee ratio score as described in Eq S1; and

M_{IFP}: Proportional majority of the dataset as described in Eq S2.

Matched catchments for each dataset partition were graphed to visualise whether partitioning the dataset changed the classifee catchment matched to each classifier and group catchments with similar responses to dataset partitioning into categories.

Table S1. Transformation of a dataset to increase variables that logically relate to water quality dynamics.

Transformation	Equation	Relationship to water quality
Daily mean	$X=\frac{\sum x1+x2+x3\ldots x24}{24}$ Where: X=daily mean of variable x_n=hourly record of variable	Standardise flow data to DIN concentration
Average daily flow rate for days 1-96 preceding each DIN record	Daily flow rate records transposed to become "x1….x96" prior day flow variables for each record	Capture flow dynamics, and therefore any nutrient exhaustion, during periods between DIN data collection.
Loads	Loads(mgDIN/s) = *WQcs* x *F* Where: *WQcs* = constituent concentrations, (i.e. DIN mg/L), *F*= corresponding flow rate (L/s)	Quantify the amount of DIN discharging from catchments.
Relative water availability	Geometric mean $x = x - \sqrt[n]{\Pi x}$ Where: $\Pi\, x$ =product of the rate of change between each baseflow or streamflow record; n=number of records in the dataset	Capture the influence of Increasing or Retreating flows (Bardgett et al., 2014, Carillo et al., 2011, Carfora et al., 2021, Li et al., 2021, Peng & Chen 2021)
Box-Cox	$y_i^{(\lambda)} = \{\begin{matrix} \lambda^{-1}(y_i^{\lambda}-1) & if\ \lambda \neq 0 \\ log(y_i) & f\ \lambda = 0 \end{matrix}$ such that, for unknown λ, $y(\lambda)=X\beta+\epsilon \quad (2) y(\lambda)=X\beta+\epsilon$ where: $y^{(\lambda)}$ = λ-transformed data; X =covariates; β =parameters; ϵ = error term ($\epsilon_1, \epsilon_2, \ldots, \epsilon_n$).	Overcome skew of datasets with frequent zeros (Box-Cox 1964, McInerney et al., 2017, Shen et al., 2020, Sudheer et al., 2003)
Log	*y*=log (*x*) where: *y* is the transformed data; *x* is the raw data.	Overcome influences caused by right skewed data.

Supplementary References

Bardgett, R. D., Mommer, L., & De Vries, F. T. (2014). Going underground: root traits as drivers of ecosystem processes. Trends in Ecology & Evolution, 29(12), 692-699.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2), 211-243.

Carfora, K., Forgoston, E., Billings, L., & Krumins, J. A. (2021). Seasonal effects on the stoichiometry of microbes, primary production, and nutrient cycling. Theoretical Ecology, 14(2), 321-333.

Carrillo, G., Troch, P. A., Sivapalan, M., Wagener, T., Harman, C., & Sawicz, K. (2011). Catchment classification: hydrological analysis of catchment behavior through process-based modeling along a climate gradient. Hydrology and Earth System Sciences, 15(11), 3411-3430.

Evans, Merran, Nicholas Hastings, and Brian Peacock. Statistical Distributions. 2nd ed. New York: J. Wiley, 1993

Khan, U., Cook, F. J., Laugesen, R., Hasan, M. M., Plastow, K., Amirthanathan, G. E., Bari, M. A. & Tuteja, N. K. (2020). Development of catchment water quality models within a realtime status and forecast system for the Great Barrier Reef. Environmental Modelling & Software, 132, 104790.

Li, X., Zhang, C., Zhang, B., Wu, D., Zhu, D., Zhang, W., Ye, Q., Yam, J., Fu, J., Fang, C., Ha, D., & Fu, S. (2021). Nitrogen deposition and increased precipitation interact to affect fine root production and biomass in a temperate forest: Implications for carbon cycling. Science of the Total Environment, 765, 144497.

McInerney, D., Thyer, M., Kavetski, D., Lerat, J., & Kuczera, G. (2017). Improving probabilistic prediction of daily streamflow by identifying P areto optimal approaches for modeling heteroscedastic residual errors. Water Resources Research, 53(3), 2199-2239

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. Water resources research, 26(7), 1465-1473.

O'Sullivan, C. M., Ghahramani, A., Deo, R. C., Pembleton, K., Khan, U., & Tuteja, N. (2022). Classification of catchments for nitrogen using Artificial Neural Network Pattern Recognition and spatial data. Science of The Total Environment, 809, 151139.

Peng, S., & Chen, H. Y. (2021). Global responses of fine root biomass and traits to plant species mixtures in terrestrial ecosystems. Global Ecology and Biogeography, 30(1), 289-304.

Shen, L. Q., Amatulli, G., Sethi, T., Raymond, P., & Domisch, S. (2020). Estimating nitrogen and phosphorus concentrations in streams and rivers, within a machine learning framework. Scientific data, 7(1), 1-11.

State of Queensland Department of Environment and Resource Management (2012) State Surfacewater Ambient Water Quality Network WMP014 version 2;

State of Queensland Department of Natural Resources, Mines and Energy (2018) Surface Water Ambient Network (Water Quality) 2018-19, WMP014 version 6 June 2018

Sudheer, K. P., Nayak, P. C., & Ramasastri, K. S. (2003). Improving peak flow estimates in artificial neural network river flow models. Hydrological Processes, 17(3), 677-686.

Supplementary Material S2: Evaluation of spatial features, and corroborated classification

Normanby.

Pioneer is the classifier to Norman using the non-partitioned dataset. While partitioning the dataset into hydrograph/flow and season retains Pioneer as the classifier for the Wet Season, the partitioned dataset for Dry Season instead identifies Burnett as a classifier, as well as Barron with Haughton and then Mary for increasing and Retreating flows datasets respectively. Interrogation of the Land Use features heatmap showed in comparison to Normanby, Pioneer had most similar spread of data away from the mean for production on natural environments. Meanwhile, Pioneer did not share any Ecounit or original vegetation feature data spread from the mean with Normanby. Rather, Burnett shared the greatest number of similar mapped Original Vegetation variable types with Normanby, these are all open forest structure on varying geologies. Of the Ecounit features, Burnett and Normanby shared deviation from the mean for production on natural environments on open woodlands (12.2 and 17.2) and dryland agriculture and irrigation on open woodlands (9.3 and 11.4) plus waterbodies on woodland drainage (16.6) for Barron. Notable spatial feature similarities with Normanby were not identified for Mary or Haughton, however similar features were observed for Barron in the Ecounit datasets in both ANN-PR and feature matching. In feature matching the similarities in deviation from the mean were for dryland agriculture on open forests on coastal lowlands and ranges.

Barron.

Partitioning the datasets show for Increasing flows, Burnett, South Johnstone and Pioneer classify for SFBF, SF and BF respectively. The Land Use heatmap showed that South Johnstone and Tully had the most similar combination of deviation from the mean as Barron for conservation on natural environments, irrigated ag, and intensive uses, however Barron additionally deviated from the mean for production on natural environments. Barron and Tully share the same deviations from the mean for Original Vegetation features. For Ecounits, Barron shares the most Ecounit features with Tully, then with Burnett, which are all production from natural environment on lands originally vegetated with open forest types. While conservation on natural environments in vine forests in the wet tropics or semideciduous are shared with South Johnstone. Burnett and Barron shared production on natural environments and waterbody on open woodlands, and Production on natural environments on open forest along drainage lines was shared for Pioneer.

North Johnstone,

South Johnstone was always classified in all datasets, however, splitting reveals similarities with Pioneer in Increasing SFBF dataset. North Johnstone shares Land Use data features with South Johnstone and Tully for conservation on natural environments. Both South Johnstone and Tully also include deviations for irrigated agriculture not shared by North Johnstone. South Johnstone was the only catchment that shared similar deviation from the mean for mapped Original Vegetation features. South Johnstone shared mostly the same deviation from mean for Ecounit features with the exceptions of conservation in originally vine forest environments shared by North Johnstone and Tully. Unlike SouthJohnstone, North Johnstone and Tully also shared intensive uses in the wet tropics. North Johnstone and Pioneer doesn’t share any deviations from the mean for any deductive features.

For North Johnstone, suitable catchments to classify share the deviation from the mean for conservation on natural environments for notophyll vine and microphyll fern forest on high peaks, as well as conservation, production on natural and dryland agriculture on wet tropics areas.

South Johnstone

South Johnstone was always matched to Barron. Below the 50% plus confidence interval there was tendency towards Tully for SFBF in Dry Season only, and North Johnstone in Wet Season and increasing flows dataset partitions.

Tully is the only catchment that shared the standard deviation spread for Land Use data with South Johnstone. This was for conservation in natural environments. South Johnstone had most similar deviations from mean for mapped Original Vegetation with North Johnstone for wet tropics vegetation and vine forest on high peaks and plateaus. South Johnstone shared these similarities with Tully, however Tully also had deviations from the mean for more coastal type vegetation on depositional or sandy areas. While Ecounits deviation spread was shared with North Johnstone for all SHAP-AD Land Uses spatial variables in the wet tropics catchments, and Tully except for intensive uses in the wet tropics. Barron and Marry also shared production on Natural Environments, on semi deciduous vine forests.

For South Johnstone, catchments that share production in natural environments on semi deciduous vine forest are a classifier, with catchments that also share costal environs for Dry Season records, and wet tropics for Wet Season and increasing flow records. Understanding of the dynamics that makes Pioneer's surface flows suitable match in the Dry Season, but its baseflows suitable in increasing flows could benefit further understanding the dynamics that drive the pattern similarities with South Johnstone.

Tully

This catchment was matched to North Johnstone for all data, and Pioneer in Wet Season and Increasing flows. The deviation spread was most similar with South Johnstone in Land Use for conservation of natural areas, Barron for Original Vegetation, and for Ecounits. Similarity was shared with conservation in natural environments in wet tropics for all mapped land uses except intensive uses. Tully also had a combination of other Ecounit variables not shared with any other catchment.

For Tully, Barron spatial data cooborated with the inductive classifiers. The fact that SF for Tully matched with Pioneer in both Wet Season and increasing Baseflows, but didn't match patterns in the Dry Season or Retreating baseflows suggests that the drivers of surface flows are important to consider during the Wet Season and increasing baseflows, and catchment similarities between Tully and Pioneer may not involve baseflow dynamics. While the relationship to North Johnstone throughout all seasons and flow events may show relationship to baseflow dynamics.

Haughton

Always classified to O'Connell, however splitting revealed pattern matches also with Burnett during dry and below flow events for SF data. Land Use SHAP-AD was most similar with Haugton for Pioneer and Normanby. These catchments shared similar deviations for production on natural environments, water bodies and unknown. For Original Vegetation datasets, Pioneer shared most similar combination of deviation for tall woodlands, and open woodlands on hilly metamorphic and acid igneous rocks as well as open forest on drainage lines and alluvial planes. Differences were Pioneer had semievergreen vine forest, while Haughton had greater spread of deviation of complex vine forest. Pioneer and Haughton share the same deviations for unknown and production on natural environments on areas mapped with the same Original Vegetation that is . O'Connell shared irrigated agriculture on microphyll vine forest.

For Haughton, the only co-oborated deductive classifiers are the similarities for irrigated agriculture on microphyll vine forest. No clear similarities were shared for Burnett. The variation between suitable classifiers in Wet Season/increasing flows, and Dry Season/Retreating flows shows that for Haughton, different catchments are relevant to classify under the differing flow regimes. With no relationship to the spatial data evaluated, the dynamics driving Haughton's differing suitability for classification under different flow regimes and seasons need to be further evaluated.

O'Connell

Matched to Plane for Wet Season and Increasing flows. However, as with Haughton, it also shows match to Burnett during Retreating flows and dry seasonss. Deviation from means on the heatmap matched with Plane for dryland agriculture. O'Connell also had deviations for conservation on natural environments not shared with Plane. For Original Vegetation spatial variables, Burnett shared the most similar deviation from the mean for open forests on coastal lowlands, while Normanby and Mary had larger devations for the same Original Vegetation type. For Ecounit data, Plane showed the most features with similar deviations from the mean. These are conservation and production on natural environments and irrigation on areas originally vegetated with notophyll or microphyll vine forests.

For O'Connell deductive classifiers show low deviation from the mean for open forest on coastal lowland and production on natural environments for dry periods, with larger deviations for the combination of both production and irrigation on natural environments for Wet Season and increasing flow times periods on areas originally vegetated with notophyll or microphyll vine forests.

Pioneer

No patterns matches were stronger than the 50% plus confidence interval line. While the tendency was towards Burnett in the non-partitioned, Wet Season and increasing flow dataset, the matches also tended toward Plane in the Dry Season and Retreating flow periods. Pioneer feature matched with Haughton and Normanby for waterbodies and production on natural environments. Original Vegetation was shared with Haughton, Normanby and Burnett. Feature matching was most similar with Haughton for open woodlands on hilly terrain of metamorphic and igneous rock geology and open forest along alluvial drainage lines. With both Normanby and Burnett for semi evergreen microphyll vine thickets. Similarities for deviations from the mean were closest to Burnett for notophyll to microphyll vine forest original vegetation, and most similar with Normanby for notophyll and mesophyll vine forests with palms along alluvia and streamlines or sandmasses. For the Ecounit data heat map similaritites were shared with Plane, O'Connell and Haughton for irrigated agriculture on previous notophyll to microphyll vine forests, and with Plane for intensive uses on the same original vegetation type. SHAP-AD evaluation found Burnett was mmost siilar for conservation on corymbia dominated open woodlands on undulating terrain. Similarities were shown for Mary, Burnett and Barron for production on open forest drainage lines on alluvial areas. Haughton had the most similar deviation for this feature followed by Mary, then Burnett then Barron.

For Pioneer, semi evergreen microphyll vine thickets, notophyll to microphyll vine, conservation on corymbia dominated open woodlands on undulating terrain production on open forest drainage lines on alluvial areas are evident for classifiers during Wet Season or increasing flows. During Dry Season and Retreating flows the deductive classifiers match to irrigated agriculture and intensive uses on previous notophyll to microphyll vine forests.

In dry and Retreating flow situation, catchments that share the deviation from the mean for moist to dry open forest on basalt areas, production on natural environments for notophyll to microphyll vine forests, or irrigation and intensive uses on the same original vegetation type may be suited to classify to Plane.

Plane

Non partitioned dataset displayed no pattern matches stronger than the 50% plus CI, however partitioning emphasised the similarity of the patters towards Burnett and Pioneer for the Dry Season and Retreating flow records. Feature matches were shared with O'Connell for production on natural environments, while Pioneer, Haughton, Normanby and Burnett shared the absence conservation dominance. Normanby and Burnett shared similar deviations from the mean only for moist to dry open forest on basalt areas, with the similarity strongest for Burnett. In regards to devations from the mean for Ecounit spatial variables Burnett only shared production on natural environments for notophyll to microphyll vine forests, while Pioneer shared the similarity for irrigation and intensive uses on the same original vegetation type. O'Connell also shared four similar deviations for Ecounit features being conservation, production and irrigation on notophyll to microphyll vine forests, and conservation of moist to dry open forest on coastal lowlands.

Burnett

Did not pattern match stronger than the 50% CI. Splitting revealed the strongest and only match beyond the 50% CI for BF with Pioneer. Tendency also towards Barron in Dry Season and Retreating flow and dry periods the deviation from the mean is for Land Use features of production on natural environments with Pioneer having the closest similarity in deviation. For Original Vegetation datasets, Burnett had the widest spread of deviation from the mean for the majority of original vegetation types. They were most similar to Pioneer for notophyll and microphyll vineforests and weakly for woodland on metamorphic and igneous rocks, while open forest on a range of coastal lowland, hills, basalt and shallow soil on weathered rocks, and drainage lines on alluvial planes were shared for Burnett and Normanby. No deviations from the mean were shared for Barron. For Ecounits however Barron shared deviations from the mean for production on natural environments for open forest on coastal lowlands, basalt areas, metamorphic and igneous rocks, drainage lines and alluvial planes and sand or depositional plains. While Normanby shared the deviation for the same landuse of production instead on open woodland on hilly weathered rocks, and sandplain or depositional areas, with conservation of corymbia open woodland on hilly areas. Pioneer only shared standard deviations for conservation on corymbia open forests on undulating terrain.

For Burnett, in Retreating flow events inductive classification results may align partially to catchments with with the spread away from the mean areas for production on natural environments on notophyll and microphyll vineforests and weakly for woodland on metamorphic and igneous rocks, and conservation on corymbia open forests on undulating terrain. Catchments with production on natural environments for open forest on coastal lowlands, basalt areas, metamorphic and igneous rocks, drainage lines and alluvial planes and sand or depositional plains may also be suitable to classify to Burnett for these Retreating and dry events. In increasing and wet events open forest on a range of coastal lowland, hills, basalt and shallow soil on weathered rocks, and drainage lines on alluvial planes and production instead on open woodland on hilly weathered rocks, and sandplain or depositional areas, with conservation of corymbia open woodland on hilly areas.

Mary

Only Pioneer had pattern matches with Mary that exceeded 50% plus the confidence interval. These only occurred in the non-partitioned dataset, Wet Season and increasing flow scenarios. While below the 50% confidence interval, Burnett had roughly equal pattern matches with Mary and Pioneer during Dry Season and Retreating flow events. Mary had the most similar deviations from the mean with Burnett for production on natural environment landuse. These similarities in Original Vegetation features with Burnett also occurred for open woodland on coastal lowlands, and weakly for wetlands. Mary only shared similar deviations for Original Vegetation data with Pioneer for notophyll and mesophyll vine forest with palms on alluvia streamlines and swamps. For Ecounit features, Mary shares standard deviation similarities with Burnet for production on natural environments for open woodland on coastal lowlands, corymbia open woodland on hilly areas, open woodland on weathered rocks, and hilly metamorphic and acid igneous rocks as well as tall forest along drainage lines and alluvial planes. Mary also shared standard deviation spread for the variety of landuses across semi deciduous mesophyll to notophyll vineforests in the wet tropics catchments, i,e, Barron, Tully, North and South Johnstone.

These results show catchments may be suitable to classify to Mary in increasing flows and the Wet Season for notophyll and mesophyll vine forest with palms on alluvia streamlines and swamps. Meanwhile during Dry Season and Retreating flow events the classification could also include catchments that share deviations from the mean with Mary for production on natural environments for open woodland on coastal lowlands, corymbia open woodland on hilly areas, open woodland on weathered rocks, and hilly metamorphic and acid igneous rocks as well as tall forest along drainage lines and alluvial planes and weakly for wetlands.

APPENDIX C: Journal Paper 3 – Supplementary Material

Supplementary Material S1: Abbreviations

A=All

ANN = Artificial Neural Network

C=Catchment

Category 1=Catchments with similar DIN patterns during increasing flows and rainy season.

Category 2=Catchments with year round similar DIN patterns.

Category 3= Catchments with similar DIN patterns during retreating flows and dry season.

CM=Mary Catchment

d=Willmotts Index

DIN = Dissolved Inorganic Nitrogen

EU=Ecounits

F=Flows

F1=Category 1 flows (Wet season/increasing flows)

G=Gauged

LU=Land use

Match=Catchments paired together for their similarities

MSE=Mean Square Error

NSE=Nash Sutcliffe Efficiency

obs=observed data

OV=Original Vegetation

OW=Open Woodlands

pde=Peak Percentage Deviation

PR=Pattern Recognition

ReLU=Rectified Linear Units

R^2=regression coefficient

RMSE=Root Mean Square Error

SHAP= Shapley Additive exPlanations

sim=simulation

WQ=Water Quality

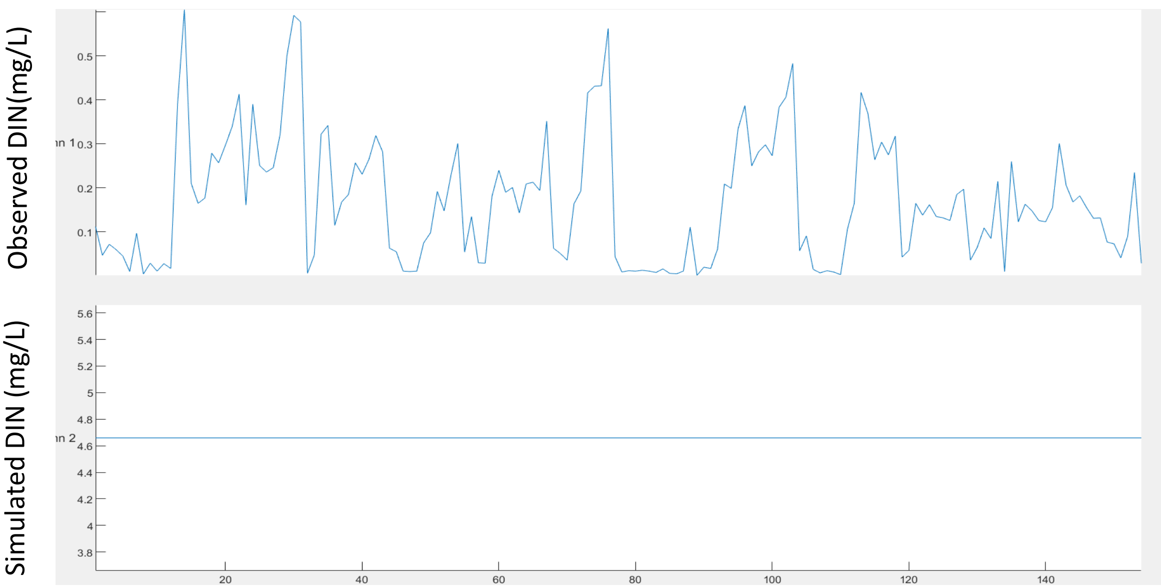
WT=Wet Tropics

XAI=eXplainable Artificial Intelligence

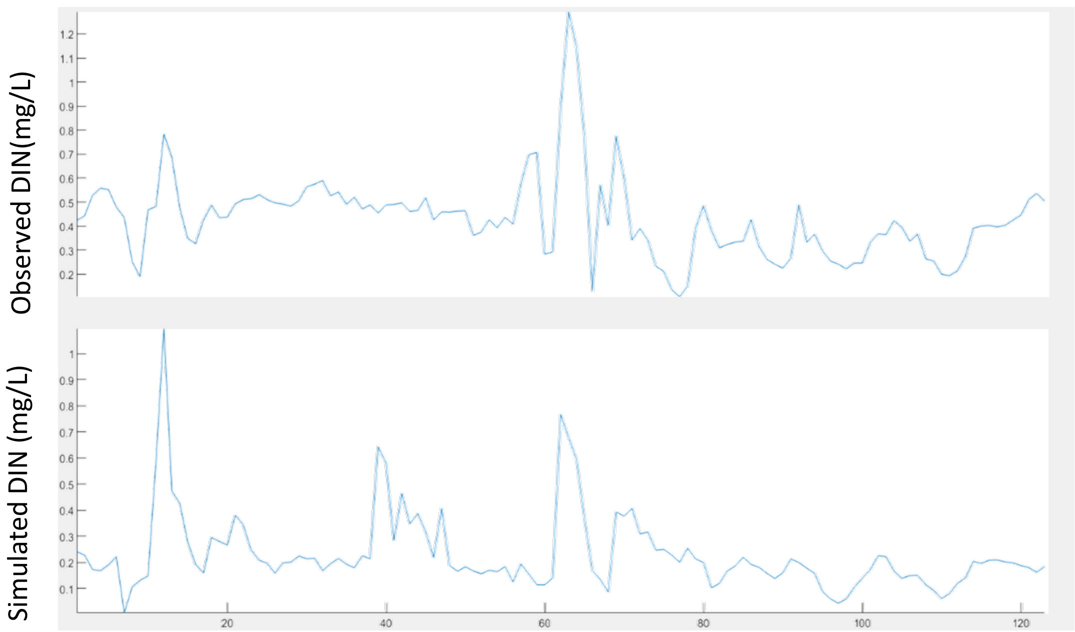
Supplementary Material Figure SF1:

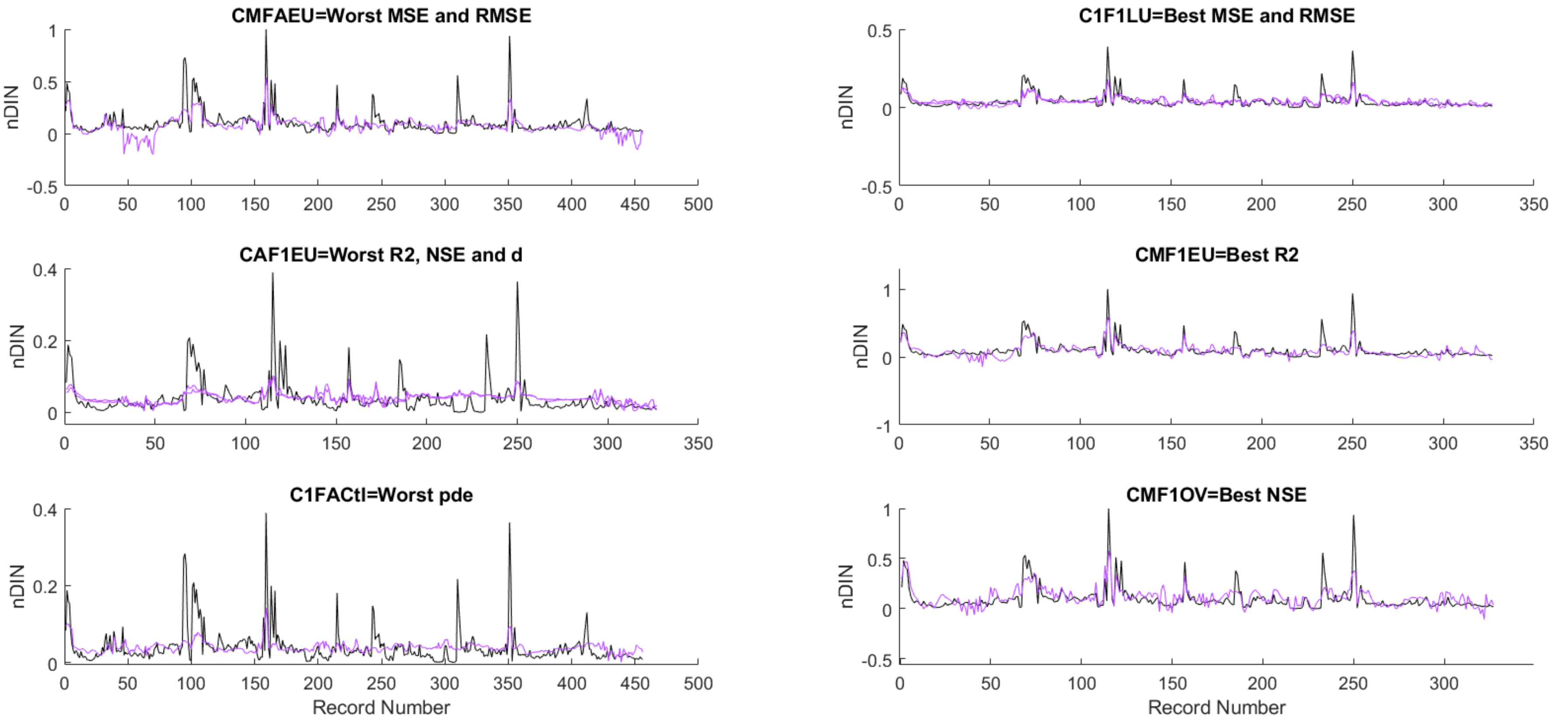
ANN_WQ simulator development. Example simulation results during code development for DIN for individual catchments identified as a Category 1 spatio temporal catchment vs Category 2 spatio temporal catchment. Simulations flatlined for testing of Category 1 catchments trained on their own data, while simulations were possible for Category 2 catchments.

a) Category 1 spatio temporal catchment



b) Category 2 spatio temporal catchment





Supplementary Material Figure SF2: Visualisation of best and worst performing scenarios for DIN simulation of Herbert. Blackline represents observed (true) data, purple lines are the simulated data for scenario that matched the best/worst performance criteria, as identified by the heading for each graph. C= catchment data included in ANN_WQ simulator training scenario, F= flow data included in ANN_WQ simulator training scenario, M = Mary Catchment Data, A=All catchment data, 1 = increasing flows and wet season data, EU= Ecounit, OV= Original Vegetation, LU= Land use,. MSE = Mean Square Error, RMSE= Root Mean Square Error, R2= Correlation Coefficient, NSE= Nash-Sutcliffe coefficient, d= Willmotts Index.

Supplementary Material Table ST1: Kruskal Wallis Test for independence of performance criteria distribution outputs from the ANN_WQ simulator. Independence exists where datasets are first discriminated to only include catchments and flow data representative of the spatio temporal regime. Inclusion of spatial data have insignificant influence where included in the training dataset.

Hypothesis	X	Sig.[a,b]	Decision
Category Groups (i.e. All, Cat 1, Cat 2, Cat 3 and Herbert Trial) have same distribution of performance criteria X	MSE	0.025	***Reject the null hypothesis.***
	R^2	0.003	***Reject the null hypothesis.***
	NSE	0.003	***Reject the null hypothesis.***
	d	0.003	***Reject the null hypothesis.***
	RMSE	0.025	***Reject the null hypothesis.***
	pde	0.045	***Reject the null hypothesis.***
Spatial Data Scenarios (i.e. Control, EU, LU, OV, All) have same distribution of performance criteria X	MSE	0.725	Retain the null hypothesis.
	R^2	0.951	Retain the null hypothesis.
	NSE	0.981	Retain the null hypothesis.
	d	0.967	Retain the null hypothesis.
	RMSE	0.725	Retain the null hypothesis.
	pde	0.453	Retain the null hypothesis.

a. Significance level is 0.05

b. Asymptotic significance is displayed

Supplementary Material Table ST2: Performance metrics for simulations of DIN for the pseudo ungauged catchment (Herbert) generated from an ANN_WQ simulator developed using the differing training dataset scenarios. Scenario abbreviation: CAll= data from all catchments, CM=data from Mary only, FAll = all flow data included, F1= flows discriminated to Category 1 spatio temporal regime of Wet Season Increasing Flows, Ctl = no spatial data included, EU= Ecounit, LU= Land Use, OV= Original Vegetation spatial data included. Performance metrics and corresponding scenario styled in: ***red bold italic*** = best performing scenario; **black bold** = poorest performing scenario; grey= failed to meet minimum standard for the corresponding performance metric. Results show scenarios trained on the single catchment matched using ANN_PR for Ecounits (Mary) and flow variables discriminated to the Category 1 spatio temporal regime of Wet Season Increasing Flows collectively achieved the best performing performance for R2 (EU), NSE (OV) and Willmots d efficiency and pde (LU). With the exception of C1F1LU and C1F1OV, and CAllFAllOV, training datasets only including data from the individual matched catchment achieved satisfactory NSE scores.

Scenario	Category	Flows	Spatial Scenario	Best score from the 1000hn trial for : MSE	R2	NSE	d	RMSE	pde
C1F1Ctl	1	1	Ctl	0.00126	0.67389	0.44019	0.69366	0.03556	-1.282
C1F1EU	1	1	EU	0.00127	0.69564	0.43751	0.69752	0.03565	45.8329
C1F1LU	1	1	LU	***0.00099***	0.78699	0.5601	0.78032	***0.03153***	21.198
C1F1OV	1	1	OV	0.00108	0.77024	0.52022	0.74909	0.03292	-2.5231
C1FAllCtl	**1**	**All**	**Ctl**	0.00147	0.59903	0.29789	0.51612	0.03831	**58.1074**
C1FAllEU	1	All	EU	0.00128	0.65115	0.38713	0.6672	0.0358	-32.352
C1FAllLU	1	All	LU	0.00112	0.71513	0.46336	0.70666	0.0335	48.6003
C1FAllOV	1	All	OV	0.00117	0.68981	0.44003	0.70817	0.03422	49.7629
CAllF1Ctl	All	1	Ctrl	0.00155	0.62667	0.31176	0.55666	0.03943	44.3145
CAllF1EU	**All**	**1**	**EU**	0.00163	**0.58198**	**0.277661**	**0.51364**	0.04040	4.47167
CAllF1LU	All	1	LU	0.00149	0.6381	0.33853	0.59382	0.03866	40.5623
CAllF1OV	All	1	OV	0.00155	0.61461	0.31426	0.55621	0.03936	9.97363
CAllFAllCtl	All	All	Ctl	0.00116	0.70186	0.44447	0.70299	0.03408	40.0596
CAllFAllEU	All	All	EU	0.00109	0.72895	0.47703	0.7417	0.03307	0.83082
CAllFAllLU	All	All	LU	0.00106	0.75433	0.49383	0.73251	0.03253	30.7716
CAllFAllOV	All	All	OV	0.00102	0.74994	0.50971	0.74708	0.03201	13.9193
CMF1Ctl	Mary	1	Ctl	0.00607	0.77443	0.59629	0.84445	0.07791	-0.0109
CMF1EU	***Mary***	***1***	***EU***	0.00586	***0.79875***	0.61034	0.83047	0.07654	-0.389
CMF1LU	***Mary***	***1***	***LU***	0.00574	0.78762	0.61823	***0.84946***	0.07576	***-0.012***
CMF1OV	***Mary***	***1***	***OV***	0.00571	0.79165	***0.62034***	0.84499	0.07555	0.11908
CMFACtl	Mary	All	Ctl	0.00673	0.72241	0.51473	0.78144	0.08201	0.08782
CMFAEU	Mary	All	EU	**0.00679**	0.74263	0.51034	0.77184	**0.08238**	-2.1616
CMFALU	Mary	All	LU	0.00642	0.7459	0.53679	0.78972	0.08012	-10.167
CMFAOv	Mary	All	OV	0.0065	0.73849	0.53101	0.78007	0.08062	16.3765

Supplementary Table ST3: Gauging allocation, DIN, and flow data availability for each of the catchments[75,76].

Catchment flowing to Great Barrier Reef	Original Vegetation Data	Landuse Data	Gauging allocation	Gauging station ID for observed data	Gauging station Latitude (decimal °)	Gauged station Longitude (decimal °)	Catchment area (km²)	Natural Resource Management Region	DIN (mg/L) at daily streamflows (averaged from hourly)for:			Mean DIN	Standard Deviation of DIN (mg/L)	DIN Record Period	DIN sampling frequency	Number of records in DIN record period
									Min	Max	Mode					
CurtisIsland	✓	✓	Ungauged	-	-	-	564	Fitzroy	-	-	-		-	-	-	-
Jardine	✓	✓	Ungauged	-	-	-		Cape York	-	-	-		-	-	-	-
JackyJacky	✓	✓	Ungauged	-	-	-	3,102	Cape York	-	-	-		-	-	-	-
GOlivePascoe	✓	✓	Psudo-Ungauged	102102A	-12.657785	143.050145	132	Cape York						-	-	-
UGOlivePascoe	✓	✓	Ungauged	-	-	-		Cape York	-	-	-		-	-	-	-
Gstewart	✓	✓	Psudo-Ungauged	104001A	-14.167489	143.394002	471	Cape York						-	-	-
UGStewart	✓	✓	Ungauged	-	-	-	2,342	Cape York	-	-	-		-	-	-	-
GNormanby	✓	✓	Gauged	105107A	-15.46	144.56	12,828	Cape York	0.012 0.011 0.069 0.017 0.006 0.016 0.005	0.013	148.24	0.055	0.128	3/10/2006- 25/08/2017	Events Jan-March	244
UGNormanby	✓	✓	Ungauged	-	-	-	11,992	Cape York	-	-	-		-	-	-	-
Jeannie	✓	✓	Ungauged	-	-	-	3,711	Cape York	-	-	-		-	-	-	-
Endeavour	✓	✓	Ungauged	-	-	-	2,214	Cape York	-	-	-		-	-	-	-
Gdaintree	✓	✓	Psudo-Ungauged	1080025A	-16.1796	145.2819	911	Wet tropics						-	-	-
UGDaintree	✓	✓	Ungauged	-	-	-	1,217	Wet tropics	-	-	-		-	-	-	-
Mossman	✓	✓	Ungauged	-	-	-	475	Wet tropics	-	-	-		-	-	-	-
GBarron	✓	✓	Gauged	110001D	-17.05	145.51	1,950	Wet tropics	0.0045	0.154	0.004 0.235	0.117	0.091	19/01/06- 15/09/17	Regular (1) monthly, Events Jan-March	318
UGBarron	✓	✓	Ungauged	-	-	-	250	Wet tropics	-	-	-		-	-	-	-
GMulgraveRussell	✓	✓	Psudo-Ungauged	111007A	-17.133361	145.764556	523.19	Wet tropics	-	-	-		-	-	-	-
MulgraveRussell	✓	✓	Ungauged	-	-	-		Wet tropics	-	-	-		-	-	-	-
G North Johnstone	✓	✓	Gauged	112004A	-17.5	145.69	926	Wet tropics	0.0035	0.157		0.147	0.073	30/01/2006- 15/09/2017	Infrequent half yearly, Events Jan-March	94
GSouth Johnstine	✓	✓	Gauged	112101B	-17.66	145.77	399	Wet tropics	0.009	0.044	0.103 0.050	0.126	0.065		Regular (1) monthly, Events Jan-March	414
UGJohnstone	✓	✓	Ungauged	-	-	-	875	Wet tropics	-	-	-		-	-	-	-
GTully	✓	✓	Gauged	113006A	-17.87	145.72	1,386	Wet tropics	0.008 0.090	0.062	0.270 0.225	0.237	0.154	13/01/2006- 19/04/2018	Frequent (>1) monthly,	723

Catchment flowing to Great Barrier Reef	Original Vegetation Data	Landuse Data	Gauging allocation	Gauging station ID for observed data	Gauging station Latitude (decimal °)	Gauged station Longitude (decimal °)	Catchment area (km²)	Natural Resource Management Region	DIN (mg/L) at daily streamflows (averaged from hourly)for:			Mean DIN	Standard Deviation of DIN (mg/L)	DIN Record Period	DIN sampling frequency	Number of records in DIN record period
									Min	Max	Mode					
															Events Jan-March	
UGTully	✓	✓	Ungauged	-	-	-	298	Wet tropics	-	-	-		-	-	-	-
Murray	✓	✓	Ungauged	-	-	-	1,107	Wet tropics	-	-	-		-	-	-	-
Gherbert	✓	✓	Psudo-Ungauged	116006B	-18.488994	145.936037	7,490	Wet tropics		0.084		0.186002571	0.213	-	-	-
UGHerbert	✓	✓	Ungauged	-	-	-	2,348	Wet tropics	-	-	-		-	-	-	-
Black	✓	✓	Ungauged	-	-	-	1,053	NQ Dry Tropics	-	-	-		-	-	-	-
Ross	✓	✓	Ungauged	-	-	-	1,696	NQ Dry Tropics	-	-	-		-	-	-	-
GHaughton	✓	✓	Gauged	119003A	-19.72	146.81	1,807	Burdekin	0.008	0.252	0.008	0.066	0.088	20/12/2012-25/09/2017	Regular monthly, Events Jan-March	80
UGHaughton	✓	✓	Ungauged	-	-	-	2,211	Burdekin	-	-	-		-	-	-	-
UGBurdekin	✓	✓	Ungauged	-	-	-	128445**	Burdekin	-	-	-		-	-	-	-
Don	✓	✓	Ungauged	-	-	-	3,698	Mackay Whitsunday	-	-	-		-	-	-	-
Proserpine	✓	✓	Ungauged	-	-	-	2,466	Mackay Whitsunday	-	-	-		-	-	-	-
GOConnell	✓	✓	Gauged	124001B	-20.77	0.008 0.090	336	Mackay Whitsunday	0.008 0.090	0.062	0.008 0.090	0.109	0.14	25/01/2007-24/08/2017	Irregularly One off Events Jan-March	87
UGOConnell	✓	✓	Ungauged	-	-	-	2,021	Mackay Whitsunday	-	-	-		-	-	-	-
GPioneer	✓	✓	Gauged	125013A	-21.23	148.74	1,464	Mackay Whitsunday	0.021 0.071 0.823 0.914 0.624 0.012 0.011	0.061	0.021 0.071 0.823 0.914 0.624 0.012 0.011	0.231	0.275	18/10/2006-13/09/2017	Frequent (>1) monthly, Events Jan-March	402
UGPioneer	✓	✓	Ungauged	-	-	-	87	Mackay Whitsunday	-	-	-		-	-	-	-
GPlane	✓	✓	Gauged	126001A	-21.24	148.94	327	Mackay Whitsunday	0.009 0.013	0.046	0.961 1.287 1.265	0.424	0.533	4/09/2009-26/08/2017	Regular (1) monthly, Events Jan-March	302
UGPlane	✓	✓	Ungauged	-	-	-	2,173	Mackay Whitsunday	-	-	-		-	-	-	-
Styx	✓	✓	Ungauged	-	-	-	2,959	Fitzroy	-	-	-		-	-	-	-
Shoalwater	✓	✓	Ungauged	-	-	-	3,535	Fitzroy	-	-	-		-	-	-	-
Waterpark	✓	✓	Ungauged	-	-	-	1,797	Fitzroy	-	-	-		-	-	-	-
UGFitzroy	✓	✓	Ungauged	-	-	-	139544*	Fitzroy	-	-	-		-	-	-	-
Calliope	✓	✓	Ungauged	-	-	-	2,193	Fitzroy	-	-	-		-	-	-	-
Boyne	✓	✓	Ungauged	-	-	-	2,441	Fitzroy	-	-	-		-	-	-	-
Baffle	✓	✓	Ungauged	-	-	-	3,992	Burnett Mary	-	-	-		-	-	-	-

Catchment flowing to Great Barrier Reef	Original Vegetation Data	Landuse Data	Gauging allocation	Gauging station ID for observed data	Gauging station Latitude (decimal °)	Gauged station Longitude (decimal °)	Catchment area (km^2)	Natural Resource Management Region	DIN (mg/L) at daily streamflows (averaged from hourly)for:			Mean DIN	Standard Deviation of DIN (mg/L)	DIN Record Period	DIN sampling frequency	Number of records in DIN record period
									Min	Max	Mode					
Kolan	✓	✓	Ungauged	-	-	-	2,838	Burnett Mary	-	-	-		-	-	-	-
GBurnett	✓	✓	Gauged	136007A	-25.73	151.28	30,724	Burnett Mary	0.004	0.281	0.004 0.119	0.161	0.318	23/10/20006- 15/09/2017	Frequent (>1) monthly, Events Jan-March	400
UGBurnett	✓	✓	Ungauged	-	-	-	1,675	Burnett Mary	-	-	-		-	-	-	-
Burrum	✓	✓	Ungauged	-	-	-	3,293	Burnett Mary	-	-	-		-	-	-	-
GMary	✓	✓	Gauged	138014A	-26.19	152.49	6,863	Burnett Mary	0.017 0.061	0.236	0.017 0.061	0.201	0.204	25/09/2013- 29/06/2018	Frequent (>1) monthly, Events Jan-March	176
UGMary	✓	✓	Ungauged	-	-	-	2,372	Burnett Mary	-	-	-		-	-	-	-

Citation:

75. State of Queensland Department of Environment and Resource Management (2012) State Surface water Ambient Water Quality Network WMP014 version 2.
76. State of Queensland Department of Natural Resources, Mines and Energy (2018) Surface Water Ambient Network (Water Quality) 2018-19, WMP014 version 6, June 2018.