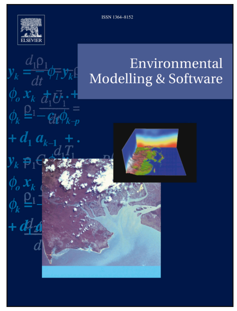
Journal Pre-proof



A pragmatic parameterisation and calibration approach to model hydrology and water quality of agricultural landscapes and catchments

Afshin Ghahramani, David M. Freebairn, Dipaka R. Sena, Justin L. Cutajar, David M. Silburn

PII: S1364-8152(18)31096-X

DOI: https://doi.org/10.1016/j.envsoft.2020.104733

Reference: ENSO 104733

To appear in: *Environmental Modelling and Software*

Received Date: 10 November 2018

Revised Date: 20 March 2020

Accepted Date: 22 April 2020

Please cite this article as: Ghahramani, A., Freebairn, D.M., Sena, D.R., Cutajar, J.L., Silburn, D.M., A pragmatic parameterisation and calibration approach to model hydrology and water quality of agricultural landscapes and catchments, *Environmental Modelling and Software* (2020), doi: https://doi.org/10.1016/j.envsoft.2020.104733.

A pragmatic parameterisation and calibration approach to model hydrology and water quality of agricultural landscapes and catchments

Afshin Ghahramani*[1], David M Freebairn[2], Dipaka R Sena[3], Justin L Cutajar[4], David M Silburn[2,5]

[1]University of Southern Queensland, Institute for Life Sciences and the Environment, Centre for Sustainable Agricultural Systems, Toowoomba, QLD, 4350, Australia

[2]University of Southern Queensland, Centre for Agricultural Engineering, Toowoomba, QLD, 4350, Australia

[3]ICAR-Indian Institute of Soil & Water Conservation, 218, Kaulagarh Road, Dehradun, Uttarakhand 248195, India

[4]Hydrobiology, Auchenflower, QLD 4066, Australia
[5]Department of Natural Resources, Mines and Energy, Toowoomba, QLD 4350, Australia

*Corresponding author: Afshin Ghahramani

Address: University of Southern Queensland, Institute for Agriculture and the Environment, Toowoomba, Queensland, 4350, Australia

Tel: T: +61 7 4631 1178, Email: Afshin.Ghahramani@usq.edu.au

Abstract

Australian and Queensland Government's Reef 2050 Water Quality Improvement Plan has set targets for improving the water quality entering the Great Barrier Reef lagoon. Given the large public investment and the deficit of data linking on-farm land management to changes in environmental outcomes, there is a need for a robust and efficient methods of quantifying links between land management and water quality. This paper explores a pragmatic approach to making this link using available data. We demonstrate that a simple parameterisation process is suitable for estimating hydrology and water quality across a wide range of land uses and management practices in agricultural landscapes. However, a manually calibrated model may still require the analysis of parameters to reduce error variances and evaluate uncertainties. Confidence in estimating hydrology and water quality in descending order is: runoff, sediment, nitrogen, phosphorous, and pesticide losses, reflecting the availability of data and inherent error propagation.

Keywords:

Great Barrier Reef (GBR), hydrological modelling, water quality, HowLeaky, nitrogen, phosphorus, pesticide, PEST, calibration, uncertainty, sensitivity

1. Introduction

Sediment, nutrients, and pesticides in runoff from agricultural landscape have impacts on the water quality and ecological functions of receiving environments (Wooldridge 2009, De'ath and Fabricius, 2010). For example, deteriorating water quality associated with agricultural enterprises (sediment, nitrogen, phosphorus, and pesticides) poses a high risk to the Great Barrier Reef (GBR), one of Australia’s iconic natural assets (Waterhouse et al., 2017). Thus, Australian and Queensland Governments Reef 2050 Water Quality Improvement Plan program (The State of Queensland Government, 2018) has set water quality targets for the Great Barrier Reef lagoon. Over 2005-2015, a significant public investment, i.e. $A 2 billion, has been spent to improve water quality entering the GBR lagoon (Commonwealth of Australia, 2015). It has been suggested that this investment requires an increase in the order of four times over 2015-2025 to reach water quality targets set out by the Water Quality Improvement Plan (The State of Queensland Government, 2018). A substantial portion of this investment is targeted at better management of agricultural landscapes and, considering the current “deficit of information publicly available”, linking and quantifying on-farm management to water quality improvement is critical for on-going support.

Hydrological and water quality experimental datasets across Australian agriculture are sparse and in most cases incomplete, causing difficulties in making an assessment of linkages between management and water quality based on empirical evidence. This paucity of data is related to the high cost and necessary long term observations from field studies to sample the variable nature of climates (Freebairn and Wockner, 1986a). However, the data are needed to provide multiple studies over a wide range of climatic and environmental conditions for the delivery of credible results. Systematic storage of such data, which is readily accessible, is also a major limitation for any model testing and application.

Objective management of water quality leaving farms requires robust linkages between soil and crop practices and hydrology, sediment, and agri-chemical transport processes. Modelling hydrology and water quality within agricultural systems can be complex and should consider the main interactions while keeping model complexity at a minimum without compromising the reliability of model estimates. Ideally, a model’s complexity should be balanced by the level of data available describing the system and our understanding of how the system operates (Grayson and Blöschl, 2000). In addition, a highly parametrised model is likely to exclude the use of less than complete datasets in its development, testing, and application, potentially weakening the empirical basis for the model. Currently, such models e.g. HowLeaky (Queensland Government, 2019) are being used for evaluation of management scenarios to guide policy and design investment portfolios.

In order to improve the credibility of models to estimate changes in water quality associated with land management changes across diverse environments, and inform policymakers, a process was needed to maximise the use of available data, regardless of completeness. This required a methodology that could accommodate datasets that ranged from detailed daily records and site descriptions to sparse and incomplete data. A model with too many parameters can be endemic and models with too many degrees of freedom incur serious risks (Jakeman et al., 2006).

In this paper, we explore the use of a simple approach for manual model parametrisation and the impact of the level of detail in system specification on the model's ability to represent a range of paddock scale land use. The paper presents and tests a pragmatic approach using a wide range of data quality and detail. We use a daily water balance model to explore how the level of detail in system specification impacts on the model's ability to represent a range of paddock scale land use and management impacts on hydrology and water quality using diverse datasets with variable quality and detail. We also compare this pragmatic approach using expert judgement with an automatic parameter optimisation technique (PEST, Doherty, 2015) to determine whether this might reduce error and uncertainty associated with input specifications.

2. Methods

2.1 HowLeaky model

HowLeaky is a water balance and water quality modelling environment built on the foundations of the PERFECT model (Littleboy et al., 1992). PERFECT has been strongly influenced by the CREAMS (Knisel et al., 1980) and EPIC (Williams, 1983) models. HowLeaky is a one dimensional, daily time step, water balance model with sub-models of the dynamics for soil erosion, phosphorus and pesticide to simulate the quality of water leaving agricultural systems at the paddock or field scale. The modelling environment is used extensively by the Queensland Government to estimate hydrology and water quality in the Great Barrier Reef catchments at the paddock scale (Carroll et al. 2012). It has also been used by other government agencies and consultants in Australia, e.g. in Victoria (Vigiak et al., 2011).

HowLeaky can be configured to describe sequences of different crops and fallow (between crop phases) management practices for a wide range of cropping systems. It is also used as the modelling engine for SoilWaterApp, a successful decision support system used by grain growers across Australia (Freebairn et al., 2018). Numerous publications describe the development, validation and application of HowLeaky, including: defining erosion-productivity relationships (Littleboy et al., 1992b, 1996); evaluating the effects of cropping systems on runoff, recharge, erosion and yield (e.g. Carroll et al., 1992, Abbs and Littleboy 1998); evaluating surface management options (e.g. Cogle et al., 1996); evaluating the effects of crop and pasture rotations on runoff, erosion and recharge

(Lawrence and Littleboy, 1990; Thornton et al., 2007; Silburn et al., 2007; Robinson et al., 2010; Melland et al.,2010); quantitative land evaluation (e.g. Thomas et al., 1995); assessing risk of soil compaction (Littleboy et al., 1998); estimating the hydrological effects of tree clearing (Williams et al., 1997) and design of land-based effluent disposal systems (Gardner et al., 1995).
HowLeaky uses 244 input variables that can be manipulated to varying degrees (The Queensland Government, 2019; also see www.howleaky.net, this will become www.howleaky.com). These include 12-40 values to describe soil water holding capacity: two for evaporation, four for runoff; two for sediment delivery; a set of curves to describe vegetation (green and dead cover and root depth); and several tillage parameters (Littleboy et al. 1989). HowLeaky includes submodels to simulate water balance, crop growth, crop residue balance, soil erosion, nitrate, and pesticide. These sub-models are briefly explained below.

2.1.1 Water balance sub-model

The water balance model used in HowLeaky has evolved from CREAMS (Knisel, 1980) which predicts soil water balance, runoff, and deep drainage from a combination of rainfall and evaporation data using the runoff model of Williams and La Seur (1976) and the soil evaporation model of Ritchie (1972). CREAMS was influential in the development of PERFECT (Littleboy et al., 1992) and later HowLeaky (McClymont et al., 2016). The latter uses the Williams-Ritchie water balance model (Williams and La Seur, 1976; Ritchie, 1972) which is a one-dimensional mechanistic model, with parameterisation strongly based on a wide range of empirical studies (Littleboy et al., 1992; www.howleaky.net). Surface runoff is estimated as a function of daily rainfall using the SCS runoff curve number model (Williams and LaSeur, 1976), soil water deficit, surface residue and crop cover. The model uses a "cascading bucket" structure where infiltration is partitioned into soil layers from the surface, filling subsequent layers to total porosity. Water flux between layers is limited by a specified daily maximum drainage rate and drainable porosity. Soil water can be removed from the profile by transpiration, soil evaporation and downwards movement from the lowest layer as deep drainage. Transpiration is a function of potential evaporation (a climate input), leaf area or percentage green cover and soil moisture. Soil evaporation removes soil water from the upper two layers. The sum of transpiration and soil evaporation (evapotranspiration) cannot exceed potential evaporation on any day.

2.1.2 Soil erosion sub-model

In cropping systems, soil erosion removes soil nutrients and reduces a soil's water-holding capacity and therefore causes damage to the receiving environment. Movement of sediment off-site carries nutrients such as nitrogen and phosphorus, as well as pesticides and solutes. Empirical models have been widely used and are considered appropriate for a wide range of agricultural systems, including analysis of the state of erosion from paddocks and predicting the changes in erosion due to changed land use. HowLeaky calculates soil erosion based on daily runoff amount and a sediment concentration-cover relationship (Freebairn and Wockner, 1986a). The model predicts soil erosion by accounting for changes in ground cover and runoff, both factors that can be controlled through management. Modified Universal Soil Loss Equation (MUSLE) (Wischmeier and Smith 1978) factors for erodibility (K, metric units), slope-length factor (LS), practice (P) and a delivery ratio are included in estimating soil erosion.

2.1.3 Phosphorus sub-model

Modelling soil phosphorus (P) losses at the paddock scale as a source of pollution, and finding options for reducing P loss are key elements of water quality assessment. HowLeaky quantifies P exports from the paddock in runoff and sediment. P in runoff can be categorised as particulate, filterable and dissolved. In agricultural systems with ample soil cover, such as well grassed areas with minimal grazing, exported P is usually dominated by soluble P (Sharpley et al., 1995, Sharpley, 2006), as high soil cover conditions typically result in low sediment concentrations in runoff. The reverse is also true; as the cover is reduced, sediment concentration rises, as does the contribution of particulate P. P losses can also be categorised based on reactivity or bioavailability. Bioavailability of P in runoff is not simply defined as there are many biotic and environmental factors that affect P availability and uptake. HowLeaky estimates total P, dissolved P and bio-available P (Queensland Government, 2019; Robinson et al. 2009).

2.1.4 Nitrate sub-model

An estimation of the surplus nitrogen that leaves farming systems is required to manage nitrogen input into the system. The nitrogen sub-model in HowLeaky is still under development and currently it can only simulate the transport of nitrate as three separate transport processes from the system: dissolved in the runoff, leaching in deep drainage, and transporting particulate nitrogen in runoff. There are also multiple options for each method of transporting dissolved nitrogen in runoff (and leaching and particulate nitrogen in runoff) implemented in the sub-model (Queensland Government, 2019).

2.1.5 Pesticide/herbicide sub-model

The pesticide module (Shaw et al. 2011) in HowLeaky evolved from CREAMS/GLEAMS (Leonard et al., 1987) with enhancements based on field observations in Australia (Silburn, 2003; Shaw et al., 2011). Pesticides can be specified as applied to the plant canopy; crop residue; or the soil, with subsequent “wash off” by rainfall to the soil below. Degradation in each pool uses an exponential decay function dependent on temperature. Loss of pesticide in runoff is based on an empirical relationship between pesticide concentration in soil and concentration in runoff (Silburn, 2003), while the partitioning of each chemical into the water and sediment phases is determined by a linear isotherm. Application of chemicals below the soil surface and losses due to leaching are not considered.

2.2 Sites

A data collection summarising approximately 140 water quality related studies across Australia (http://howleaky.net/index.php/library) was available as an empirical basis for this analysis. Fifteen sites were used in this analysis, with 12 having sediment data, seven with soil water data and four with nutrient or pesticide data (Table 1, Fig 1). Some datasets only recorded average annual or annual values while other sites had time series of daily observations for varying durations (3-35 years). This collection of sites provided 46 observations of mean annual runoff, based on 456 site-years of data across a range of land uses, climate and soil types.

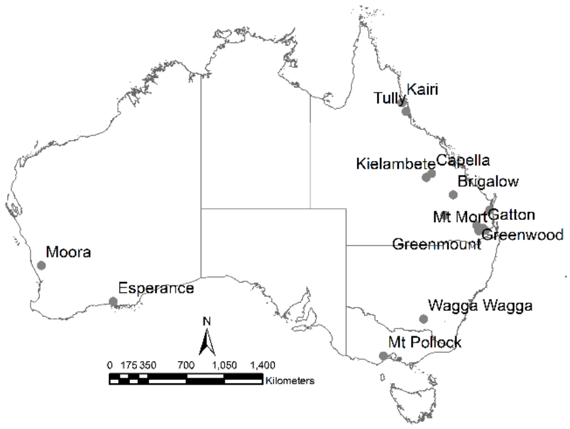


Fig 1. Location of fifteen sites used in this paper (details are shown in Table 1).

Table 1. Summary of site conditions, data type and special features of sites used in this analysis. PAWC: Plant available water capacity. L: slope length. Datasets available at www.howleaky.net (this will become www.howleaky.com)

Site	Annual rainfall (mm)	Soil	Site descriptions/Dataset feature	Data	Data period	Location	Reference	in this paper used for evaluation of
Kairi Research Station	1230	Red Ferrosol, PAWC: 143 mm	Slope 6%, L=20 m, 12 bounded plots (20x5m): bare soil, peanuts, maize, and pasture. Conventional & reduced tillage / Semi-arid tropics	Cover, runoff, suspended sediment, total nitrogen and total phosphorus	1998-2001	Queensland	Cogle et al. 2011	Contrasting environments (sub-tropical); Diverse environments; Manual calibration: runoff, erosion, phosphorus.
Capella	535	Black Vertosol, PAWC: 177 mm	Slope ~2%, L=130m. Nine contour bay catchments: Zero, reduced conventional tillage, wheat, sorghum, and sunflower / Shallow soil, cropping	Cover, runoff, suspended sediment	1984-1989	Queensland	Carroll et al. 1997	Diverse environments; Manual calibration: runoff, erosion.
Brigalow Research Station	655	Black Grey Vertosol, PAWC: 138 mm	Slope 3%, L=100m, Four ~13 ha catchments: Brigalow scrub, buffel grass pasture, opportunity cropping, and legume pasture / three vegetation types	Cover, runoff, suspended sediment, phosphorus, dissolved inorganic nitrogen, atrazine, 2,4-D	1965-current	Queensland	Cowie et al 2007, Thornton et al. 2007; Thornton and Elledge 2013	Diverse environments; Manual calibration: runoff, erosion, DIN, Atrazine, 2,4,D.
Keilambete	656	Rubyvale Red Sodosol, PAWC: 87 mm	Slope 4.3%, L=20m, Bounded plots, pasture at high and medium utilisation / Grazing study with extremes in management	Runoff and total soil erosion	1994-2000	Queensland	Waters 2009	Contrasting environments (grazing systems); Diverse environments; Manual calibration: runoff, erosion.
Wallumbilla	570	Brown Vertosol, PAWC: 188 mm	Slope ~2%, L=100m, Four contour bay catchments (4-6 ha), Aggressive and conservative tillage, wheat, pasture/tillage treatments and pasture	Soil water, cover, runoff, suspended sediment	1982-2000	Queensland	Freebairn et al. 2009	Soil details; Diverse environments; Manual Vs Automated calibration
Greenmount	793	Black Vertosol, PAWC: 288 mm	Slope 6% L=60 m. Five contour bay catchments (~1.2ha): winter crop, burnt, incorporated, mulch, no-till fallow, pasture/tillage treatments and pasture	Soil water, cover, runoff, soil movement, suspended sediment	1976-1990	Queensland	Freebairn and Wockner 1986a	Diverse environments; Manual calibration: runoff, erosion.
Greenwood	657	Grey Vertosol, PAWC: 243 mm	Slope ~5%, L = 35m. Four contour bay catchments (~0.6ha): winter crop, burnt, incorporated, mulch, no-till summer fallow / Tillage treatments	Soil water, cover, runoff, soil movement, suspended sediment.	7 years of data	Queensland	Freebairn and Wockner 1986a	Vegetation dynamic/semi-static.
Gatton Research Station	732	Black Vertosol, PAWC: 195 mm	Slope 6-8%, L=30 m. Twelve bounded plots (3x3m), four treatments, 3 replicates: Wheeled stubble mulch, Wheeled zero tillage, Controlled traffic stubble mulch, Controlled traffic, zero tillage / Controlled traffic (with and without compaction) and tillage treatments	Soil water, cover and runoff, detailed soil physical characterisation	1995-1997	Queensland	Tullberg et al. 2001; Li et al. 2008	Contrasting environments (Tillage/compaction); Diverse environments; Applied for manual calibration: runoff.
Mt Mort.	768	Sodosol, PAWC: 190mm	Slope 12%, L=10m, Bounded plots, pasture at three cover levels: bare, grazed, and excluded / Pasture study with extremes of management including bare soil	Soil water, cover, runoff, suspended sediment loss	1993-1999	Queensland	Silburn, 1994; Rattray et al. 2006	Vegetation &soil detail; Diverse environments; Applied for manual calibration: runoff, erosion.
Goomboorian	1063	Fine sandy loam, PAWC: 86 mm	Slope 5%, L=36m. Three treatments of pineapple management; bare soil, furrow mulching and conventional bare furrow / Horticulture (pineapples) on steep slopes	Soil water, runoff, cover, sediment loss	1991-1995	Queensland	Ciesiolka et al. 1995; Coughlan and Rose 1997	Diverse environments; Manual calibration: runoff, erosion.
Wagga Wagga Research Station	563	Yellow Sodosol, PAWC: 280 mm	Slope ~11%, L=100m, Paired paddock scale catchments: Treated -good pasture management, contour ripping, fertilizer applied, Untreated -fixed heavy grazing, no fertiliser two "paddock" scale catchments	Runoff, suspended sediment	1952-1973	New South Wales	Adamson 1974	Diverse environments; Manual calibration: runoff, erosion.
Mt Pollock	486	Sodosol, PAWC: 109 mm	Slope <2%, L 100m, nine 0.2 ha plots (20m x100m), three replicates, three treatments: Raised bed, conventional cultivation and deep cultivation / raised beds, atrazine	Runoff, sediment, atrazine, nutrients	2000-2004	southern Australia	Holland et al. 2012; Wightman et al. 2005; T Johnston per. comm.	Diverse environments; Manual calibration: runoff, phosphorus, Atrazine.
Esperance Research Station	1063	Sodosol, PAWC: 67 mm	Slope <1%, Waterlogging and poor soil structure , L 130 m duplicate treatments, 7 rows each, treatments consisted of raised beds and a normal no-till seed bed	Runoff, intermittently over 3 winters	2000-2001	Western Australia	Bakker et al. 2005	Diverse environments; Manual calibration: runoff, erosion.
Moora	430	Dermosol, PAWC:61 mm	Slope <2%, L=130 m, nitrogen flows in pasture-wheat and lupin-wheat rotations / Deep drainage and nitrate leaching	Soil water, deep drainage and nitrogen losses in drainage	1994-1996	Western Australia	Anderson et al. 1998	Diverse environments; Manual calibration: runoff, erosion, Nitrate.

2.3 Soil parameters and profile characterisation

Soil water holding characteristics were based on databases such as experimental site descriptions where available, including APSoil (APSoil, 2012; Dalgliesh et al., 2006), soil surveys and qualitative assessments of profile hydrology based on local knowledge. Model sensitivity to detail in soil specification was compared using two, four and six-layers. The parameter that determines partitioning of rainfall into infiltration and runoff, curve number (CN) was adjusted to best reflect observed runoff patterns and amount. A default value for soil erodibility (K) was based on Loch and Rosewell (1992) and Loch et al. (1998) and adjusted if soil erosion or sediment data were available. A single value of K was adopted for each site unless there was evidence that soil stability was grossly modified by land use, such as thick swards of pasture and root mass or soil surface armouring associated with weathering. This means we adjusted CN for soils with a heavy pasture sward as they behaved differently to cultivated soil beyond the change in CN associated with cover, and may be considered as a weakness in parameterisation. The delivery ratio was assumed to be 0.15 for catchments where sediment deposition occurred before water sample collection, typical of some small catchment studies (Freebairn and Wockner 1986b). This delivery ratio value was based on sediment size analysis of a wide range of soil types (Mark Silburn pers. comm.) while a value of 1 was used if runoff was sampled at the end of a slope before deposition could occur.

All soils require estimates of maximum daily drainage rate between each layer along with two infiltration/runoff parameters: Curve number (CNIIbare) (SCS runoff curve number for moisture condition II bare soil); and CN response to 100% cover; and two evaporation parameters (Cona and U) (Littleboy et al., 1992). The three soil specifications require 9-29 input parameters to describe water storage and movement into and within a soil profile. The CNII bare and CN response to cover were calibrated to observed runoff data subject to availability. A single soil type is used for each experimental site unless there is a clear reason to do otherwise, such as deep swards or mats of roots changing the nature of a soil or compaction changing internal drainage rates e.g. Li et al. (2001).

2.4 Parametrisation of Phosphorous and Nitrate models

Phosphorous availability for transport in runoff is based on the following soil phosphorus tests: Total P; Colwell P (Moody, 2007); Phosphorus Buffering Index and an enrichment ratio based on clay content or constant ratio function (Robinson et al., 2007). Nitrification of soil organic matter is dependent on complex interactions between soil organic matter, crop residues, soil texture, soil biology, soil moisture content, and soil temperature. The ephemeral nature of losses of nitrate to gas or leaching and rapid uptake by crops and weeds make soil nitrate concentrations difficult to simulate accurately. The inclusion of these processes into a model comes with a large burden of model specification above that desired for HowLeaky (Keating et al., 2003). The database of sites with nitrogen in soil and runoff is limited, so a simpler process for specifying soil nitrate leaching as the main mobile nitrogen form was applied.

A monthly time series of soil profile nitrate was constructed, based on the knowledge of experts in agronomy and soil nutrition who have a good understanding of soil nitrate accumulation over fallows and fertiliser practice. The time series constructed in this approach allows for average N processes such as fertiliser and plant uptake but does not consider a year to year variation. Nitrate N (kg/ha) is assumed to be uniformly distributed in the plant available water capacity (PAWC) and when leaching occurs, a proportion of the total nitrate is lost based on:

Nitrate N loss (kg/ha/day) = Deep drainage (mm) * Nitrate concentration in soil water contributing to leaching (mg/L) * efficiency coefficient * 0.001.

The efficiency coefficient was added to accommodate empirical data as it became available. Deep drainage is available from the water balance calculations and a value of 0.5 is assumed for the efficiency coefficient. A coefficient less than 1 indicates preferential loss of moisture relative to nitrate, for example, due to incomplete mixing of rainfall with the soil solution during transit through the profile.

2.5 Parametrisation of Pesticide model

The pesticide model implemented in HowLeaky (Shaw et al., 2011) requires date and rate of application, placement (soil, residue, or crop), and pesticide properties including pesticide half-life, degradation activation energy, pesticide sorption coefficient, and a runoff extraction coefficient. Parameter values were sourced from Kookana et al. (1998), Rattray et al. (2006), Silburn (2003), the Footprint pesticides properties database (http://www.eu-footprint.org/ppdb.html) and pesticide labels. Pesticide parameter values were not calibrated.

2.6 *Approach to comparing detail in the model specification*

To explore the impact of detail in model specification, we compared model performance where detail in soil description (number of layers, treatment/plot vs. site average) and vegetation (semi-static cover specification vs. dynamic LAI based model) were used to specify the model. Each land-use-management system was described using available soil type, crop, and pasture descriptions which in some cases were incomplete. These comparisons were applied to assess losses in model performance when each system was specified in less detail.

2.6.1 Setup of the soil profile

Three levels of soil description (Fig 2) were used to explore the impact of detail of soil description on simulated runoff and sediment loss from two land uses. The three descriptions are:

- Two-layers, with soil water holding specified by three inputs: available soil water (% volumetric) of the surface layer; depth of effective evaporation; depth of soil water extraction; and total porosity for both layers (specified parameters n=5) (water between air dry and wilting point is ignored);
- Four-layers with inputs: air dry content in the top two layers, wilting point, drained upper limit, saturated water content for all layers (n=26); and
- Six-layers with inputs: air dry content in the top two layers, wilting point, drained upper limit, saturated water content for all layers (n=32);

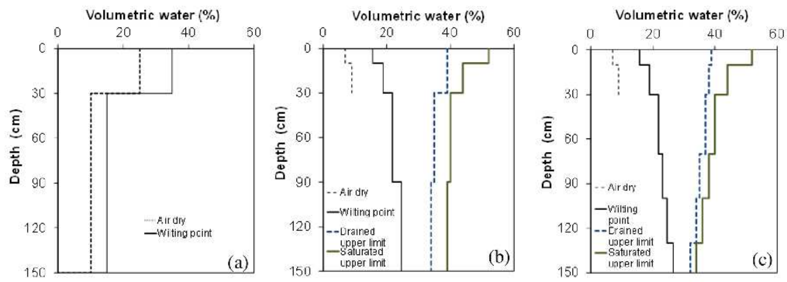


Fig 2. Three levels of soil water description (a) 2-layer using five soil water input variables (n=5); (b) 4 layers (n=26), and (c) six layers (n=32). n is the number of variables describing water holding properties.

2.6.2 Vegetation growth model: Semi-static vs dynamic

Two approaches were used to describe the vegetation sub-model: (i) a simple, semi-static cover description, which typically is average monthly values of plant cover and dead and root depth repeated each year (Fig 3a) or a multi-year cropping cycle as shown in Fig 3b); (ii) Leaf Area Index (LAI) (Fig 3c). The simpler cover model description of green and dead cover reduces complexities associated with land use and management specification which require planting and tillage rules to deal with common agronomic practices being evaluated (Freebairn and Wockner 1986a). The simpler cover model allows predefining annual or multi-year profiles of green cover (%), residue cover (%) and root depth (mm). Information relating to building the simpler cover model is gained from local expert agronomic knowledge. The LAI model explicitly specifies a crop's potential growth pattern including responses to heat sum and heat and water stress. The LAI model approach is used in a wide variety of cropping systems models such as EPIC (Williams, 1983) and APSIM (Holzworth et al., 2014) and hydrological models with cropping components such as SWAT (Neitsch et al., 2001). The LAI model requires planting and tillage rules to mimic agronomic practices while a greater understanding of crop physiology and agronomy of each agricultural system, similar to models such as APSIM (Holzworth et al., 2014) is required where the focus may be on detailed crop specification (Fig 3c). A practical difference in these two approaches is that the semi-static cover description can be gleaned from local knowledge while the dynamic LAI model requires an understanding of crop physiology and agronomy, as codified within the model and requires a depth of experience. The cover model is easier to specify, transparent, stable, and not being controlled by arbitrary planting and tillage rules.

In specifying a static cover model, "% green cover" controls transpiration while total cover (green and dead) influences runoff, evaporation and soil erosion. This approach also supports specification of a wide range of land use and management options without reference to databases of model parameters such as used in EPIC (Williams, 1983).

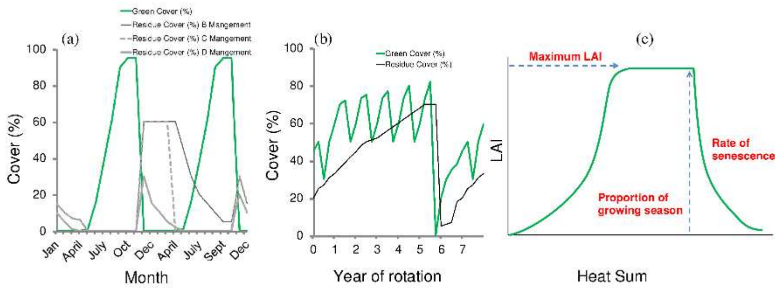


Fig 3. Example of vegetation cover models (a) static vegetation, showing average monthly crop and residue cover over a two-year wheat-canola rotation with three levels of crop residue management; (b) static vegetation for a 7 year banana rotation including annual harvests and a replant in year 6; and (c) a typical LAI development model specified by twenty one parameters. Legend of (a): B= Best management, C= Current practice, average management practice and D= poor management

2.6.3 Soil setup & Vegetation cover

Case study 1: pasture

The impact of detail in specifying vegetation and soil description was explored using data from a pasture study near Mt Mort in SE Queensland (Silburn, 1994, Rattray et al., 2006) where runoff and soil erosion were monitored on three land treatments (grazing excluded, grazed and bare soil). Three levels of system specification were compared:

• L1 - observed green and residue cover levels and soil descriptions of individual plots for each treatment were used to describe the three land uses - a literal re-enactment of experimental conditions;

• L2 - observed green and residue cover levels and a single soil description for all treatments - assuming uniform soil across the site; and

• L3 – a generalised monthly distribution of green and residue cover reflecting treatments and a single soil description - a broad description of three pasture management systems.

Case study 2: tillage

We compared two approaches to crop and management specification for a tillage study at Greenwood experimental site (see Table 1 for details of the site): a semi-static representation of green and residue cover; and a dynamic LAI model where the model adjusts cover (green and residue) on a daily basis through feedbacks from soil water, temperature, crop growth and tillage. Both models specify ~30-40 variables but the static model’s input is more transparent to the non-specialist (i.e. non-modeller).

2.7 Manual and automatic calibration

In this paper, we have performed a manual calibration and an automatic calibration approach. An experienced modeller who is familiar with each study site generally conducts manual calibration. It is a process to adjust the value of the model parameters manually to match the outputs to the observation visually (Pechlivanidis et al., 2011) or to evaluate the performance by using statistical measures, e.g. Nash Sutcliff Efficiency and percent bias (Moriasi et al, 2007; Wang et al., 2012). This method can be time-consuming and is best suited to experienced modellers while it is difficult or even impossible to use for watershed models (Sorooshian and Gupta, 1995) but not necessarily the case for a paddock scale models, such as HowLeaky, which has a relatively smaller number of parameters as described in section 2.1 (Queensland Government, 2019). Automated calibration is performed using algorithms, which in general, include objective functions that measure differences between observation and simulation values, with control variables which help to decide on the range of each parameter and an algorithm to optimise parameters (Efstratiadis and Efstratiadis, 2010). This method offers consistency in performance by removing likely biases associated with modellers’ skills (Boyle et al., 2000) but nevertheless requires a well-trained modeller to supervise the process and perform judgment of the modelling outputs (Gupta et al., 1999).

2.7.1 Manual calibration strategy

The database of field studies had a range of detail from daily observations and detailed experimental descriptions to annual values with little detail, short duration and incomplete records. In tuning the model with observed data, estimates of runoff were compared with the observed event, annual or cumulative values, depending on data quality. For example, some sites reported annual runoff and soil loss values, while six sites had daily observations (Table 1). While data completeness varied across studies, all sites were given a similar level of effort after available data was compiled.
The model was calibrated for each site by:

- Using any readily available hydrologic and site descriptive data. If site rainfall data was not available, weather data for the closest site was accessed from Silo (www.longpaddock.qld.gov.au) (Jeffrey et al. 2001). Soil and land management descriptions were estimated from local knowledge and data custodians if site-specific information was not available;
- Adjusting CNIIbare to achieve a visual fit of cumulative predicted and observed runoff using one value of CNIIbare to describe all treatments at a site (e.g. Figure 5b);
- Adjusting soil erodibility (K) value to match estimates of long-term sediment loads;

- Applying published values for phosphorous and pesticide parameter values and expert opinion to describe soil nitrate time series.

The models were not calibrated beyond an initial system specification, allowing algorithms within the model to deal with treatment impacts (crop and residue cover, pesticide application rates and dates) without further calibration. This qualitative “reasonable fit” approach allowed for an efficient examination of model performance for each experimental site. Cumulative plots of observation and model estimates were used to minimise errors associated with timing errors common in rainfall-runoff records and applying a pragmatic view that a model should capture the main processes, accepting that hydrology and erosion are impacted by rainfall rates at sub-daily time intervals. For example, a daily model cannot distinguish between a runoff event over several hours and an event with similar total rainfall over several days. The data manager is challenged by allocating runoff and soil erosion when arbitrary cut-off times are allocated to a day (typically 9:00 am). As models must conserve mass, viewing model performance using a cumulative plot recognises these necessary shortcoming in the model and data collection and recording practice. As a check for serious data and interpretation errors a draft of most site’s (Table 1) model performance was reviewed by data custodians and adjustments made if more information became available. These reports are placed on www.howleaky.net). Once data was assembled, model assessment typically took 1-2 hours per site to complete. Little attempt was made to improve model performance if model estimates were visually similar in pattern and amount to observed cumulative values. If model output describing the main hydrologic processes survived a visual inspection for sensibility (absolute values and trends), parameter sets were accepted.

While a further iteration of parameter values would likely improve the coincidence of modelled and measured values, the aim of the study was to generate a set of useful parameter values to be applied with fair confidence in other applications, especially in allocating hydrology and water quality signatures to a range of land use and management practices to estimate their performance toward water quality targets (The State of Queensland, 2018).
Details of the manual calibration strategy:

1. Create a soil description based on an observed plant available water (PAW), estimated using a site lower limit (LL), drained upper limit (DUL) and soil bulk density to estimate Plant Available Water Capacity (PAWC);
2. A fixed plant date was applied to cropping studies rather than year specific dates with a mid-maturity crop selected to avoid uncertainties associated with planting rules and varieties, aiming to provide a generalised description of a winter cereal crop. LAI model parameters were based on literature (Kiniry et al., 1995) and field observations;
3. Available observations of soil cover were used to derive tillage and stubble decay parameters;

4. Plant available soil water (PAW) observations were used to check that the model was predicting the water balance sensibly (Freebairn et al., 2018);
5. Cumulative predicted and observed runoff were compared and CNIIbare and Cover-CN were calibrated to achieve Observed: Predicted (O: P) ratio near 1; and
6. Cumulative predicted and observed sediment losses were compared and soil erodibility (K) and the delivery ratio were calibrated to achieve an O: P ratio near 1.

2.7.2 Multi-parameter calibration using PEST

When calibration is performed without prior knowledge of the parameters and their associated bounds, manual calibration can be associated with uncertainties in the parametrisation of some parameters e.g. degree days from planting to harvest, harvest index; while parameters such as PAWC (if measured) are known with reasonable certainty.

Interactions between parameters add further uncertainty (Ghahramani et al., 2011) leading to the propagation of errors through the process pathways shown in Fig 4. Inherent error shifts to the next sub-model through model input while adding to internal errors within each sub-model (Fig 4). The output of each sub-model provides input to the next sub-model through the modelling process. This cascading of errors adds uncertainty to each simulation (McMillan et al., 2011). Further, there are inherent limitations of handling simultaneous and multiple variable observations at once using traditional manual calibration processes as interactions between parameters are seldom understood and difficult to trace. Over recent decades, there have been efforts to substitute manual calibration approaches with unbiased and potentially more efficient automated mathematical procedures. Bounds for parameter values can be defined by expert judgement and then complimented further by establishing posterior distribution ranges built by uncertainty analysis. Without completing a parametric uncertainty investigation, a well-calibrated model may have uncertain outputs when used for predicting outcomes of partially understood systems and interactions. Fine manual calibration based solely on expert knowledge may also require a sensitivity analysis and calibration of parameters to reduce the error variances associated with them (Doherty, 2015).

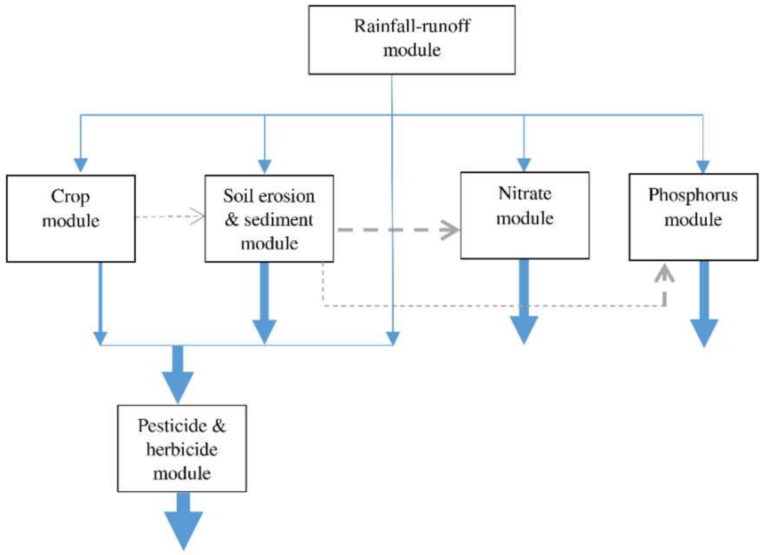


Fig 4. Schematic view for inherent error propagation in modelling water quality. Inherent error shifts to the next sub-model through model input while added to the internal error of the sub-model. Each arrow shows the output of a module (model component). The arrow below a module is the output of the module and the size of the arrow represents the magnitude of the likely associated error with sub-model output.

To calibrate model parameters, we used the parameter estimation software tool, PEST++ (Welter et al., 2012) and utilities of PEST v 15.0 (Doherty, 2016 a,b). To demonstrate the effect of uncertainty in the model application, PEST was applied to a dataset (daily data during 1983-1991) with reliable observations of soil water, hydrology, crop management, crop and crop residue cover and sediment loss from a 3-4 ha agricultural catchments near Wallumbilla, representing a spectrum of field conditions associated with alternative tillage regimes (Freebairn et al., 2018). PEST was used to adjust parameters with expert defined upper and lower bounds based on judgement and literature (Freebairn et al., 2009).

The HowLeaky model was linked with PEST++/ PEST through a series of intermediate files and compatible code scripts, allowing for an iterative dialogue between model simulations and PEST. Multiple observation databases were grouped to build an objective or loss function and each of the observed variables was assigned a preferred weight. The loss function employs the Least Square technique to minimise the difference between observed and predicted variables (Doherty, 2016a). Available multiple observations representing varying environmental and management conditions were used to ensure that the system parameters converge to near field conditions with minimum error variances. Each simulation was allowed a “warm-up” period of one year (1982) to provide initial conditions (Ghahramani et al., 2015). Sequences of management practices were set to those from the field study. Calibrations were performed from the multiple observed data samples from Wallumbilla for an eight-year period (1983–1991) followed by validation with four years of observations (1992-95).

An identifiability analysis of the parameters was carried out to analyse whether it is possible to assign unique values to the parameters from datasets (Guillaume et al., 2019). Identifiability is a defined scalar term for relative estimability that ranges from 0 to 1 (Doherty 2015). In other words, it is a method to restrict the search for a nearly unique solution to the inverse problem by sampling a parameter set to a number that allows the model to effectively reproduce the observations. If the identifiability of a parameter is 1.0, then that parameter is completely estimable on the basis of the current calibration dataset. This does not mean that its estimation is without error; however it means that measurement noise, and not an information deficit in the calibration dataset, is responsible for this error. Alternatively, if a parameter has an identifiability of 0.0, then the calibration dataset is completely uninformative of that parameter; thus the parameter is completely insensitive as far as the calibration dataset is concerned. On the other hand, if the identifiability of a parameter is between 0.0 and 1.0 then information within the calibration dataset that pertains to that parameter is shared between it and other parameters; the parameter can therefore not be resolved uniquely (Guillaume et al., 2019).

2.8 Model validation

2.8.1 Statistical validation

Two statistical techniques were applied to calculate goodness of fit and to evaluate modelling performance. These statistical evaluations of the simulation performance included Nash Sutcliff Efficiency (NSE) (Nash and Sutcliffe 1970) and Percent Bias (PBIAS) (Moriasi et al., 2007). Equations and performance ratings are presented in Table 2.

Table 2 General model performance rating approach used in this paper

	Performance rating			
Equation	Very good	Good	Satisfactory	Unsatisfactory
$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{mean}\right)^2}$	0.75 – 1.00	0.65 – 0.74	0.50 – 0.64	<0.50
$PBIAS = \frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right) \times 100}{\sum_{i=1}^{n}\left(Y_i^{obs}\right)}$	< ±10%	±10% – ±15%	±15% – ±25%	> ±25%

Y_i^{obs} is the observed data, Y_i^{sim} is the simulated data, and Y^{mean} is the mean of observed data. PBIAS values of greater than 0 indicates model overestimation and negative values indicate model underestimation bias.

2.8.2 Temporal resolution of the model output

Issues of temporal scale (e.g. daily, monthly) can affect how observed data are compared against simulation results (Daggupati et al., 2019). Field observations of runoff, soil erosion, and suspended sediment from two long-term experimental sites of Greenmount and Greenwood (Table 1) were used to explore effect of temporal resolution in the evaluation of the model performance. This evaluation was performed for catchments with relatively long-term observations for different management and land use i.e. cropping with stubble burnt after harvest with little soil cover.

2.9 Response to land use and management

A key application of simulation models is to explore and quantify management options for improving water quality. Models aim to add value to empirical studies by stretching short records of observations and exploring the effect of variations in management options on the hydrological responses that are beyond the resources of any one research program. Contrasting land uses and a range of management conditions for cropping and pasture are used to assess the model's robustness in describing the impact of management on hydrology and water quality. In all cases, the model's internal management algorithms were used to estimate management impacts, with no "treatment" specific model calibration being applied.

3. Results and discussion

3.1 Comparison of model input details and resolutions in the model specification

3.1.1 Effect of Soil description

Applying the model with three levels of soil profile definition (2, 4 and 6 layers) resulted in similar estimates of runoff and erosion at Wallumbilla, all being in reasonable agreement with measured values; however, the soil profile with 2 layers was an exception, showing that simulation results were underestimated (Table 3). The three levels of soil description also provide similar estimates in terms of absolute values and rankings for two contrasting land uses – crop and pasture. On reflection, this is not a surprise as runoff is mainly triggered when the soil is wet and runoff is dominated by a few large rainfall events in these semi-arid environments. The distribution of water in the soil (i.e. number of layers) is less likely to impact on runoff, with the CN value dominating runoff prediction. We arbitrarily used a four-layer soil description as a default specification, even when high resolution (more detailed) soil descriptions were available at some sites, while other sites did not have a specified plant available water capacity (PAWC) profile. The number of parameters required to describe a soil profile were 5, 26, and 32 for a 2, 4, and 6-layer soil respectively, representing almost a 6-fold increase in values with no apparent improvement in predictive capacity. From an operational viewpoint, it is simpler and less error-prone to use fewer variables describing a soil's water holding properties as long as the PAWC value is representative of the site. Conversely, there is no disadvantage in using more detailed descriptions if available. However, this can be still dependent on the expert knowledge of the modeller who is familiar with the site and experimental data.

3.1.3 Effects of semi-static vs. dynamic vegetation

Predictions of runoff and soil erosion from two model approaches, dynamic and semi-static, are compared with observations comparing alternative fallow management strategies between annual wheat crops at Greenwood (Table 3). Both calibration approaches capture treatment differences with absolute values varying between sets of simulations. While a statistical comparison of the two approaches might suggest a more refined optimum specification, further calibration 30+ parameter values is unlikely to inform the basic proposition being explored that relatively simple system descriptions are adequate for estimating hydrology and water quality at the paddock scale. From an operational viewpoint, we found that capturing data similar to that shown in Fig 3 from local experts was efficient and reliable. On the other hand, our experience with the implementation of a dynamic model with detailed crop specifications, including planting and tillage rules, is prone to model instability, especially when planting rules aim to mimic farmer behaviour. Given this result, static cover descriptions were used for all other analyses in this study.

Table 3. Comparison of observed and predicted runoff and soil erosion for three sites with a range of resolution in soil and vegetation specification.

Location/ Management	Average Annual runoff (mm)				Average Annual soil erosion (t/ha)			
	Observed	Prediction			Observed	Prediction		
Soil resolution [1] (Wallumbilla, SW Qld)		2 layer	4 layer	6 layer		2 layer	4 layer	6 layer
Winter crop, tilled	55	44	54	57	3.3	2.5	3.4	3.6
Pasture, light grazing	14	27	18	17	0.3	0.2	0.2	0.2
Vegetation/soil resolution [2] (Mt Mort, SE Qld)		L1	L2	L3		L1	L2	L3
Bare soil	136	159	147	142	46	50	65	54
Grazed pasture	22	18	15	20	0.5	0.7	0.4	0.7
Un-grazed pasture	3	4	12	17	0.06	0.1	0.2	0.2
Vegetation dynamic/semi-static[3] (Greenwood, South Qld)		Static	Dynamic			Static	Dynamic	
Stubble burnt	85	86	95		39	37	39	
Disc tillage	71	80	81		10	15	24	
Blade tillage	65	70	65		5	7	5	
No tillage	66	63	60		3	3	4	

Notes:

1. 5 ha catchments, 3-15 years data, Brown Sodosol Wallumbilla, Queensland (Freebairn et al. 2009). 2, 4 and 6 layer soil descriptions are shown in Fig 2.
2. 10 m long bounded plots, 6 years of data, grazing study: three levels of vegetation description (L1, L2, L3) (Silburn, 1994).
3. 0.8 ha catchments, 6 years data, winter crops with a range of stubble management: model specified as a) static annual cover pattern, and b) dynamics crop LAI and residue cover, tillage and planting dates specified (Freebairn and Wockner 1986a).

3.1.4 Effects of Vegetation and soil type

In comparing the impact of resolution (and effort) in describing a grazing trial at Mt Mort, Table 3 presents observed and predicted runoff and soil erosion values with three levels of soil profile description. Agreement between observations and predictions was variable, with no one specification being superior, with over-prediction in the runoff of 4-16% and soil erosion of 8-41%, for 2 and 6 layer models. The most generalised model specification (i.e. L3) captured the influence of management adequately and represents a generic system description which is applicable to a wider range of conditions compared to the highly specified "model re-enactment" of experimental conditions described by L1. This result supports the proposition that a generic system specification is well suited for application across a wide range of environments (with specified climate and soils), providing a broad environmental assessment of impacts of management on environmental flows (water, sediment, pesticides and nutrients). This analysis demonstrates that if green and residue cover distributions are in agreement with average field conditions, a water balance model estimates runoff and resultant soil erosion in agreement with field observations without tuning each "treatment", allowing the models algorithms to deal with changes in soil and cover conditions.

3.2 Contrasting environments and land uses

This section describes the HowLeaky model's performance across a wide range of environments and management scenarios using daily observed data and applying a pragmatic and manual calibration approach described above. Some sites were evaluated using coarser model outputs from published studies (Waters 2009; Tullberg et al. 2001; Cogle et al, 2011).

3.2.1 Sub-tropical landscapes (crops and pasture)

Observed and estimated runoff, sediment and total phosphorus loss in runoff from small plots at the Kairi Research Station (Cogle et al., 2011) over 3 years for four soil management conditions: bare soil; cropping aggressive tillage, cropping with reduced tillage; and pasture are shown in Table 5. Generic system descriptions (soil, green and residue cover) were used to describe field conditions. The model simulated three-fold differences in runoff and four-fold differences in soil erosion for contrasting soil conditions in a semi-arid tropical environment where data of this type is scarce. The model captured the effects of different land management well. It also predicted the sediment and phosphorous runoff losses without an increase in error being pasted along.

3.2.2 Tillage, stubble, and compaction management of cropping lands (Southern Queensland)

Three tillage systems: stubble mulch; minimum till and zero till; with and without wheel track compaction created a range of soil cover and compaction conditions on small plots near Gatton (Tullberg et al., 2001). Within the model, tillage was described in terms of residue cover while compaction was described by adjusting the internal drainage rate of layer 2, based on disk permeameter measurements in this study (McHugh et al., 2009; Li et al., 2001). Observed and predicted daily runoffs are shown in Fig 5. Minimal calibration using two model parameters (CN and internal drainage rates) resulted in a reasonable agreement with field observations. Table 4 shows model performance was either very good or satisfactory, although estimates for no compaction and no tillage were over-predicted although this is not evident in Figure 5b. Daily runoff predictions are poorer for the no wheel traffic treatments, with under prediction of large events and over prediction of many small events.

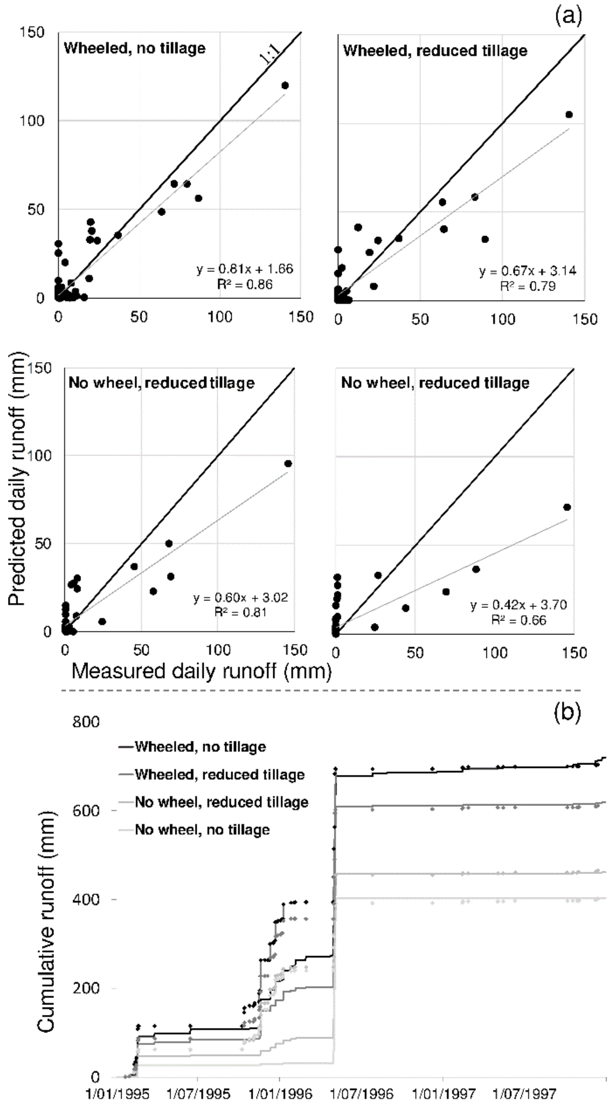


Fig 5. (a) Observed Vs predicted daily runoff for four compaction/tillage treatments, (b) Observed (dots) and predicted (lines) cumulative runoff for four tillage/compaction treatments at Gatton (data from Tullberg et al., 2001).

Table 4. Model performance for daily simulation outputs presented in Fig 5a

	Wheeled, no tillage		Wheeled, reduced tillage		No wheel, reduced tillage		No wheel, no tillage	
Method	value	Agreement	value	Agreement	value	Agreement	value	Agreement
NSE	0.86	Very good	0.78	Very good	0.75	Very good	0.57	Satisfactory
PBIAS	-2.21	Very good	-1.39	Very good	0.27	Very good	-0.51	Very good

3.2.3 Grazing landscapes

Runoff and soil erosion from plots with three levels of pasture utilisation were monitored in central Queensland (Fraser and Waters 2004, Waters 2009). The model developed for this site used a static cover description and a common four-layer soil description. Table 5 summarises model predictions using CN/cover reduction values (95/10) and soil erodibility (K = 0.22) for all conditions. The model captured the essence of hydrology and water quality responses to grazing management and climate with minimal adjustment.

Table 5 Average annual observed and predicted runoff, sediment and total phosphorus loss from three studies with a range of land use and management conditions.

Site and Management description	Runoff (mm)		Sediment (t/ha)		Phosphorus (kg/ha)	
	Observed	Predicted	Observed	Predicted	Observed	Predicted
Crop, pasture, north Queensland, Kairi Research Station (Cogle et al, 2011)						
Bare soil	282	269	21	17	23	31
Cropped & tilled	116	93	5	4	7	8
Cropped & reduced tillage	93	91	3	3	6	6
Pasture, high cover	77	78	0.8	0.5	2	2
Crop, southern Queensland, Gatton Research Station* (Tullberg et al. 2001)						
Wheeled, reduced tillage	237	240				
Wheeled, no tillage	217	206				
No wheel, reduced tillage	154	150				
No wheel, no-tillage	134	134				
Pasture, central Queensland Keilambete site (Waters 2009)						
Heavy grazing	171	171	5	6		
Moderate grazing	81	80	2	2		
No grazing	42	42	0.5	0.2		

*Daily modelling results of this site is presented in Fig 5.

3.3 Temporal resolutions effect on model uncertainty

Model artefacts, data error and parameter uncertainty are inherent in modelling hydrology and water quality, particularly across multiple experimental sites, as shown in Figs 4, 5. Additionally, uncertainty increases when processes are less-understood or insufficient data are available for developing robust empirical relationships. For example, confidence in estimating transport of sediment and chemicals can be lower compared to estimating well known processes such as soil water balance (Daggupati et al., 2019), although this is not the case in many for the simulations presented here.

Uncertainty in model prediction typically increase when simulating longer time series as long sample periods naturally include more extremes such as droughts and intense storms that may trigger additional processes (Baffaut et al., 2015). However, this was not a case in our manual calibration as we have used the full range of observations to calibrate the model to be stationary, this means, hydrological responses of simulations against observations were evaluated without changes in system conditions e.g. no changes in the land management thus the model structure could capture the processes.

The HowLeaky model uses a daily time step, thus ignores the impact of rainfall intensity (within a 24 hour period) on runoff, soil erosion, nutrient and chemical generation and transport. HowLeaky has been used at a daily temporal resolution at a field scale as the modelling engine for SoilWaterApp (Freebairn et al., 2018), with soil water being a relatively stable predictor in non-rain periods. In contrast, water quality improvement indices require temporal resolution of the model outputs at monthly or yearly intervals when used in a policy setting (Carroll et al. 2012; Commonwealth of Australia, 2015).

Fig 6 and Table 6 shows that statistical model performance improves when output at a monthly time step is used, relieving timing errors associated with the somewhat arbitrary partitioning runoff to a 24 hour period (9am to 9am) and also timing errors common in datasets such used here.

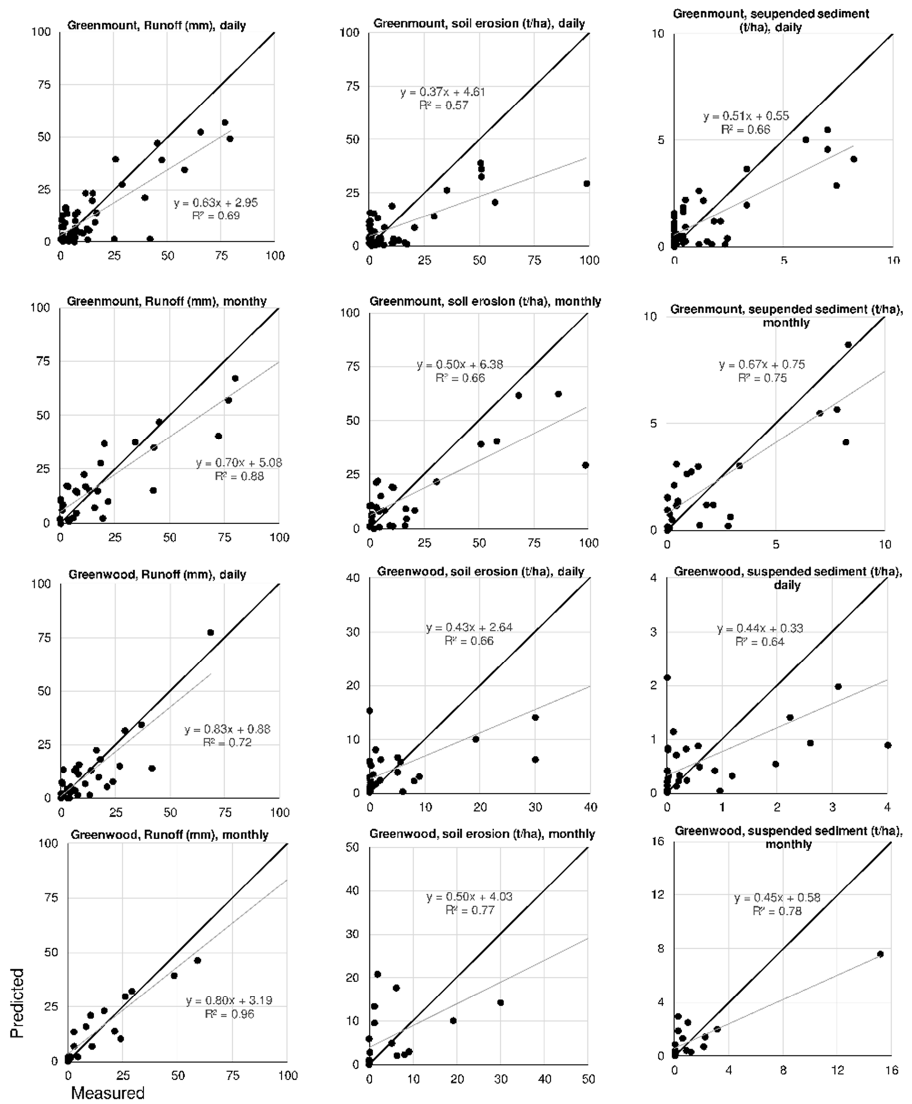


Fig 6 Observed Vs predicted daily and monthly runoff, soil erosion, and suspended sediment at two catchments: Greenmount, Greenwood.

Table 6. Model performance for data showing in Fig 6. An overall improvement in monthly simulations compared to daily.

		Daily				Monthly			
		NSE		PBIAS		NSE		PBIAS	
Site	Simulated item	Value	Agreement	Value	Agreement	Value	Agreement	Value	Agreement
Greenmount	Runoff	0.67	Good	-16.82	Satisfactory	0.83	Very good	-10.50	Good
	Soil erosion	0.47	Unsatisfactory	-26.76	Unsatisfactory	0.61	Satisfactory	-15.64	Satisfactory
	Suspended sediment	0.62	Satisfactory	-13.46	Good	0.74	Good	0.81	Very good
Greenwood	Runoff	0.69	Good	-9.64	Very good	0.93	Very good	-5.02	Very good
	Soil erosion	0.60	Satisfactory	-14.90	Good	0.70	Satisfactory	-10.20	Good
	Suspended sediment	0.56	Satisfactory	-21.08	Satisfactory	0.64	Satisfactory	-17.37	Satisfactory

3.4 Manual and automatic calibration, parametric sensitivity and model uncertainty

In addition to model performance measures as described above, here we used model sensitivity and uncertainty analysis to compare performance between manual and auto-calibration. The main comparison was to identify undetected aspects of parameter interaction that may be better described by an automatic calibration and identify the role of parameter contributions to error and uncertainty. This approach also identifies the optimum number of parameters that are sufficiently sensitive for use in model calibration, instead of choosing a large set of parameters that consume computational resources.

A total of 21 parameters (Table 7) describing vegetation, soil, and tillage conditions were used to explore sensitivities of parameters and their interactions at the Wallumbilla site (Table 1). Parameters were grouped into 5 classes (Table 7) owing to their membership to various structural modules of the model. Parameters were manually calibrated using these 21 parameters in an expert defined parameters' bound which was used as a *prior* information on parametric ranges during automatic calibration using PEST. Automatic calibration investigated the level of uncertainty over those five groups of observation time series, i.e. runoff, soil erosion, plant available soil water (PAW) in 0-150 cm depth and crop and residue cover %. This calibration used data for an 8 year period (1983 – 1990 with 1982 as a "warm-up" year.

The weight of observation groups assigned were according to the relative contribution to the loss function derived using the parameters from manual calibration. Once the model was auto calibrated using PEST/PEST++, the *posterior* standard deviation (σ_{post}) was calculated to define the *posterior* parameter bound using the equation $\mu_{post} \pm 2.\sigma_{post}$. Where μ_{post} is the *posterior* best fit parameters estimated by automatic calibration. Further the *posterior* parameter range along with best fit parameters were used to analyse the sensitivity and uncertainty in the model with a final adjustment made to the weights of the observation groups derived from their relative contribution to objective/ loss function.

Among those parameters assessed by PEST over the manual calibration, the solution space confirmed that seven parameters were identifiable in range between 0.5 and 0.94 (Fig 7 and Table 7) with 14 referred to the *null space*. None of the parameters were entirely unique to the calibration dataset suggesting that the model parameters were structurally locally identifiable with data i.e. not unique globally (Kreutz, 2018).

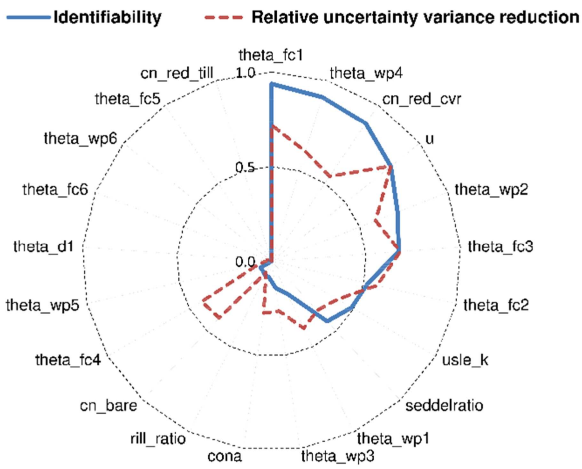


Fig 7. Identifiability and uncertainty of the model parameters post automatic calibration. (The parameter codes were as defined in Table 7)

The relative uncertainty variance reduction (Table 7) is a similar statistic that suggests the role of a parameter in reducing uncertainty errors in variance (relative contribution from structural and measurement errors) in the overall estimates of outputs corresponding to a set of measured observations. In this case, parameters U, field capacity of layer 1 and 2 (Table 7) are very important in terms of their contribution in reducing the total uncertainty error variance. Thus the inclusion of these parameters for model calibration with the available observed data is likely to reduce the variances in uncertainty. The reduction in the prediction uncertainty accrued through calibration is primarily due to the reduction in uncertainty in those parameters. However, as mentioned in the analysis of identifiability, a non-unique nature of a parameter technically share its contribution to the error uncertainty with other sets of parameters (Guillaume et al., 2019).

To describe uncertainty, we investigated the role of CN values in uncertainty analysis which is nearly non-identifiable. Table 8 shows that the uncertainty error contribution of various parameter groups was due to the inclusion of CN as a calibration parameter. In manual calibration, the predictive uncertainty variance was only detected from the runoff group. It failed to detect contributions from the other groups in the uncertainty reduction. When subject to PEST calibration, predictive uncertainty variance has come down to 0.00036 collectively from 0.1744 with apportioned contributions from other parameter groups. This underpins the structural interaction of a non-unique parameter within the model while operating over the process modules. These parameters are interdependent and contribute jointly to variance reduction, a fact that often remains undetected in manual calibration.

Table 7. Parameter values before and after automated (PEST) calibration, and their relative sensitivities, and relative uncertainty reduction.

Parameter	Parameter group	Manual	PEST calibration	Lower Limit	Upper Limit	Relative Sensitivity	Rank (Sensitivity)	Identifiability	Relative uncertainty variance reduction	Rank (uncertainty)
Bare soil Curve Number (CN)	Runoff	85.00	88.50	87.14	89.86	1.000	1	0.068	0.406	9
Field Capacity, layer 2	Soil water	39.00	40.00	37.80	42.20	0.176	2	**0.522**	0.571	6
Field Capacity, layer 3	Soil water	36.50	39.71	36.92	42.50	0.157	3	**0.673**	0.682	3
Field Capacity, layer 1	Soil water	32.30	40.00	37.41	42.59	0.115	4	**0.939**	0.719	2
Stage 2 soil evap. (CONA)	Evaporation	4.00	3.20	2.95	3.45	0.105	5	0.087	0.276	13
Wilting point, layer 3	Soil water	21.80	27.00	24.24	29.76	0.059	6	0.142	0.265	14
Stage 1 soil evaporation (U)	Evaporation	4.00	3.09	2.68	3.51	0.048	7	**0.806**	0.806	1
Airdry, layer 1	Soil water	7.00	10.00	7.72	12.28	0.042	8	0.001	0.025	17
Field Capacity, layer 4	Soil water	35.80	36.69	32.38	40.99	0.034	9	0.067	0.425	8
Field Capacity, layer 6	Soil water	35.60	36.42	31.48	41.36	0.034	10	0.000	0.024	18
Wilting point, layer 5	Soil water	26.60	27.00	24.03	29.97	0.033	11	0.002	0.069	15
Wilting point, layer 6	Soil water	26.00	21.00	18.03	23.97	0.029	12	0.000	0.013	20
Wilting point, layer 1	Soil water	15.70	15.44	13.14	17.75	0.026	13	0.188	0.391	11
Wilting point, layer 2	Soil water	18.90	18.76	16.40	21.13	0.026	14	**0.715**	0.591	5
Field Capacity, layer 5	Soil water	35.60	36.91	32.03	41.80	0.021	15	0.000	0.021	19
Wilting point, layer 4	Soil water	24.60	25.62	22.79	28.45	0.020	16	**0.908**	0.607	4
Reduction in CN by cover	Runoff	10.00	5.11	2.07	8.15	0.013	17	**0.883**	0.544	7
Soil erodibility of USLE (K)	Erosion	0.40	0.20	0.10	0.30	0.000	18	0.487	0.401	10
Sediment Delivery Ratio	Erosion	0.50	0.50	0.31	0.69	0.000	19	0.431	0.354	12
Rill ratio	Erosion	1.00	1.00	0.55	1.45	0.000	20	0.081	0.066	16
Max CN reduction by tillage	Tillage	10.00	10.00	7.50	12.50	0.000	21	0.000	0.000	21

Note: HowLeaky uses metric units for soil erodibility rather than SI. Metric K is 9.81 times larger than K in SI system

Table 8. Predictive uncertainty variance of different parameter groups due to the inclusion of CN as a calibration parameter

Parameter group	Manual Calibration	PEST calibration
Runoff	0.174429	1.766E-04
Erosion	0	6.100E-10
Soil water	0	1.273E-04
Evaporation	0	5.733E-05
Tillage	0	2.000E-11

Daily time series of measured runoff, soil erosion, plant available water, and cover with their predictions through manual and automated calibrations are presented in Fig 8. The comparison of results from manual calibration to those from automated PEST calibration accounts for a subtle statistical departure in the performance in reproducing the observed data sets (Fig 9). At first glance, this comparison favours the pragmatic approach. However, a given set of parameters that immediately satisfies a complex environmental variable using manual calibration may not be perceived to be free from the problem of equifinality, an issue where different parameter sets within a model can reproduce the observations (Beven and Freer, 2001). This is because complex environmental problems are illposed (Doherty, 2015) without a unique parameter set and inverse modelling that yields an alternative set of parameters. Figs 8 and 9 shows that PEST provided a superior calibration for runoff, PAW, soil erosion and cover compared to manual calibration. A reasonably poor prediction of event soil loss and total cover is indicated in Fig 9, for both manual and PEST calibration. This would appear to indicate that use of a static temporal cover pattern has failed to capture the variation in total cover over time which led, in part, to the poor estimates of soil loss. For instance, in at least one year no wheat crop was planted and the ground cover was low for the following summer even though the simplistic model setup ignored this extreme physical outcome. This poor performance can also be related to the fact that a model with a daily time step does not consider the impact of rainfall intensity on runoff rate and subsequent soil erosion. The structure of a static cover model thus prevented PEST from optimising the results; no amount of calibration could overcome this defect in the model structure. Also, the validation period coincided with the worst soil erosion predictions (in part due to the poor cover prediction). If periods were resampled at random (multiple times) a clearer result might be obtained. However, the results presented in Table 3 and 4 indicate that the model did capture the average annual soil loss and the management effects on them.

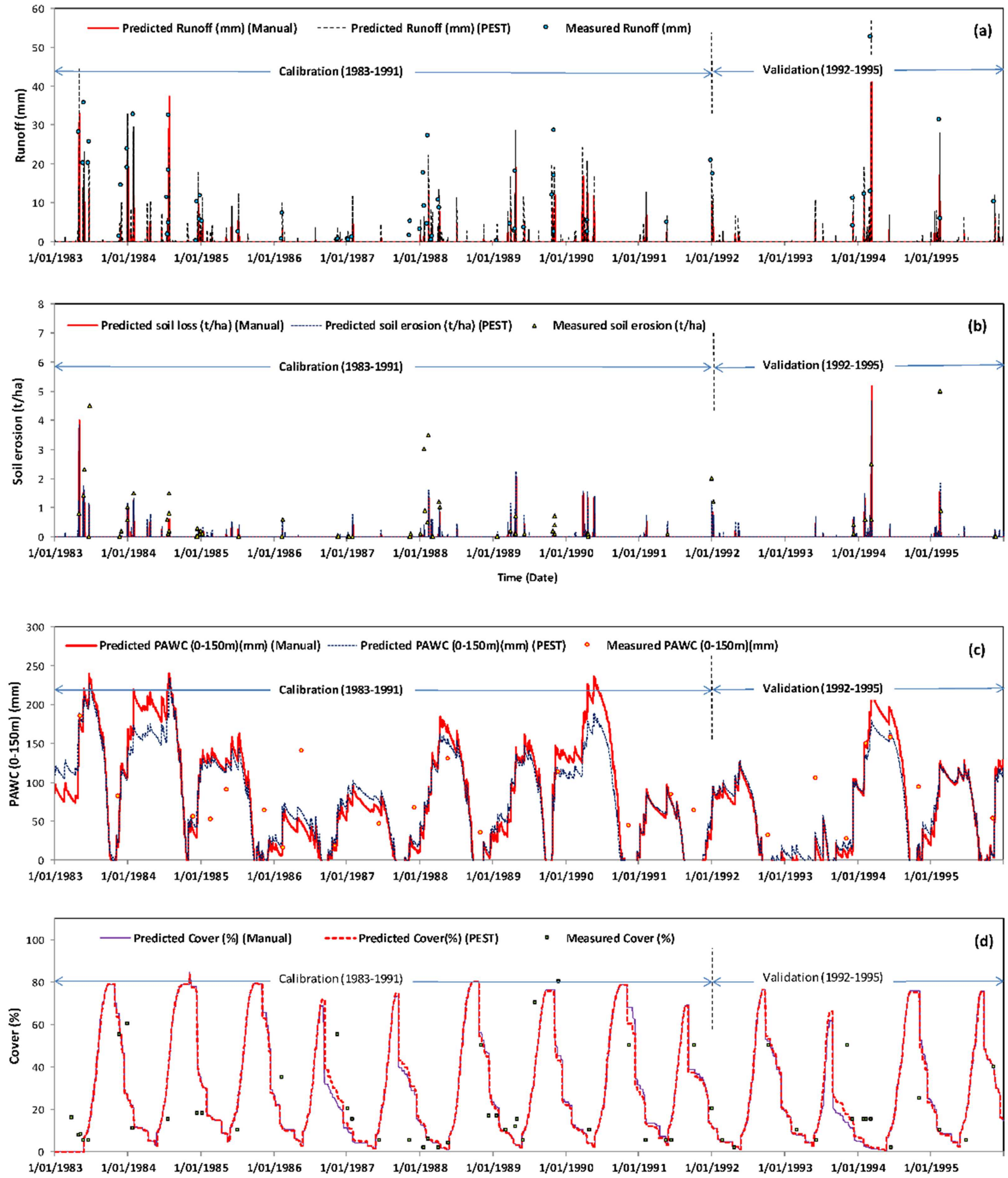


Fig 8. Daily time series at Wallumbilla site for comparison of observation variables (a) runoff (mm), (b) soil erosion (t/ha), (c) plant available water content (PAW) (mm) and (d) Plant (wheat) and residue cover (%) with their predicted counterparts during calibration (1983-1991) (both manual and PEST) and validation (1992-95).

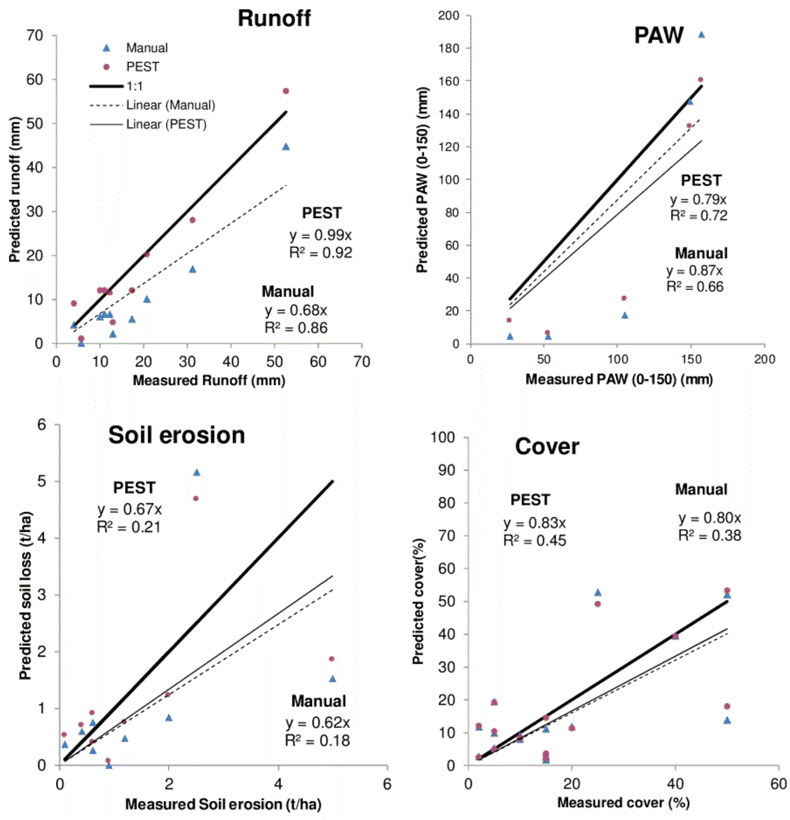


Fig 9. Comparison between simulation results of manual calibration and automated PEST supported calibration at Wallumbilla. PAW: plant available water. Calibration was performed for daily data from 1983-1991 and validations (this figure) for daily data of 1992-1995.

3.5 Diverse environments, applying manual calibration at other locations

In order to demonstrate the robustness of manual calibration in estimating water quality signatures across a range of land uses, management practices and environments, observed hydrology, sediment, nutrient and pesticide losses are compared to model estimates for all available datasets on an average annual basis (Fig 10). While the number of datasets available in southern and Western Australia is less than in north-eastern Australia, each dataset adds to the confidence in model application across environments, even when data is incomplete. Where data is sparse, there is an opportunity to fill knowledge gaps when empirical data is added. In this study, 46 average annual hydrology, 37 sediment and 12 nutrient or pesticide observations were available. This represents many site-years of data from 14 sites described in Table 1 (~2500 plot years of field monitoring).

The plot of observed and predicted runoff (Fig 10) for the available studies demonstrates the models' general applicability for describing water balances across diverse environments, soil types and management practices. At each site, estimated average annual runoff values were calibrated to observations, with the impact of management within each site described by HowLeaky's runoff algorithm which are sensitive to soil type, crop and residue cover and driven by local daily weather data. All treatments within a site had shared soil and crop parameters except those variables explicitly describing management (i.e. crop type, tillage, grazing). Where available, soil and site descriptions were used without modification. Hydrology calibrations were restricted to two parameters: CN and CN response to cover, and one soil description was applied across each site (rather than treatment by treatment calibration) and management conditions were simply described by a generic crop and residue cover distribution.

Fig 10 presents predicted and observed values of mean annual sediment loss for 45 diverse catchments. Sediment loss was calibrated using the soil erodibility factor, K, and a default delivery ratio (0.15 where water was sampled after some ponding such as at the end of a graded channel or 1 if water was collected immediately below a slope such as in plot studies). Table 5 and Fig 10 show that for any one site, the model captures differences in observed soil erosion associated with various management practices and land treatment.

The observed and predicted mean annual nutrient and pesticide losses for six available data sets are shown in Fig 10. The estimates of phosphorus and pesticide losses were in broad agreement with measured values (absolute values and responses to management) as shown in Table 9 but it is acknowledged that the empirical database is small and errors are larger than for runoff and sediment losses.

Estimates of nitrate in drainage and total N losses are rudimentary at this stage as the database is too small for any confidence to be given to model predictions, but the simple model structure and system description is well suited to accommodate new data when it becomes available. In all cases, errors in estimated runoff lead to errors in water quality; however, it should be noted that parameters for nutrients and pesticide losses were not calibrated (due to limited observations) while parameters for runoff and sediment were calibrated. This water balance based approach to estimating water quality attributes of various land use and management options was able to integrate wide variations in climate, soil type, vegetation, and agronomic management.

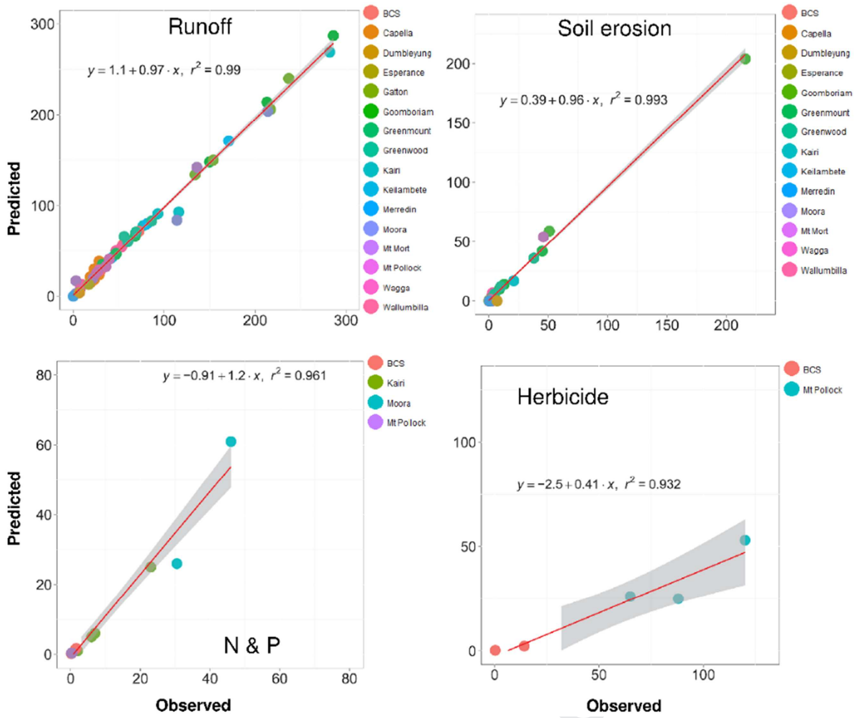


Fig 10. Observed and predicted mean annual runoff, soil erosion, nutrient, and herbicide loss.

Table 9 Model performance for data presented in Fig 10

	Runoff		Soil erosion		N & P		Pesticide	
	value	Agreement	value	Agreement	value	Agreement	value	Agreement
NSE	0.99	Very good	0.99	Very good	0.90	Very good	-0.01	Unsatisfactory
PBIAS	-1.59	Very good	-1.09	Very good	8.22	Very good	-62.92	Unsatisfactory

3.6 How pragmatic was an expert calibration?

We used a relatively simple water balance model to describe the general characteristics of each sites hydrology and water quality and importantly, the impact of management for a wide range of environments and management conditions. A simplified approach to describing site conditions was sufficient to allow for model-based estimates of daily and annual patterns of runoff for a wide range of climates (annual average rainfall of 1230 mm at Kairi catchment in north Queensland to 430 mm at Moora in Western Australia). Land uses included annual crops, pastures and horticulture, all with varying soil conditions while soil types ranged from heavy clays to deep sands. This diversity of conditions was described using a simple and efficient process. To evaluate the performance of the manually calibrated method using expert judgement to establish the parameter range, the modelling results were compared to those using a more sophisticated mathematical calibration approach with minimum user intervention. Model performances were similar and at an acceptable level suggesting visually observed history matching is a good alternative to the more sophisticated PEST parameter estimation.

The manual calibration approach to each dataset was to use data that was readily available, knowing the time-consuming nature of data extraction when there was no structured database available. Data collation generally involved contacting data monitoring teams and “custodians”, discussing data management, site conditions and the exchange of reports and spreadsheets.

A large number of datasets and variable completeness required an efficient process. Vegetation patterns were described using an average monthly time series of green and residue cover, well suited to engaging non-modellers using common language and system descriptions. A “generic” description of vegetation (green and residue cover) and soil type was adequate in specifying water balance rather than a traditional literal description of specific conditions during an experimental period. Adoption of generic soil and vegetation descriptions results in sets of parameter values that have broader application beyond the experimental sites.

Predicted runoff and deep drainage patterns were similar to observed values without major adjustments to the model, while a single CN value was generally used for each site. The model dealt with the dynamics of soil water and cover impacts on hydrology while the relatively simple water quality algorithms were able to capture site and management impacts.

While data for phosphorus and nitrogen loss data were scarce, estimates from the model were in reasonable agreement with observed values and responses to management in agreement with observations. Cook et al. (2005) and Grayson and Blöschl (2000) discussed the trade-offs between model complexity and data availability, indicating that a balance needs to be found between model resolution and data availability. In this case, the effort to tune a model using observed values across a wide range of datasets required a degree of pragmatism.

While there were differences in the runoff, soil erosion and water quality estimates from a model with a range of detail in system description, the differences between model estimates are generally small compared to the differences associated with management. It has been demonstrated that simpler or more generic model inputs will result in similar estimates of water quality compared to high resolution system specifications, in particular, for annual predictions which are being used for policy making in GBR catchments.

A caveat would be that the model practitioner should have a good understanding of the real world and expected results. Certainly, a model with fewer inputs is easier to set-up, diagnose, and apply. More importantly, in here, the expert knowledge of the modeller, and their understanding of the interactions between the site’s hydrologic, land use, soil loss and nutrient behaviours under storm conditions and management use is a significant pillar. In the absence of this expertise, the pragmatic approach could be at risk. This requirement is a fundamental prerequisite to apply pragmatic approaches for successful modelling.

This study has demonstrated that this water balance approach, with a pragmatic method of system specification, was able to add value to datasets, regardless of resolution. The ability to estimate runoff for a representative period of climate record is a useful starting point for adding value to a short record of observations. For example, an understanding of daily volumes and seasonal distribution of runoff can be a guide for when management interventions are most likely to be effective. When knowledge of hydrology is combined with simple soil erosion and water quality models, water quality can be estimated in absolute terms with “reasonable” confidence and impact of management can be determined with confidence.

Confidence in the estimation of the various components of water quality at the average annual scale, discussed in this paper in descending order is: hydrology, soil erosion and suspended sediment loss, phosphorus loss, pesticide loss and finally nitrate and total nitrogen losses. This order of confidence is roughly in the same order as data availability, indicating that our confidence in model performance is limited by data as much as the model itself. As more data becomes available, we would expect to refine algorithms.

3.7 Complexity versus simplicity

This process of land use and management specification has facilitated the capture of knowledge from technical experts and farmers alike. Static descriptions of vegetation systems are compared to more dynamic descriptions common to most simulation models. The static descriptions of crop, pasture or tree crops avoid much of the complexities associated with specifying land use systems in models such as planting and tillage rules, detail in crop rotations and pasture grazing and crop physiology descriptors, a feature of cropping system models with greater complexities such as APSIM (Holzworth et al., 2014).

This analysis shows that relying on existing, relatively simple models (e.g. HowLeaky) can be sufficient for decision making, at least for the purpose of estimating water quality signatures for various management options, e.g. in the Great Barrier Reef catchments. But a requirement for reducing uncertainties may push the user to a particular approach suited to the model i.e. a more complex parameter calibration procedure.

This result does not reject requirements for complex modelling when complexities are unavoidable, e.g. mixed farming systems that include physical and biophysical, cropping, animal, and economic structures (e.g. Ghahramani and Bowran, 2018). The evidence is presented here to show that a simple model configuration can be reliably compared to more complex setups over a wide range of environments regardless of the complexities of the physical system. Greater complexity may only increase the uncertainties related to the over-parametrisations or use of the models by people with insufficient skills (Jakeman et al., 2006). Also, a simpler modelling approach is better suited to simulating datasets that have poorly described experiential conditions and incomplete data.

4. Conclusion

A pragmatic approach for applying water balance simulation for a wide range of data qualities is demonstrated. This process produced sets of parameters describing soil type, vegetation, nutrient and pesticide behaviour which can be used with reasonable confidence to further explore the impact of management practices and land treatment on hydrology and water quality for simply specified management practices. Water balance as a methodology is demonstrated to be robust in that water-flows (runoff and deep drainage) can be estimated without detailed measurements of soil properties as long as there are some estimates of hydrology at a local or regional scale. Parameter values based on similar soil-land use-vegetation combinations improve confidence in model estimates. Model estimates of average annual runoff, sediment, and nutrients were similar to measurements from catchment studies in both magnitude and responses to management. Empirical studies are fundamental to ensuring credible analysis of hydrology, sediment, phosphorus, nitrogen and pesticide movement at the paddock scale. Given the current hydrology and water quality database of field studies at the paddock scale, the current water quality algorithms are probably of sufficient complexity (or simplicity) to deal with this available data. However, there are gaps (in data and algorithms), such as for modelling dissolved nitrogen runoff. Ideally, more data dealing with hydrology, soil erosion, sediment loss, and nutrient and pesticide movement will become available to inform the development of more reliable models.

Acknowledgements

This paper was funded by the partnership program between Queensland Government, Department of Environment and Science and Department of Natural Resources, Mines and Energy, and University of Southern Queensland (Resilient landscape), HowLeaky re-development project.

The authors acknowledge the support of the Australian Government's Caring for Our Country initiative (Project: "Capturing historic small catchment study data to support quantification of management impacts on water quality on the Great Barrier Reef" (RRRD011) and the Grains Research and Development Corporation project: "Quantifying Impacts of Management on Sediment, Nutrient and Pesticide Loss from Grain Farms" (RPS00002). We are grateful to our many colleagues who generously contributed data and information for the www.howleaky.net data collection, including key unpublished datasets that made the modelling in this paper possible, particularly Craig Thornton, Lex Cogle, David Waters, Chris Carroll, Jeff Tullberg, Tim Johnston and Dirk Bakker. Authors acknowledge useful comments made by four anonymous reviewers.

References

Abbs, K., Littleboy, M., 1998. Recharge estimation for the Liverpool Plains. Aust. J. Soil. Res. 36:335-357.

Adamson, C.M., 1974. Effects of soil conservation treatment on runoff and sediment loss from a catchment in south western New South Wales, Australia. Proc. of the Paris Symposium. IAHS—AISH Publ. No 113, 3-14.

Anderson, G.C., Fillery, I.R.P., Dunin, F.X., Dolling, P.J.C., Asseng, S., 1998. Nitrogen and water flows under pasture -wheat and lupin-wheat rotations in deep sands in Western Australia, 2. Drainage and nitrate leaching. Aust. J. Soil. Res. 49, 345-61.

APSoil. (2012) On line database for soil descriptions http://www.apsim.info/Products/APSoil.aspx.

Bakker, D.M., Hamilton, G.J., Houlbrooke, D.J., Spann, C., 2005. The effect of raised beds on soil structure, waterlogging, and productivity on duplex soils in Western Australia. Aust. J. Soil. Res. 43, 575-585.

Baffaut, C., Dabney, S.M., Smolen, M.D., Youssef, M.A., Bonta, J.V., Chu, M.L., Guzman, J.A., Shedekar, V.S., Jha, M.K. and Arnold, J.G., 2015. Hydrologic and water quality modeling: Spatial and temporal considerations. Transactions of the ASABE. 58, 1661-1680.

Beven, K. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249, 11-29.

Carroll, C., Littleboy, M., Halpin, M., 1992. Minimising soil erosion and runoff by maximizing cropping opportunities. Math. Comput Simulat. 33, 427-432.

Carroll, C.M., Halpin, N., Burger, P., Bell, K., Sallaway, M.M., Yule, D.F.,1997. The effect of crop type, crop rotation, and tillage practice on runoff and soil loss on a Vertisol in central Queensland. Aust. J. Soil. Res. 35, 925-939.

Carroll, C., Waters, D., Vardy, S., Silburn, D.M., Attard, S., Thorburn, P.J., Davis, A.M., Schmidt, M., Wilson, B., Clark, A., 2012. A Paddock to Reef Monitoring and Modelling framework for the Great Barrier Reef: Paddock and Catchment component. Marine Pollution Bulletin. 65, 136–149.

Ciesiolka, C.A.A., Coughlan, K.J., Rose, C.W., Smith, G.D., 1995. Hydrology, sediment and erosion in pineapples in steep areas of south eastern Queensland (Imbil). Soil. Technol. 8, 245-258.

Cogle, A.L., Littleboy, M., Rao, K.P.C., Smith, G.D., Yule, D.F., 1996. Soil management and production of Alfisols in the semi-arid tropics. III Long term effects on water conservation and production. Aust. J. Soil. Res. 34,113-126.

Cogle, A.L., Keating, M.A., Langford, P.A., Gunton, J., Webb, I.S., 2011. Runoff, soil loss, and nutrient transport from cropping systems on Red Ferrosols in tropical northern Australia. Aust. J. Soil. Res. 49, 87-97.

Cook, F.J., Schilizzi, S., Campbell, A.P., Asseng, S., Wardell-Johnson, A., Rixon, A., Su, X., Nancarrow, B., Carlin, G.D., 2005. Uncertainty in modelling human-landscape interactions. In: International Congress on Modelling and Simulation: advances and applications for management and decision making (MODSIM05), 12 - 15 December, Melbourne, Australia.

Coughlan, K.J., Rose, C.W., 1997. A new soil conservation methodology and application to cropping systems in tropical steeplands. ACIAR Technical Reports No. 40, 147.

Cowie, B.A., Thornton, C.M., Radford, B.J., 2007. The Brigalow Catchment Study: I. Overview of a 40-year study of the effects of land clearing in the Brigalow bioregion of Australia. Aust. J. Soil. Res. 45, 479-495.

Dalgliesh, N.P., Wockner, G., Peake, A., 2006. Delivering soil water information to growers and consultants. Proceedings of the 13th Australian Agronomy Conference. Western Australia, Australian Society of Agronomy, Perth.

Daggupati, P., Pai, N., Ale, S., Douglas-Mankin, K.R., Zeckoski, R.W., Jeong, J., Parajuli, P.B., Saraswat, D. and Youssef, M.A., 2015. A recommended calibration and validation strategy for hydrologic and water quality models. Transactions of the ASABE, 58, 1705-1719.

De'ath, G., Fabricius, K., 2010. Water quality as a regional driver of coral biodiversity and macroalgae on the Great Barrier Reef. Ecol. Appl. 20, 840-850.

Department of Premier and Cabinet., 2009. Reef Water Quality Protection Plan 2009 (Reef Plan). http://www.reefplan.qld.gov.au/about/assets/reefplan-2009.pdf Published by the Reef Water Quality Protection Plan Secretariat, September 2009, 100 George Street, Brisbane Qld, 4000.

Doherty, J., 2015. Calibration and uncertainty analysis for complex environmental models. Published by Watermark Numerical Computing, Brisbane, Australia. 227pp. ISBN: 978-0-9943786-0-6.

Doherty, J., 2016 a. PEST: Model-independent parameter estimation Part I: PEST, SENSAN and Global Optimisers, User Manual. 6th ed. Brisbane, Queensland, Australia: Watermark Numerical Computing.

Doherty, J., 2016 b. PEST: Model-independent parameter estimation Part II: PEST Utility Support Software, User Manual. 6th ed. Brisbane, Queensland, Australia: Watermark Numerical Computing.

Efstratiadis, A. and Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. Hydrological Sciences Journal–Journal Des Sciences Hydrologiques. 55, 58-78.

Fraser, G.W., Waters, D.K., 2004. Modelling runoff and erosion processes in central Queensland grazing lands. In ‘Conserving Soil and Water for Society: Sharing Solutions’. Proceedings 13th International Soil Conservation Organisation Conference. Brisbane, 2004. Paper 749. (Eds SR Raine, AJW Biggs, NW Menzies, DM Freebairn, PE Tolmie) (ASSSI/IECA: Brisbane, Qld).

Freebairn, D.M., Wockner, G.H., 1986a. A study of soil erosion on Vertisols of the Eastern Darling Downs, Queensland. I Effects of surface conditions on soil movement within contour bay catchments. Aust. J. Soil. Res. 24, 135-58.

Freebairn, D.M., Wockner, G.H., 1986b. A study of soil erosion on Vertisols of the Eastern Darling Downs, Queensland. II The effect of soil, rainfall and flow conditions on suspended sediment. Aust. J. Soil. Res. 24, 159-172.

Freebairn, D.M., Wockner, G.H., Hamilton, A.N., Rowland, P., 2009. Impact of soil conditions on hydrology and water quality for a brown clay in the north-eastern cereal zone of Australia. Aust. J. Soil. Res.47, 389–402.

Freebairn, D.M., Ghahramani, A., Robinson, J.B. and McClymont, D.J., 2018. A tool for monitoring soil water using modelling, on-farm data, and mobile technology. Environmen. Modell. Softw. 104, 55-63.

Gardner EA, Littleboy M and Beavers P. 1995. Using a water balance model to assess the hydrological implications of on-site effluent disposal. 16th Federal Convention of the Australian Water and Waste Water Association, April 1995, Sydney, Australia.

Ghahramani, A., Ishikawa, Y., Gomi, T., Shiraki, K. and Miyata, S., 2011. Effect of ground cover on splash and sheetwash erosion over a steep forested hillslope: A plot-scale study. Catena, 85, 34-47.

Ghahramani, A., Kokic, P.N., Moore, A.D., Zheng, B., Chapman, S.C., Howden, M.S. and Crimp, S.J., 2015. The value of adapting to climate change in Australian wheat farm systems: farm to cross-regional scale. Agr. Ecosyst. Environ. 211, 112-125.

Ghahramani, A., Bowran, D., 2018. Transformative and systemic climate change adaptations in mixed crop-livestock farming systems. Agr. Syst. 164, 236-251.

Grayson, R., Blöschl, G., 2000. Spatial Patterns in Catchment Hydrology: Observations and Modelling, Cambridge University Press, 404 pp., Cambridge.

Griffin, N.R.M., 2001. Evaluation of Investment in Landcare Support Projects For Agriculture Fisheries and Forestry Australia (http://www.agriculture.gov.au/ag-farm-food/natural-resources/landcare/publications/eval-support-projects Accessed 4/03/2017).

Guillaume, J.H., Jakeman, J.D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M.C., Jakeman, A.J., Keesman, K.J., Razavi, S. and Stigter, J.D., 2019. Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. Environmen. Modell. Softw. 119, 418-432.

Gupta, H.V., Sorooshian, S. and Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. J. Hydrol. Eng. 4, 135-143.

Holland, J.E., Johnston, T.H., White, R.E., Orchard, B.A., 2012. An investigation of runoff from raised beds and other tillage methods in the high rainfall zone of south-western Victoria, Australia. Aust. J. Soil. Res. 50, 371-379.

Holzworth, D.P., Huth, N.I., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D. and Brown, H., 2014. APSIM - evolution towards a new generation of agricultural systems simulation. Environ. Modell. Softw. 62, 327-350.

Jakeman, A.J., Letcher, R.A. and Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environ. Modell. Softw. 21, 602-614.

Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environmen. Modell. Softw. 16, 309-330.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N., Meinke, H., Hochman, Z. and McLean, G., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron., 18, 267-288.

Kiniry, J.R., Major, D.J., lzaurralde, R.C., Williams, J.R., Gassman, P.W., Morrison, M, Bergentine, R., Zentner, R.P., 1995. EPIC model parameters for cereal, oilseed, and forage crops in the northern Great Plains region. Can. J. Plant Sci. 75, 679-688.

Knisel, W.G., (Ed.) 1980. CREAMS: A Field-Scale Model for Chemicals, Runoff and Erosion from Agricultural Management Systems. United States Department of Agriculture, Conservation Research Report 26, 640 pp.

Kookana R.S., Baskaran S., Naidu R. (1998) Pesticide fate and behaviour in Australian soils in relation to contamination and management of soil and water: a review. Aust. J. Soil. Res. 36, 715-764.

Kreutz, C., 2018. An easy and efficient approach for testing identifiability. Bioinformatics. 34, 1913-1921.

Leonard, R.A., Knisel, W.G., Still, D.A., 1987. GLEAMS: Groundwater Loading Effects of Agricultural Management Systems. Transactions American Society Agricultural Engineers. ASAE 30, 1403-1418.

Li, X.Y., Tullberg, J.N., Freebairn, D.M., 2001. Traffic and residue cover effects on infiltration. Aust. J. Soil. Res. 39, 239–247.

Li, X.Y., Tullberg, J.N., Freebairn, D.M., McLaughlin, N.B., Li, H.W., 2008. Effects of tillage and traffic on crop production in dryland farming systems: I. Evaluation of PERFECT soil-crop simulation model. Soil. Till. Res. 100, 15-24.

Littleboy, M., Silburn, D.M., Freebairn, D.M., Woodruff, D.R., Hammer, G.L. 1989. PERFECT: A computer simulation model of Productivity, Erosion, Runoff Functions to Evaluate Conservation Techniques. Queensland Department of Primary Industries, Brisbane QLD, Bulletin QB89005, 119 pages.

Littleboy, M., Silburn, D.M., Freebairn, D.M., Woodruff, D.R., Hammer, G.L., Leslie, J.K., 1992a. Impact of soil erosion on production in cropping systems. I. Development and validation of a simulation model. Aust. J. Soil. Res. 30, 757-74.

Littleboy M., Freebairn D.M., Hammer G.L. and Silburn D.M. 1992b. Impact of soil erosion on production in cropping systems. II Simulation of production and erosion risks for a wheat cropping system. Aust. J. Soil. Res. 30: 775-788.

Littleboy, M., Cogle, A.L., Smith, G.D., Yule, D.F., Rao, K.P.C., 1996. Soil management and production of alfisols in the semi-arid tropics. I. Modelling the effects of soil management on runoff and erosion. Aust. J. Soil. Res. 34, 91-102.

Littleboy, M., McGarry, D and Bray, S. 1998. A combination of modelling and soil data to determine risk of soil compaction. Proceedings, National Soils Conference, Brisbane, pp 462-465. Australian Society of Soil Science Incorporated.

Loch, R.J., Rosewell, C.J., 1992. Laboratory methods for measurement of soil erodibilities (K factors) for the Universal Soil Loss Equation. Aust. J. Soil. Res.30, 233-248.

Loch, R.J., Slater, B.K., Devoil, C., 1998, Soil erodibility (K(m)) values for some Australian soils. Aust. J. Soil. Res. 36, 1045-1055.

McClymont, D., Freebairn, D.M., Rattray, D.J., Robinson, J.B., White, S., 2016. Howleaky:Exploring Water Balance and Water Quality Implication of Different Land Uses. Software V5.49.19. http://howleaky.net/. (Accessed 23 January 2018).

McHugh, A.D., Tullberg, J.N., Freebairn, D.M., 2009. Controlled traffic farming restores soil structure. Soil Till. Res. 104, 164–172.

McMillan, H., Jackson, B., Clark, M., Kavetski, D. and Woods, R., 2011. Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. J. Hydrol., 400, 83-94.

Melland, A.R., Vigiak, V., Roberts, A.M., Rattray, D., Whitford, J., 2010. Evaluation of a static water balance model in cropped and grazed systems of temperate Australia. Environmen. modell. Softw. 25, 1682-1691.

Moody, P.W., 2007. Interpretation of a single-point P buffering index for adjusting critical levels of the Colwell soil P test. Aust. J. Soil. Res. 45, 55-62.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., Veith, T. L. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. T. Asabe. 50, 885-900.

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. J. hydrol. 10, 282-290.

Neitsch, S. L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2001. Soil and Water Assessment Tool, Theoretical Documentation. Temple, Texas, United States Department of Agriculture.

Pechlivanidis, I.G., Jackson, B.M., McIntyre, N.R. and Wheater, H.S., 2011. Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. Global. Nest. J. 13, 193-214.

Queensland Government, 2019. HowLeaky Model V5 Technical Documentation: Version 1.06. Department of Natural Resources, Mines and Energy and the Department of Environment and Science.

Rattray, D.J., Silburn, D.M., Owens, J., Barlow, L., 2006. Combining ‘paddock models’ and geographical information systems can be a powerful too in catchment planning. 30th Hydrology and Water Resources Symposium. 4-7 Dec 2006, Launceston, Tasmania, Institution of Engineers Aust.

Ritchie, J.T., 1972. Model for predicting evaporation from a row crop with incomplete cover. Water Resour. Res. 8, 1204-1213.

Robinson, J.B., Rattray, D.J., Freebairn, D.M., Silburn, D.M., McClymont, D., 2007. Using a Simple Hydrologic Model to Link Management to Nutrient Concentrations and Loads in Runoff. MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December. 2007, 74-80.

Robinson, J.B., Silburn, D.M., Rattray, D., Freebairn, D.M., Biggs, A., McClymont, D., Christodoulou, N., 2010. Modelling shows that the high rates of deep drainage in parts of the Goondoola Basin in semi-arid Queensland can be reduced with changes to the farming systems. Aust. J. Soil. Res. 48, 58-68.

Rosewell, C.J., 1992. The erodibility of five soils in New South Wales. In Proceedings of the 5th Australian Soil Conservation Conference. Vol. 3, pp. 112–115. (Eds. GJ Hamilton, KM Howes, R Attwater) (Department of Agriculture: Perth).

Sharpley A.N. 1995. Dependence of runoff phosphorus on extractable soil phosphorus. J. Environ. Qual. 24, 920-926.

Sharpley, A.N., 2006. Modeling phosphorus movement from agriculture to surface waters. In: D.E. Radcliffe and M.L. Cabrera, Modeling phosphorus in the environment. CRC Press, Boca Raton, FL.

Shaw, M., Silburn, D.M., Thornton, C., Robinson, J.B., McClymont, D., 2011. Modelling pesticide runoff from paddocks in the Great Barrier Reef using HowLeaky 19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011.

Silburn, D.M., 1994. Field trip notes Eastern Darling Downs and Lockyer Valley (http://www.howleaky.net/index.php/library/supersites/65-library/site-summaries/pastures/mt-mort/188-mt-mort).

Silburn, D.M., 2003. Characterising Pesticide Runoff from Soil on Cotton Farms using a Rainfall Simulator. Unpublished Ph.D., Faculty of Agriculture, Food and Natural Resources, University of Sydney.

Silburn, D.M., Robinson, J.B., Freebairn, D.M., 2007. Why restore marginal cropland to permanent pasture? - land resource and environmental issues. Trop. Grasslands. 41, 139-153.

Sorooshian S., Gupta V.K., 1995. Model Calibration, Computer models of watershed hydrology, edited by Singh, V.P., Water Resources Publications, USA.

The State of Queensland Government, 2013. Reef Water Quality Protection Plan 2013 Published by the Reef Water Quality Protection Plan Secretariat, July 2013.

The State of Queensland Government, 2018. Reef 2050 Water Quality Improvement Plan, 2017-2022.

Thornton, C.M., Cowie, B.A., Freebairn, D.M., Playford, C.L., 2007. The brigalow catchment study: II. Clearing brigalow (Acacia harpophylla) for cropping or pasture increases runoff. Aust. J. Soil. Res. 45,496–511.

Thornton, C.M., Elledge, A.E., 2013. Runoff of Nitrogen, Phosphorus and Sediment from Pasture Legumes: An Enhancement to Reef Catchment Modelling (Project RRRD009). Report to the Reef Rescue Research and Development Program. ISBN: 978-1-925088-11-3.

Thomas EC, Gardner EA, Littleboy M and Shields PJ. 1995. The cropping systems model PERFECT as a quantitative tool in land evaluation: An example for wheat cropping in the Maranoa area of Queensland. Aust. J. Soil. Res. 33,535-554.

Tullberg JN, Ziebarth PJ, Li Y (2001) Tillage and traffic effects on runoff. Aust. J. Soil. Res. 39, 249-257.

Vigiak, O., Newham, L.T., Whitford, J., Roberts, A.M., Rattray, D. and Melland, A.R., 2011. Integrating farming systems and landscape processes to assess management impacts on suspended sediment loads. Environmen. modell. softw. 26, 144-162.

Wang, X., Williams, J., Gassman, P., Baffaut, C., Izaurralde, R., Jeong, J. & Kiniry, J. 2012. EPIC and APEX: Model use, calibration, and validation. T. Asabe. 55, 1447-1462.

Waterhouse, J., Brodie, J., Tracey, D., Smith, R., Vandergragt, M., Collier, C., Petus, C., Baird, M., Kroon, F., Mann, R., Sutcliffe, T., Waters, D., Adame, F., 2017. Scientific Consensus Statement 2017: A synthesis of the science of land-based water quality impacts on the Great Barrier Reef, Chapter 3: The risk from anthropogenic pollutants to Great Barrier Reef coastal and marine ecosystems. State of Queensland, 2017.

Waters, D.K., 2004. Grazing management implications and runoff and erosion processes in semi-arid Central Queensland. ISCO 2004 – 13th International Soil Conservation Organisation Conference Brisbane July 2004.

Waters, D.K., 2009. Keilambete Brief summary. Notes prepared as part of a modelling exercise using the Howleaky model with Keilambete (www.howleaky.net).

Welter, D.E., Doherty, J.E., Hunt, R.J., Muffels, C.T., Tonkin, M.J., Schreuder, W.A., 2012. Approaches in highly parameterized inversion: PEST++, a parameter estimation code optimized for large environmental models. US Geol Surv Tech Methods, Book 7, Section C5, 47 pp.

Wightman, B., Peries, R., Bluett, C., Johnston, T., 2005. Permanent raised bed cropping in southern Australia: practical guidelines for implementation. In ACIAR Proceedings No. 121 – Evaluation and performance of permanent raised bed cropping systems in Asia, Australia and Mexico, Eds. CH Roth, RA Fischer and CA Meisner, 173-190.

Williams, J., Bui, E.N., Gardner, E.A., Littleboy, M. and Probert, M.E., 1997. Tree clearing and dryland salinity hazard in the upper Burdekin catchment of north Queensland. Soil. Res. 35, 785-802.

Williams, J.R. and LaSeur, W.V., 1976. Water yield model using SCS curve numbers. J. Hydr. Div-ASCE. 102, 12379.

Williams, J.R., 1983. 'EPIC, The Erosion-Productivity Impact Calculator, Volume 1. Model Documentation.' Agricultural Research Service, United States Department of Agriculture.

Wischmeier, W. H. and Smith, D. D. 1978. Predicting rainfall soil erosion losses – a guide to conservation planning. U.S. Department of Agriculture, Agriculture Handbook No. 537

Wooldridge, S.A., 2009. Water quality and coral bleaching thresholds: Formalising the linkage for the inshore reefs of the Great Barrier Reef, Australia. Marine Pollution Bulletin. 58, 745-751.

Web references

The State of Queensland Government, 2018, Scientific Consensus Statement. Available at (https://www.reefplan.qld.gov.au/about/reef-science/scientific-consensus-statement). (accessed 2.10.2018).

Jane Waterhouse, Britta Schaffelke, Rebecca Bartley, Rachel Eberhard, Jon Brodie, Megan Star, Peter Thorburn, John Rolfe, Mike Ronan, Bruce Taylor and Frederieke Kroon. 2017. 2017 Scientific Consensus Statement. Land use impacts on Great Barrier Reef water quality and ecosystem condition. Available at (https://www.reefplan.qld.gov.au/about/assets/2017-scientific-consensus-statement-summary.pdf. (accessed 2.10.2018).

HowLeaky webiste, University of Southern Queensland, www.howleaky.net, this will become www.howleaky.com.

20 March 2020

Environmental Modelling & Software
Dear Professor Jakeman

Hereby, I am writing on behalf of all authors that there is no any conflict of interest related to our manuscript:

"A pragmatic parameterisation and calibration approach to model hydrology and water quality of agricultural landscapes and catchments".

Sincerely yours,

Dr. Afshin Ghahramani

Yours sincerely

Dr Afshin Ghahramani
Research Fellow | Agricultural Systems Modelling
Institute for Life Sciences and the Environment| University of Southern Queensland
| Toowoomba, Queensland | 4350 | Australia
E: Afshin.Ghahramani@usq.edu.au
T: +61 7 4631 1178 | M: +61 402 485 046

- A pragmatic modelling by Howleaky is suitable to link management to water quality
- A manually calibrated model may still require a reduction in error variances
- Data availability and inherent error propagation determined confidence in modelling