

The influence of model simplicity on uncertainty in the context of surface – groundwater modelling and integrated assessment

R.S. Blakers^{a,b}, **B.F.W. Croke**^{a,b} and **A.J. Jakeman**^a

^a *Integrated Catchment Assessment and Management (iCAM) Centre, Fenner School of Environment and Society, National Centre for Groundwater Research and Training, Australian National University, Australia*

^b *Department of Mathematics, Australian National University, Australia*

Email: rachel.blakers@anu.edu.au

Abstract: In the past, water resources in Australia have been over-allocated and this has led to the implementation of water reforms. An integrated assessment project is underway in the Namoi Catchment, located in New South Wales, Australia, that aims to assess the social, economic and ecological impacts of reductions in water allocations under current and future environmental conditions. Surface and groundwater resources in the Namoi are heavily utilised for irrigation agriculture and farming, are an essential source of drinking water and support ecologically significant riparian zones.

The integrated assessment project considers the impacts of a range of climate scenarios and water policy options and investigates the benefits of adaptations in land management practises and innovations in water use efficiency. This paper focuses on the development of the hydrological component of the integrated assessment framework. The requirements of the hydrological model are that it should take climate and water extraction data and produce predictions of daily surface water flows and monthly groundwater levels. These hydrological impacts are transferred to an assessment of water availability and its social, economic and ecological outcomes.

The paper discusses the trade-offs between model simplicity and complexity with respect to predictive uncertainty, equating model complexity with the number of parameters. A particular issue in hydrological modelling is that there is often insufficient data available to define the system, leading to underdetermined models and non-uniqueness of parameter values. The implication is that, while highly parameterised models generally contain enough flexibility to fit observation data well, they may have poor predictive performance due to parameter uncertainty. There are two main methods of dealing with the ‘complexity paradox’, the first being model parsimony and the second being the employment of a regularisation approach to stabilise the inversion of a highly parameterised model. While regularisation can be highly effective, there are costs associated with complexity that may include onerous data requirements and time-consuming development.

The socioeconomic component of the integrated assessment project divides the Lower Namoi Catchment into a number of large zones that are considered homogenous with respect to the variables of interest. This paper argues that, for the purposes of the project, a simple model of surface and groundwater interactions is likely to perform as well as a more complex model when measured at the coarse spatial scale at which the results will be utilised. The proposed model utilises a catchment scale conceptualisation of the hydrological processes occurring, and is an extension of the IHACRES rainfall-runoff model. Finally, the paper presents a discussion of the uncertainty analysis approach that will be used to assess the performance of the model.

Keywords: *Surface-groundwater interactions, integrated assessment, uncertainty, IHACRES.*

1. INTRODUCTION

Water resources in Australia are under increasing pressure due to rising water consumption and changing climatic conditions. In the past, more water entitlements have been allocated than is sustainable. The over-allocation of water resources has led to the implementation of water reforms, resulting in reductions in surface and groundwater allocations being set in many catchments. Catchment managers and the planning process require a means of assessing the likely social, economic and ecological impacts of water allocation decisions under current and future environmental conditions. The breadth and complexity of the management problem lends itself to integrated, often model based, assessment approaches.

In response to this, an integrated assessment project is underway in the Namoi Catchment to assess the socioeconomic and ecological consequences of scenarios and adaptations that consider a changing and varying climate, water policy options, farm production activities and innovations in water use efficiency. The project engages the expertise of industry and government groups in the Namoi Catchment to encourage knowledge sharing, identify major risks to water security, and explore opportunities for policy and landholder innovations that minimise socioeconomic and environmental impacts. At the core of the assessment is a hydrological model that predicts the effects of surface and groundwater extraction regimes on surface water flows and aquifer storage and discharge. These hydrological impacts are transferred to an assessment of water availability and its social, economic and ecological outcomes.

Uncertainty in hydrological modelling, and environmental modelling in general, is often very large. Errors in the model structure, inputs, parameters and outputs arise from insufficient understanding of the system, simplifications in the representation of the processes occurring, spatial and temporal discretisation, limitations in the data collection process and inherent system variability. Hence, to be useful in decision-making, the integrated model must provide a measure of the uncertainty in the output variables of interest.

The focus of this paper is the development of the hydrological component of the integrated modelling framework. The trade-offs between a pragmatic modelling approach and minimising uncertainty in the model are discussed and it is argued that, in this particular application, a simple, conceptual model is likely to yield the most benefits.

2. CASE STUDY AREA

The Namoi Catchment is located in the Murray-Darling Basin in northern New South Wales, Australia. It is an important agricultural region supporting industries including cotton, livestock production, grain and hay, poultry and horticulture. The catchment has one of the highest levels of groundwater extraction in Australia, accounting for around 15% of the total annual groundwater use from the Murray-Darling Basin, and groundwater levels in some areas of the catchment have dropped by more than 10 metres during the last 40 years (Giambastiani *et al.*, 2009).

The Lower Namoi Catchment, being the focus of this paper, is an alluvial floodplain occupying an area of around 5,100 km². The alluvial sediments are generally divided into three formations of which the Gunnedah and Cubbaroo Formations provide the productive aquifers for irrigation. The sediments of the Cubbaroo Formation infill a narrow palaeochannel running along the northeast boundary of the catchment. The Gunnedah Formation overlies the Cubbaroo Formation, covering the bedrock surface in the remainder of the catchment. Exposed at the ground surface is the Narrabri Formation. Refer to Williams (1986) and McLean (2003) for further details of the hydrogeology.

3. MODEL COMPLEXITY ISSUES

Hydrological systems, as with environmental systems in general, are inherently complex. In an effort to describe and understand this complexity, and enabled by rapid advances in computing technology, scientists are drawn to increasingly detailed models of the real world processes. Generally, as the level of model conceptual complexity increases so does the number of parameters. Groundwater models in particular tend

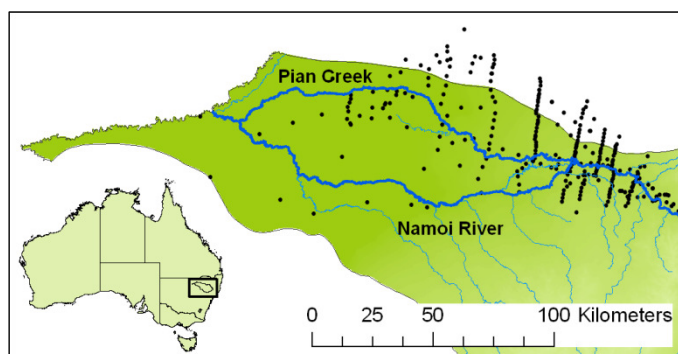


Figure 1. Lower Namoi Catchment map showing the river network and the distribution of groundwater monitoring sites.

to be highly parameterized, involving the numerical solution of physics-based partial differential equations across intensively discretised spatial domains. Examples include the finite difference software MODFLOW (Harbaugh *et al.*, 2009) and the finite element software FEFLOW (DHI-WASY, 2009). These types of models are valuable tools in developing process understanding and they are often applied to prediction and forecasting. However, there is considerable debate regarding the appropriate level of model complexity and the way in which model development should be approached (e.g. Hill, 2006). In this paper, the influence of model parameterisation on uncertainty is of particular interest, with the number of parameters being used as a measure of complexity.

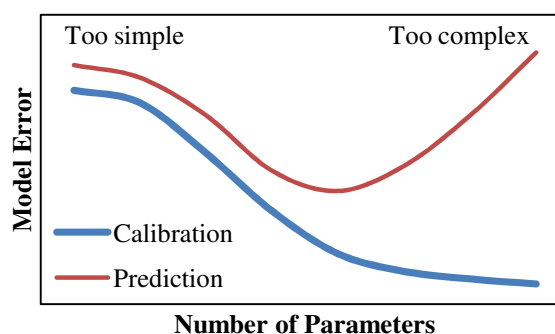


Figure 2. An idealised view of the influence of model complexity on calibration and prediction error.

In hydrological modelling, there is often limited high quality data available with which to parameterise, calibrate and validate models. As a result, model identification tends to be an issue in that there is insufficient information in the data to estimate parameter values uniquely. The implication of this is that while highly parameterised models generally contain enough flexibility to fit observations well, the selected parameter set may not be physically meaningful. Secondly, the model may be fitting variations in the observations that represent noise rather than system behaviour. These factors contribute to the poor predictive performance of some highly parameterised models (stylised in Figure 2) and lead to the complexity paradox; “the closer a model comes to capturing the full range of processes and parameters in the system being modelled, the more difficult it is to ascertain whether or not the model faithfully represents that system” (Oreskes, 2003).

There are two main methods of addressing the complexity issue. The first is model parsimony – using only as many parameters as can be uniquely identified from the data. The disadvantage of this approach is that structural error may dominate model predictions due to over-simplification of both the system processes and spatiotemporal variability. The second method is to parameterise the model with more parameters than can be uniquely determined from the data and employ one of the many regularisation methods available to stabilise the model inversion, guaranteeing convergence to a unique, or at least better defined, parameter set. Note that the impact of the regularisation method on the uncertainty in predictions needs to be assessed. The hybrid regularisation scheme of Tonkin and Doherty (2005), for example, is computationally efficient and allows the information content of the data to drive the model parameterisation and the level of spatial variability. The disadvantage remains that highly parameterised models tend to be time consuming and expensive to develop, data intensive and computationally demanding. From a practical perspective, a simpler model that performs equally well, or nearly as well, as a more complex model with respect to the objectives of the modelling exercise is often preferable.

4. HYDROLOGICAL MODEL OBJECTIVES

The integrated assessment framework for the project requires a hydrological model that can predict:

- daily surface water flows through the Lower Namoi Catchment, including contributions from surface runoff and baseflow
- monthly groundwater levels at selected locations in the catchment, for use in ecological modelling and water availability assessment.

The socio-economic model, as is common with such models, divides the catchment into a number of large regions that are considered homogenous with respect to land use and farm management practices (e.g. Letcher *et al.* 2004). The coarse representation is partly due to practical constraints, which include the time and resources available to develop the model and the need to limit run times of the integrated model to reasonable levels. The outcome is that uncertainty only needs to be assessed at a large spatial scale.

The standard in groundwater modelling for the Lower Namoi Catchment is currently a three-layer MODFLOW model, which represents the system as an upper unconfined aquifer overlying two semi-confined aquifers (Merrick, 2001; Williams *et al.*, 1989). The hypothesis presented in this paper is that a much simpler hydrological model will perform well when measured at the coarse spatial scale at which the results will be utilised.

5. HYDROLOGICAL MODEL FORMULATION

The proposed hydrological model employs a catchment scale conceptualisation of the Lower Namoi and it is comprised of:

1. a surface and groundwater model that takes rainfall, temperature and extraction data, and estimates stream flow and groundwater storage volumes
2. a stream flow routing module that uses the variable parameter lag-route method of Croke *et al.* (2006) to route flow downstream through the sub-catchment network
3. a groundwater level module that converts groundwater storage volumes to levels via a regression analysis.

The integrated surface-groundwater model is based on the IHACRES model (Jakeman and Hornberger, 1993; Croke and Jakeman, 2004) and is described in detail here. The basic IHACRES model is a catchment scale rainfall-runoff model consisting of two modules. The first is a non-linear loss module that takes rainfall and temperature data and produces ‘effective rainfall’, accounting for losses due to evapotranspiration. The second is a linear routing module that models stream flow as a linear combination of antecedent stream flow and effective rainfall. A number of versions of the IHACRES model have been developed over the years and it has been successfully extended to modelling surface and groundwater interactions (e.g. Ivkovic *et al.*, 2009; Herron and Croke, 2009b). The model presented in this paper is based on the IHACRES_GW version of the linear routing module (Ivkovic *et al.* 2005, 2009) coupled with the CMD (catchment moisture deficit) version of the non-linear loss module (Croke and Jakeman, 2004). The primary contribution of this paper is the extension of the linear routing module to allow specification of aquifers whose boundaries do not coincide with catchment boundaries. A second modification to the Ivkovic *et al.* (2005) model is the addition of a parameter representing infiltration of surface flows to groundwater.

The structure of the model is illustrated in Figure 3 and the equations of the new linear routing module are described below. In the equations, bracketed superscripts indicate whether a variable is an attribute of a sub-catchment, an aquifer, or both. Each sub-catchment is assigned a number $i = 1, \dots, N$, where N is the number of catchments. The catchments are ordered with catchment number 1 being at the head of the stream network and catchment N being the furthest downstream. Groundwater aquifers are numbered $j = 1, \dots, M$, where M is the number of aquifers. For time dependent variables, the subscript k represents the time step, where $k = 1, \dots, T$ and T is the total number of time steps.

The IHACRES linear module converts effective rainfall to stream flow via two parallel transfer functions representing a quick-flow pathway (equated to surface runoff) and a slow-flow pathway (representing groundwater discharge). The sum of the two exponential functions gives the modelled stream flow. The contribution of rainfall-runoff and groundwater discharge to stream flow, $Q_k^{(i)}$, in the i th catchment is:

$$Q_k^{(i)} = Q_k^{q,(i)} + Q_k^{s,(i)} \quad (1)$$

where $Q_k^{q,(i)}$ is quick-flow volume, $Q_k^{s,(i)}$ is slow-flow volume.

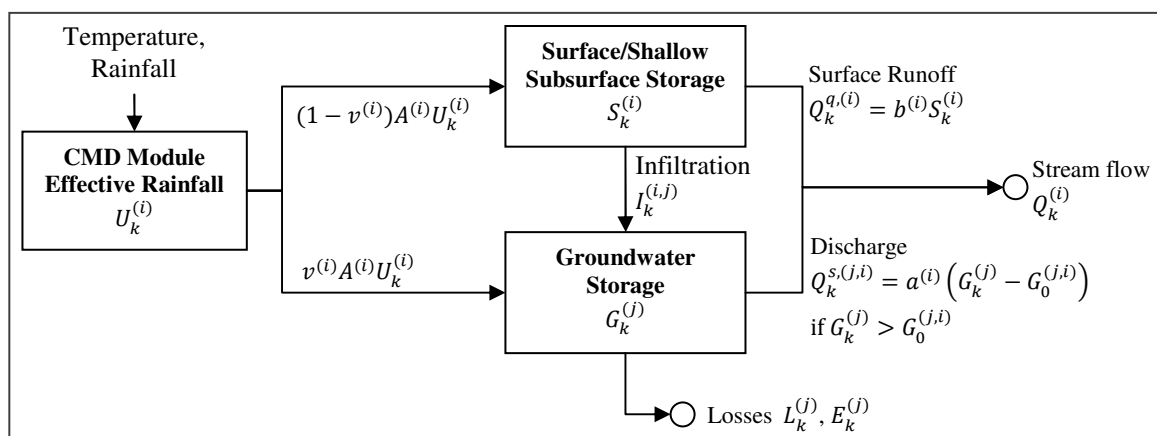


Figure 3. Structure of the new linear module of the IHACRES model, showing a single catchment connected to a single aquifer ($f^{(j,i)} = 1$).

5.1. Quick-flow Equations

The volume of surface runoff contributing to stream flow is modelled as a constant proportion of the storage:

$$Q_k^{q,(i)} = b^{(i)} S_k^{(i)} \quad (2)$$

where $S_k^{(i)}$ is the quick-flow storage volume, $b^{(i)}$ is a dimensionless constant representing the fraction of surface runoff and $0 \leq b^{(i)} \leq 1$. The mass balance equations for quick-flow storage are:

$$\begin{aligned} S_k^{(i)} &= S_k^{*(i)} - Q_k^{q,(i)} - I_k^{(i)} \\ S_k^{*(i)} &= S_{k-1}^{(i)} + (1 - v^{(i)}) A^{(i)} U_k^{(i)} \end{aligned} \quad (3)$$

where $S_k^{*(i)}$ is the storage volume at the k th time step before accounting for losses to the stream and infiltration to groundwater, $I_k^{(i)}$ is the infiltration to groundwater, $A^{(i)}$ is the catchment area, $U_k^{(i)}$ is the total effective rainfall per unit area, $(1 - v^{(i)})$ is the proportion of effective rainfall partitioned to quick-flow and $0 \leq v^{(i)} \leq 1$.

Infiltration to groundwater from the stream is assumed to absorb all the available water until a maximum infiltration rate, $I_{max}^{(i)}$, is reached:

$$I_k^{(i)} = \begin{cases} I_{max}^{(i)} & \text{if } S_k^{*(i)} > I_{max}^{(i)} \\ S_k^{*(i)} & \text{otherwise} \end{cases} \quad (4)$$

The infiltration is apportioned to each of the aquifers receiving recharge from the catchment. The infiltration received by the j th aquifer from the i th catchment is:

$$I_k^{(j,i)} = f^{(j,i)} I_k^{(i)} \quad (5)$$

where $f^{(j,i)}$ is the recharge fraction, $0 \leq f^{(j,i)} \leq 1$ and $\sum_j f^{(j,i)} = 1$.

5.2. Slow-flow Equations

Discharge from the aquifer to the stream is assumed to occur when the aquifer volume exceeds a threshold. As with quick-flow, the discharge volume is modelled as a constant fraction of the excess. The slow-flow contribution, $Q_k^{s,(j,i)}$, of the j th aquifer to stream flow in the i th catchment is:

$$Q_k^{s,(j,i)} = \begin{cases} a^{(i)} (G_k^{(j)} - G_0^{(j,i)}) & \text{if } G_k^{(j)} > G_0^{(j,i)} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where $G_k^{(j)}$ is the groundwater storage volume, $a^{(i)}$ is a dimensionless storage coefficient, and $G_0^{(j,i)}$ is the threshold volume below which no discharge occurs within the i th catchment. The groundwater discharge to each of the catchments is calculated iteratively, starting with the catchment at the head of the flow network and proceeding downstream. Note that the groundwater storage volume is relative to the discharge threshold for the catchment furthest downstream, and as such does not represent the total volume of the aquifer.

The mass balance equations for groundwater storage are:

$$G_k^{(j)} = G_{k-1}^{(j)} + \sum_i (f^{(j,i)} v^{(i)} A^{(i)} U_k^{(i)}) + \sum_i (f^{(j,i)} I_k^{(i)}) - \sum_i (Q_k^{s,(j,i)}) - L_k^{(j)} - E_k^{(j)} \quad (7)$$

where $v^{(i)}$ is the proportion of effective rainfall partitioned as slowflow, $L_k^{(j)}$ represents losses to deep groundwater, and $E_k^{(j)}$ is the groundwater extraction volume.

For the i th catchment, the total baseflow contribution to stream flow, $Q_k^{s,(i)}$, is the sum over the contributions from each of the aquifers:

$$Q_k^{s,(i)} = \sum_j Q_k^{s,(j,i)} \quad (8)$$

In its current form, the model allows representation of an aquifer as a single vertically connected layer only and assumes that there is no flow between aquifers. These two aspects of the model are likely to require refinement.

5.3. Numerical Procedure

The above formulation cannot be solved via a time-stepping numerical procedure due to the simultaneous dependence of quick-flow, $Q_k^{q,(i)}$, and the surface water storage volume, $S_k^{(i)}$, and the simultaneous dependence of slow-flow, $Q_k^{s,(i)}$, and the groundwater storage volume, $G_k^{(j)}$. The solution of the equations follows a similar method to that given in Ivkovic *et al.* (2005), hence just the results are presented here.

The quick-flow contribution to stream flow is calculated as:

$$Q_k^{q,(i)} = \frac{b^{(i)}}{1 + b^{(i)}} (S_k^{*(i)} - I_k^{(i)}) \quad (9)$$

The quick-flow storage volume (before losses to the stream and infiltration) is given by:

$$S_k^{*(i)} = \frac{1}{b^{(i)}} Q_{k-1}^{q,(i)} + (1 - v^{(i)}) A^{(i)} U_k^{(i)} \quad (10)$$

The groundwater storage volume of the j th aquifer, before accounting for discharge to the stream, is given by:

$$G_k^{*(j,0)} = G_{k-1}^{(j)} + \sum_i (f^{(j,i)} v^{(i)} A^{(i)} U_k^{(i)}) + \sum_i (f^{(j,i)} I_k^{(i)}) - L_k^{(j)} - E_k^{(j)} \quad (11)$$

The storage volume, $G_k^{*(j,i)}$, of the j th aquifer after discharge to the i th catchment is:

$$G_k^{*(j,i)} = \begin{cases} \frac{1}{1 + a^{(i)}} (G_k^{*(j,i-1)} + a^{(i)} G_0^{(j,i)}) & \text{if } G_k^{*(j,i-1)} > G_0^{(j,i)} \\ G_k^{*(j,i-1)} & \text{otherwise} \end{cases} \quad (12)$$

The corresponding slow-flow contribution, $Q_k^{s,(j,i)}$, of the j th aquifer to stream flow in the i th catchment is:

$$Q_k^{s,(j,i)} = \begin{cases} a^{(i)} (G_k^{*(j,i)} + G_0^{(j,i)}) & \text{if } G_k^{*(j,i-1)} > G_0^{(j,i)} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

The final groundwater storage volume accounting for all discharges is $G_k^{(j)} = G_k^{*(j,N)}$.

6. UNCERTAINTY ANALYSIS APPROACH

The purpose of uncertainty analysis of the hydrologic model is to assess confidence in the predictions with respect to the criteria (e.g. impact on management decisions) of other components of the integrated framework. A separate analysis of the project objectives and the integrated model as a whole will identify detailed criteria. A distinguishing feature of the modelling exercise is that the socioeconomic and ecological models utilise coarse spatial and temporal scales. The proposed uncertainty analysis will explore the efficacy of a simple hydrological model and the level of structural detail that minimises predictive uncertainty in the variables of interest.

To investigate the effect of spatial scale the analysis will consider model performance when the Lower Namoi Catchment is represented as a single hydrological zone compared to dividing the region into varying numbers of smaller sub-catchments and aquifers. While this paper presents a formulation of the model that allows just one vertically connected aquifer, the model can easily be extended to include multiple groundwater stores (Herron and Croke, 2009b). The analysis will assess the benefits of allowing up to two or three stores representing the layered aquifer structure of the Lower Namoi Catchment.

To reduce the size of the uncertainty analysis problem, a sensitivity assessment of the model inputs, parameters and structure will identify critical components. This will be followed by an analysis of uncertainty in the model inputs, which include rainfall, temperature, stream flow and groundwater extraction data. Uncertainty in areal rainfall estimates, in particular, is likely to be a limiting factor in the accuracy of the model (Herron and Croke, 2009a). To the extent possible, the performance of the model will be compared to that of other more complex models in the same region.

7. CONCLUSION

There are advantages and disadvantages in both model simplicity and complexity, the optimal approach depending on the objectives of the particular modelling exercise and the resources available to the modeller.

This paper has presented a discussion of uncertainty in the context of developing a hydrological model suitable for integration with socioeconomic and ecological modelling. The hypothesis is presented that a very simple model of surface and groundwater interactions will provide adequate predictions given the tightly defined objectives of the integrated assessment project. The structure and equations of the proposed model are presented and the future uncertainty assessment approach discussed. Rigorous uncertainty analysis of this model will provide a valuable guide in the choice of hydrological model components for future integrated assessment projects of this type.

REFERENCES

- Croke, B.F.W., and Jakeman, A.J., (2004). A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environmental Modelling & Software*, 19, pp. 1 – 5.
- Croke, B.F.W., Letcher, R.A., and Jakeman, A.J., (2006). Development of a distributed flow model for underpinning assessment of water allocation options in the Namoi River Basin, Australia. *Journal of Hydrology*, 39, pp. 51 – 71.
- DHI-WASY, (2009). FEFLOW finite element subsurface flow and transport simulation system - Reference manual. Technical Report, DHI-WASY GmbH, Berlin. (<http://www.feflow.info>).
- Giambastiani, B.M.S., Kelly, B.F.J., The, C., Andersen, M.S., McCallum, A.M., Acworth, R.I., (2009). 3D time and space analysis of groundwater head change for mapping river and aquifer interactions. In Anderssen, R.S., Braddock, R.D., and Newham, L.T.H., (eds), proceedings of the 18th World IMACS Congress and MODSIM09, Cairns, Australia, July 2009, pp. 2377-2383.
- Harbaugh, A.W., Banta, E.R., Hill, M.C., and McDonald, M.G., (2000). MODFLOW-2000, the U.S. Geological Survey modular ground-water model – User guide to modularization concepts and the ground-water flow process: U.S. Geological Survey Open-File Report 00-92, 121 p.
- Herron, N., and Croke, B.F.W., (2009a). Including the influence of groundwater-exchanges in a lumped rainfall-runoff model. *Mathematics and Computers in Simulation*, 79, pp. 2689-2700.
- Herron, N., and Croke, B.F.W., (2009b). IHACRES-3S – A 3-store formulation for modelling groundwater-surface water interactions. In Anderssen, R.S., Braddock, R.D., and Newham, L.T.H., (eds), proceedings of the 18th World IMACS Congress and MODSIM09, Cairns, Australia, July 2009, pp. 3081-3087.
- Hill, M.C., (2006). The practical use of simplicity in developing ground water models. *Ground Water*, 44(6), pp. 775–781.
- Ivkovic, K.M., Croke, B.F.W., Letcher, R.A., and Evans, W.R., (2005). The development of a simple model to investigate the impact of groundwater extraction on river flows in the Namoi Catchment, NSW, Australia, in: “Where Waters Meet” NSHS-IAH-NSSSS Conference, 28 November – 2 December, Auckland, New Zealand.
- Ivkovic, K.M., Letcher, R.A., and Croke, B.F.W., (2009). Use of a simple surface–groundwater interaction model to inform water management. *Australian Journal of Earth Sciences*, 56, pp. 71 – 80.
- Jakeman, A.J., and Hornberger, G.M., (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resources Research*, 29(8), pp. 2637-2649.
- Letcher, R.A., Jakeman, A.J., and Croke, B.F.W., (2004). Model development for integrated assessment of water allocation options. *Water Resources Research*, 40, W05502, (doi:10.1029/2003WR002933).
- McLean, W.A., (2003). Hydrogeochemical evolution and variability in a stressed alluvial aquifer system: Lower Namoi River Catchment, NSW. PhD Thesis, School of Biological, Earth & Environmental Science, University of New South Wales, Sydney, Australia.
- Merrick, N.P., (2001). Lower Namoi Groundwater Flow Model: Calibration 1980-1998. Insearch Limited Report for NSW Department of Land and Water Conservation, Project No. C99/44/001, August 2001, 91p.
- Oreskes, N., (2003). The role of quantitative models in science, pp. 13-31 in *Models in Ecosystem Science*, edited by Canham, C.D., Cole, J.J., and Lauenroth, W.K.. Princeton University Press, Princeton.
- Tonkin, M.J., and Doherty, J., (2005). A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resources Research*, 41, W10412, doi:10.1029/2005WR003995.
- Williams, R.M., (1986). The Cainozoic geology, hydrogeology and hydrochemistry of the unconsolidated sediments associated with the Namoi River in the Lower Namoi Valley, NSW Department of Water Resources, Hydrogeology Report 1986/4.
- Williams, R.M., Merrick, N.P., and Ross, J.B., (1989). Natural and induced recharge in the Lower Namoi Valley, New South Wales. In Sharma, M.L. (ed.), *Groundwater Recharge*, Proceedings of the Symposium on Groundwater Recharge, Mandurah, 6-9 July 1987. A.A. Balkema, Rotterdam, pp. 239-253.