Exploring the utility of multi-response calibration in river system modelling

T. Micevski [a], J. Lerat [b], D. Kavetski [a], M. Thyer [c], and G. Kuczera [a]

[a] School of Engineering, University of Newcastle, New South Wales
[b] Land and Water, CSIRO, Australian Capital Territory
[c] School of Civil Environmental and Mining Engineering, University of Adelaide, South Australia
Email: tom.micevski@newcastle.edu.au

Abstract: Water allocation models can be used to compare water sharing scenarios in regulated catchments to evaluate the effects on both the water users and the environment. These models include a representation of the physical system with modules such as flow routing, rainfall-runoff modelling or groundwater/surface water interactions, as well as management components to take into account infrastructure such as dams, canals or extraction points. Water allocation models can be complex modelling structures with a large number of parameters to be calibrated on limited datasets, especially regarding the management aspects. Additionally, these models are used as a tool in the making of long-term decisions with important social and environmental impacts. As a result, the assessment of uncertainty becomes a critical task to inform the decision-makers about the likely robustness of the model analysis and predictions.

Calibration of these models is currently problematic. In particular, the errors affecting system observations are often not properly accounted for, which is a concern since these errors may be quite large. Furthermore, calibration is often performed separately on various components of the system, resulting in inconsistencies when the components are linked. These deficiencies make it difficult to quantify the uncertainty in the predictions of the entire system performance.

The Bayesian approach provides a platform to directly address the sources of uncertainty (input, output, and model error) in the model calibration and prediction process. This study seeks to develop a Bayesian multi-response method for use with river system models, allowing joint calibration to all sources of information available in a particular application. Unlike the traditional approach, joint calibration forces consistency in performance across the entire system. Moreover, the Bayesian approach provides a framework for a proper accounting of uncertainty both in the inferred parameters and in the model predictions.

This study illustrates the application of the Bayesian multi-response calibration approach to the STICKMAN model, a simplified river system model which describes key aspects of complex river basin models such as IQQM but is computationally less demanding. The model was calibrated using a Weighted Least Squares method in a synthetic data study. Model calibration used both single and multiple response data (eg. streamflow at the outlet and at internal system nodes, reservoir time series, etc.) to investigate the improvements in parameter estimation associated with the inclusion of additional responses.

The use of multiple response data during model calibration was generally found to reduce parameter uncertainty. However, the extent of reductions in uncertainty depended on which responses were included, highlighting that some sources of data are more informative than others. This supports the findings of Kuczera and Mroczkowski's (1998), who conclude that the value of new sources of response data should be assessed *a priori* before embarking on (potentially expensive) field campaigns. This study reports the first findings in this project. Future work will explore the effects of multiple response data on model predictive performance, further develop the STICKMAN model to better represent processes and errors, and finally consider IQQM case studies.

Keywords: Uncertainty, river system model, model calibration, Bayesian analysis

1. INTRODUCTION

Water allocation models are used to compare water sharing scenarios in regulated catchments in terms of the impacts for the water users and the environment. These models include a representation of the physical system with modules such as flow routing, rainfall-runoff modelling or groundwater/surface water interactions, as well as management components to take into account infrastructure such as dams, canals or extraction points. Water allocation models can be complex modelling structures with a large number of parameters to be calibrated on limited datasets, especially regarding the management aspects. At the same time, water allocation models are used to aid in the making of long-term decisions with important social and environmental impacts. As a result, the assessment of uncertainty becomes a critical task to inform the decision makers about the strengths and weaknesses of the model.

Calibration of these models is currently problematic. Errors affecting system observations are often not properly accounted for (and these errors may be quite large). Further, calibration is often performed separately on the sub-modules of the system, resulting in inconsistencies with the observed data when the sub-modules are linked. This makes it difficult to quantify the uncertainty in the predictions of the entire system performance.

The Bayesian approach provides the opportunity to directly address the sources of uncertainty (input, output, and model error) in the model calibration process. While our previous work has applied Bayesian techniques to hydrologic models (Kavetski et al. 2002, 2006; Renard et al., 2010), this study seeks to develop a Bayesian multi-response method for use with a river system model. This involves joint calibration to all sources of information, a process which entails fitting the parameters of the model simultaneously to all data. Unlike the traditional approach, joint calibration forces consistency in performance across the entire system. Moreover, joint calibration offers the well-established benefit of using multiple response data to better identify model parameters and structure (eg. Sorooshian and Gupta, 1983; Kuczera and Mroczkowski 1998). Finally, the Bayesian approach offers the opportunity to allow proper accounting of uncertainty.

In this paper, an initial implementation of the Bayesian joint calibration is demonstrated using a simple river system model named STICKMAN. While there are numerous existing river system models, with IQQM (Hameed and Podger 2001) and Source River (Welsh et al. 2011) being prominent examples, the use of a simplified model provides a simple means to explore issues affecting parameter identifiability and predictive uncertainty in river basin models. Model calibration is performed using multiple responses with a synthetic data study being used to evaluate model performance and improvements in parameter uncertainty associated with the inclusion of additional responses.

2. STICKMAN MODEL

2.1. Model description

The initial investigations reported in this paper are based on a simplified river model. This avoids the heavy setup and computational burden of modelling more complex river networks (e.g. using tools such as IQQM or Source River). The simplified model is known as STICKMAN, with Figure 1 providing a schematic of the model. In its final form, this model will be constructed to contain the majority of the features of more complex models, while being sufficiently efficient in a computational sense to enable systematic assessment of the issues affecting complex river basin models.

Figure 1. Schematic of the STICKMAN model.

The STICKMAN model consists of a reservoir and a number of reaches, several internal nodes, and an extraction and a release node. The reservoir inflow may be assumed to be gauged directly, or may be estimated using a rainfall-runoff model. The release from the reservoir enters the top of first river reach, with an extraction point located at the bottom of this reach — the extraction aims to meet, but not exceed, a known demand (say for cropping purposes). The second reach begins immediately below the extraction point and has a known environmental flow requirement at its outlet.
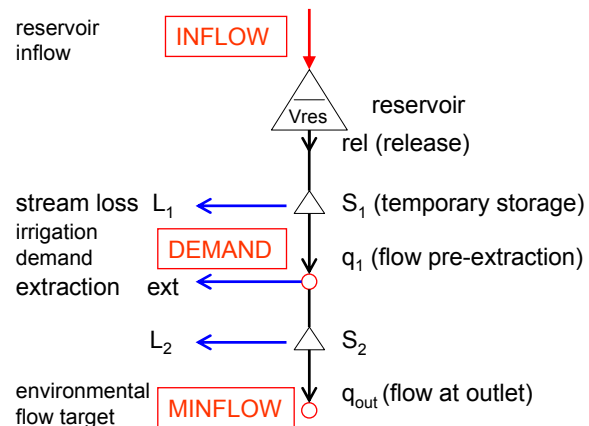
In each river reach, it is assumed that rates of outflow and losses are linear functions of the stream storage, so that the rate of change of storage in the reach is

$$\frac{dS}{dt} = I - Q - L = I - k_Q S - k_L S \tag{1}$$

where S is the stream storage (m^3), I, Q, and L are rates of inflow, outflow, and losses (m^3/s), and k_Q and k_L are storage and loss coefficients (1/s).

Assuming constant reach properties, in particular a rectangular cross section in each reach, the coefficients (or time constants) can be reparameterised. For the outflow:

$$k_Q S = VA = V(S/l)$$
$$k_Q = V/l \tag{2}$$

where V is the reach velocity (m/s), A is the reach cross-sectional area (m^2), and l is the reach length (m).

For the loss:

$$L = k_{SAT} A_{CON} = k_{SAT}(lP) = k_{SAT}l(\beta A) = k_{SAT}l\beta(S/l) = k_{SAT}\beta S \tag{3}$$

where k_{SAT} is the saturated hydraulic conductivity (m/s), A_{CON} is the streambed area in contact with water (m^2), P is the wetted perimeter (m) and β is the inverse hydraulic radius (1/m), taken to be a constant. Thus:

$$k_L = k_{SAT}\beta \tag{4}$$

The reservoir inflow is either obtained from direct measurement (eg. from a stream gauge and rating curve) or is obtained through indirect means (eg. derived from rainfall-runoff model simulation). Regardless of the way it is determined, the observed inflow is subject to errors. In many cases, we expect the observed inflow to be biased. In the simplified model setup, this is represented by scaling the "observed" inflow time series by a constant, which becomes a calibrated parameter in the inference scheme. Thus, the observed inflow is corrected to obtain the "true" inflow:

$$I = k_I I_{OBS} \tag{5}$$

where I is the true (corrected) inflow, I_{OBS} is the observed inflow and k_I is a correction factor.

In summary, the STICKMAN model consists of 5 parameters to be calibrated (V_1, k_{SAT1}, V_2, k_{SAT2}, and k_I). It also requires 4 reach characteristics to be defined (l_1, β_1, l_2, and β_2).

2.2. Model algorithm and implementation

The basic model algorithm over a single time step is:

1. The corrected inflow given by equation (5) is added to the reservoir (updating its current volume, Vres).
2. Choose a release, noting that the reservoir may spill or run dry.
3. Route release through both river reaches.
4. Check that demand and environmental flow constraints are met. If all are met then stop. Otherwise go to step 2 and iterate.

Note that, at each time step, routing the release through the reaches requires solving the differential equation (DE) given by equation (1). This DE for the linear storage in each reach was solved numerically using the implicit (backward) Euler method. Implicit methods are preferred because they remain stable for any time step size Δt, while explicit methods suffer from potential instabilities and other computational weaknesses (Kavetski and Clark 2010). Note that in the current setup the DEs are linear and hence could be solved exactly. However, we favoured more general numerical approximations over specific analytical solutions because in the future we envisage using nonlinear representation of the flows in each reach (eg. making the losses a nonlinear function of storage), which would preclude analytical solutions.

The algorithm was implemented in Fortran 95 and was solved using bisection to determine the optimal reservoir release to satisfy the constraints.

3.   DATA

The forcing data is derived from a 12-year period in the Namoi catchment in NSW. The reservoir inflow is an aggregation of multiple river tributaries (see Figure 2 for inflow time series). The demand time series was computed using the Namoi IQQM model and follows a seasonal pattern to reflect agricultural usage with a cycle of 8 months on and 4 months off, varying at a monthly scale.   The environmental flow requirement is set at a constant value of 0.2 m^3/s.

Synthetic data were generated such that the reservoir experiences periods of spilling and periods of extended drawdown (but never such that the reservoir runs dry) in a reasonably realistic manner (see Figure 2). The internal model states were recorded at 5 model "locations": outflow *qout*, volume of reservoir *vres*, release *rel*, extraction *ext*, and intermediate flow before extraction *q1*. These model states were selected because they should be measurable or estimable in a typical river system.
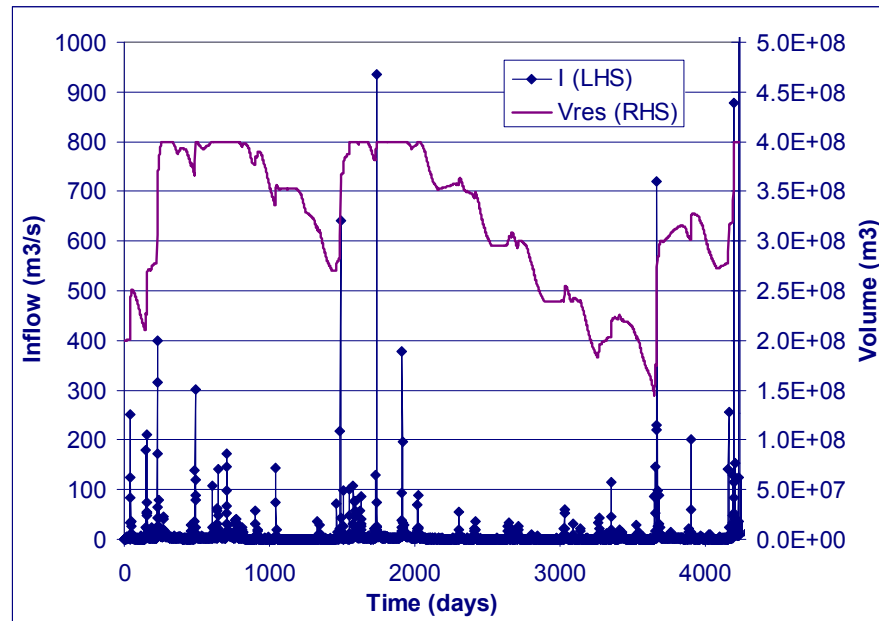

Figure 2. Time series of synthetic data for reservoir inflow (*I*; left-hand axis) and volume (*vres*; right-hand axis). Note that reservoir capacity is 400×10^6 m^3.

3.1.   Response data errors

The synthetic data were corrupted with heteroscedastic noise, where the noise (ie. residual) was of the form:

$$\sigma_\varepsilon = a + bX; \quad X' \leftarrow \mathrm{N}(X, \sigma_\varepsilon^2) \tag{6}$$

where σ_ε is the standard deviation of the noise, $a > 0$ is a fixed offset (to prevent zero-sized errors for zero-valued observations), b is the scaling factor applied to the observation X, and X' is the corrupted observation. The true model parameters used for the synthetic study are given in Table 1.

4.   CALIBRATION METHODOLOGY

Model calibration is based on Bayesian multi-response inference. The advantage of Bayesian methods is that they provide a formal framework for describing the error processes affecting the various components of the model and observation systems and incorporating them into the calibration process:

- Input errors: The input observation system is imperfect, so there will be uncertainty in the measured model input. For example, there may often only be a single raingauge within a catchment, so there will be uncertainty in the areal rainfall values over the entire catchment.
- Output errors: The output observation system is imperfect, so there will be uncertainty in the true model output. For example, streamflow is often estimated using a rating curve and a flood may exceed the levels over which the rating curve was derived, resulting in uncertainty in streamflow values.
- Model errors: The river system model is necessarily a simplification of a very complex system, so there will be considerable uncertainty due to simplifications and idealisations made within the model.

The primary objective of model calibration is to identify the model parameters given the observed time series, the observed forcing time series, and any prior information. In the Bayesian context, this inference problem is described by a posterior pdf. Since the posterior may be difficult to characterize analytically, it is often approximated using Markov chain Monte Carlo methods (eg. Gelman et al. 1995).

In this study the Bayesian multi-response formulation was expressed as follows:

$$p(\theta \,|\, \tilde{Q}, \tilde{R}, M) \propto p(\tilde{Q} \,|\, \tilde{R}, \theta, M) p(\theta \,|\, M) \tag{7}$$

where $\tilde{Q}$ is a matrix of observed response times series (corresponding to some combination of outflow *qout*, volume of reservoir *vres*, release *rel*, extraction *ext*, and intermediate flow before extraction *q1* in Figure 1), $\tilde{R}$ is a matrix of observed forcing (corresponding to I_{OBS} in Figure 1), θ is a vector of parameters of model M to be inferred (corresponding to the STICKMAN model parameters V_1, k_{SAT1}, V_2, k_{SAT2}, and k_I and the response error model parameters *a* and *b*), $p(\theta \,|\, M)$ is the prior pdf assigned to θ describing what is known about the parameters before calibration, $p(\tilde{Q} \,|\, \tilde{R}, \theta, M)$ is the likelihood function which describe the probability model used to generate the observed responses, $p(\theta \,|\, \tilde{Q}, \tilde{R}, M)$ is the posterior pdf of the parameters conditioned on the observed data, the model and any available prior knowledge. Non-informative priors were used for all model parameters.

The form of the likelihood function depends on the error models affecting the observed responses. In this synthetic study, the errors corrupting responses were sampled from independent, heteroscedastic Gaussian distributions given by the equation (6). Accordingly, the likelihood function takes the following weighted least squares (WLS) form (Kuczera and Mroczkowski 1998)

$$p(\tilde{Q} \,|\, \tilde{R}, \theta, M) = \prod_{t=1}^{T} \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi(a_k + b_k Q_{kt}(\theta, \tilde{R}, M))^2}} \exp\left[-\frac{1}{2}\left(\frac{\tilde{Q}_{kt} - Q_{kt}(\theta, \tilde{R}, M)}{a_k + b_k Q_{kt}(\theta, \tilde{R}, M)} \right)^2 \right] \tag{8}$$

where the *t* and *k* subscripts refer to time and response indices respectively.

5. CALIBRATION USING SYNTHETIC DATA

The corrupted synthetic data were analysed using the BATEA toolkit software (Kavetski et al. 2006), which implements inference setups ranging from simple WLS schemes to complex hierarchical setups. In this work, the WLS method was used, with the weights for the individual observations derived from the inverse variances predicted using the residual error model (ie, exactly as in equation (8)).

The first 800 days of data were used for calibration, with the remainder being saved for later prediction purposes (not shown). Various combinations of the available responses (single and multiple) were analysed to determine the effect of multi-response calibration on parameter uncertainty. Table 1 provides a summary of the analyses performed for five calibration scenarios.

The table shows that the use of additional responses in the calibrations generally leads to reduced parameter uncertainty (ie. reductions in posterior standard deviation of the parameters). This is expected since the use of additional responses provides additional information to the parameter inference scheme. Moving from 1 to 2 responses (from scenario A to scenario B or C) significantly reduces parameter uncertainty. It was found that adding the intermediate flow response (*q1*) as the second response resulted in a greater reduction in uncertainty than adding the reservoir volume (*vres*). It appears that this arises because the information provided by *q1* (the flow at the outlet of reach 1) is much more informative about the overall "state" of reach 1 compared to that provided by *vres* (the reservoir volume upstream of reach 1). As such, *q1* "sees" the loss in reach 1, while *vres*, being upstream of reach 1, has no information about the reach loss.

Interestingly, scenarios C and D have effectively the same parameter uncertainty, even though C only uses 2 responses while D uses 3 responses. The responses *qout* and *q1* are common to both, while scenario D adds *vres*. In this case, the extra response (*vres*) provides no extra useful information about the system. This illustrates that using extra responses will not always lead to more precise parameter estimates — ideally the extra responses should be assessed *a priori* to ascertain whether they will significantly augment the information content of the calibration data (Kuczera and Mroczkowski, 1998, p. 1488).

Scenario E uses all 5 available responses and leads to a massive reduction in parameter uncertainty, even compared to the 3-response case (scenario D). In this case, most of this improvement is due to the inclusion of the release response (*rel*). The improvement is both a consequence of the relatively small noise applied to this response (see Table 1) and of the simplicity of the model — knowing the reservoir release allows the loss in reach 1 to be determined quite accurately, and in turn for reach 2.

The single-response calibration to *vres* was investigated next. In this case, the calibrated response (*vres*) is upstream of both reaches, so we would not necessarily expect the model parameters to be well identified. However, Figure 3 shows that the parameters for reach 1 are, in fact, quite well identified (although the parameters for reach 2 are highly uncertain). This result suggests that knowledge of *vres* provides indirect knowledge of *rel* (and hence the flows entering reach 1). This illustrates that, even with limited data, it may be able to reasonably identify some (but not all) parameters of the model.

Table 1. Summary of posterior mean and standard deviation of parameters for various calibration scenarios.

STICK-MAN parameter	True value of transformed parameters	Calibration scenario				
		A	B	C	D	E
		1 response	2 responses	2 responses	3 responses	5 responses
		qout	qout+vres	qout+q1	qout+vres+q1	qout+vres+rel+ext+q1
$\ln V_1$ (m/s)	-1.204	-1.194 (0.260)	-1.426 (0.140)	-1.101 (0.096)	-1.137 (0.105)	-1.205 (0.004)
$\ln k_{SAT1}$ (m/s)	-12.206	-12.801 (0.440)	-12.469 (0.151)	-12.120 (0.106)	-12.126 (0.114)	-12.209 (0.005)
$\ln V_2$ (m/s)	-0.916	-0.866 (0.217)	-0.068 (0.165)	-0.927 (0.099)	-0.898 (0.100)	-0.906 (0.016)
$\ln k_{SAT2}$ (m/s)	-12.206	-11.899 (0.243)	-11.987 (0.204)	-12.224 (0.138)	-12.232 (0.146)	-12.190 (0.020)
$\ln k_I$ (-)	0.0	-0.017 (0.009)	-0.002 (0.003)	0.000 (0.003)	0.001 (0.002)	0.000 (0.000)
Error model						
$\ln a$(qout)	0.0	-0.013 (0.033)	-0.012 (0.033)	-0.012 (0.032)	-0.013 (0.032)	-0.014 (0.032)
$\ln b$(qout)	-1.204	-1.353 (0.096)	-1.357 (0.093)	-1.349 (0.094)	-1.347 (0.092)	-1.349 (0.085)
$\ln a$(vres)	12.206		10.605 (0.929)		10.645 (0.932)	10.625 (0.929)
$\ln b$(vres)	-4.605		-4.593 (0.035)		-4.594 (0.034)	-4.594 (0.033)
$\ln a$(rel)	-3.219					-3.200 (0.057)
$\ln b$(rel)	-4.605					-4.643 (0.068)
$\ln a$(ext)	-3.912					-3.867 (0.043)
$\ln b$(ext)	-2.303					-2.354 (0.033)
$\ln a$(q1)	0.0			0.010 (0.059)	0.010 (0.057)	0.013 (0.057)
$\ln b$(q1)	-1.204			-1.224 (0.068)	-1.223 (0.066)	-1.229 (0.067)

5.1. Discussion

Additional work is needed for the STICKMAN model to become a closer approximation of real river system models (while remaining computationally efficient). Future enhancements include using a rainfall-runoff model to estimate the reservoir inflows. This coupled model could then be calibrated jointly, by simultaneously estimating the parameters of the rainfall-runoff model and of STICKMAN.

Lateral inflows downstream from the reservoir may also be added to make the model more representative of real-world scenarios. This may be achieved through adding an extra inflow between the two reaches or by coupling existing models together in the appropriate configuration.

6. SUMMARY

This paper demonstrated the use of Bayesian multi-response calibration approaches to a river system model. The approach was applied to the STICKMAN model, a simplified river system model which was introduced in the paper. The model was calibrated using a Weighted Least Squares method in a synthetic data study. Model calibration used both single and multiple response data to investigate the reductions in parameter uncertainty resulting from the inclusion of additional responses into the inference.

The use of multiple response data during model calibration was generally found to reduce parameter uncertainty. However, the reductions in uncertainty were variable, highlighting the fact that some sources of data are more informative than others. This provides further weight to Kuczera and Mroczkowski's (1998) conclusion that the value of new sources of information should be assessed *a priori* before

embarking on a field campaign. This study reports the first findings in this project. Future work will explore the effects of multiple response data on model predictive performance, further develop the STICKMAN model to better represent processes and errors, and finally consider IQQM case studies. The approach is general and independent of the river-system model. We expect it can be applied to river system models with a large number of responses and parameters — in other studies we have calibrated hierarchical models with hundereds of parameters using the BATEA toolkit software.
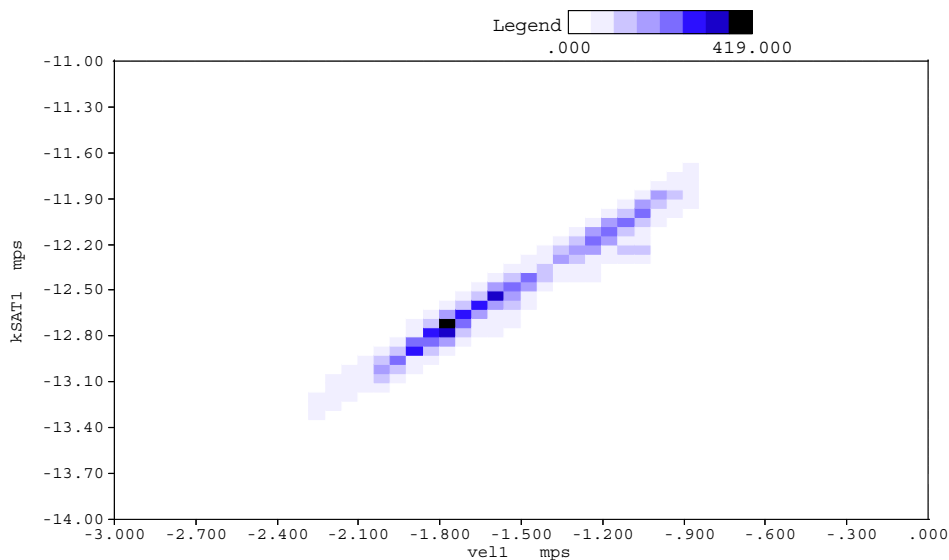

Figure 3. Marginal posterior density plot for reach 1 parameters for single-response calibration to vres. True parameter values are V_1 = -1.2 and k_{SAT1} = -12.2.

REFERENCES

Gelman, A.B., Carlin, J.S., Stern, H.S., and Rubin, D.B. (1995). *Bayesian Data Analysis*, Chapman & Hall, London.

Hameed, T. and Podger, G. (2001). Use of the IQQM simulation model for planning and management of a regulated river system. *IAHS Publication*, 272, 83-89.

Kavetski, D. and Clark, M.P. (2010). Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resources Research*, 46(10), W10511.

Kavetski, D., Franks, S.W., and Kuczera, G. (2002), Confronting input uncertainty in environmental modeling. In: Duan, Q., Gupta, H.V., and Sorooshian, S. (eds), *Calibration of Watershed Models*, AGU Series, vol. 6, 49-68.

Kavetski, D., Kuczera, G., and Franks, S.W. (2006). Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resources Research*, 42(3), W03407.

Kuczera, G. and Mroczkowski, M. (1998). Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. Water Resources Research, 34(6), 1481-1489.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S.W. (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resources. Research*, 46(5), W05521, doi:10.1029/2009WR008328.

Sorooshian, S. and Gupta, V.K. (1983). Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness. *Water Resources Research*, 19(1), 260-283.

Welsh, W.D., Vaze, J., Dutta, D., Rassam, D., Rahman, J.M., Jolly I.D., Wallbrink P., Podger, G.M., Bethune, M. Hardy, M.J., Teng, J. and Lerat, J. (2011) An integrated Modelling Framework for Regulated River Systems. *Environmental Modelling and Software*, (in press).