Modelling precipitation uncertainties in a multi-objective Bayesian ecohydrological setting

Yating Tang\textsuperscript{a}, Lucy Marshall\textsuperscript{a,b*}, Ashish Sharma\textsuperscript{a}, Hoori Ajami\textsuperscript{b}

\textsuperscript{a} UNSW Water Research Centre, School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia
\textsuperscript{b} Department of Environmental Sciences, University of California Riverside, Riverside, CA 92521, USA

A R T I C L E   I N F O

Keywords:
Ecohydrological modeling
Bayesian inference
Input error
Multi-objective calibration
Uncertainty analysis

A B S T R A C T

Recent studies have demonstrated that hydrological model calibrations are impaired by uncertainties in observations and model structures. For a rainfall-driven model, the error in precipitation observations can lead to biased parameter estimates and predictions. For a conceptual ecohydrological model, an appropriate description of input error is essential because rainfall controls both hydrological processes and vegetation growth in the model. However, to date the impact of precipitation uncertainty on ecohydrological model parameters and outputs has not been widely explored (Fuentes et al., 2006; Shields and Tague, 2012). The increased dimensionality of these types of models and the uncertainties associated with calibration data can make traditional approaches for characterizing precipitation errors problematic. Our study aims to investigate the impact of precipitation uncertainty for a Bayesian multi-objective calibration approach in an ecohydrological modeling study. A conceptual model that combines a hydrologic model and a modified bucket grassland model is implemented for a forested catchment in Australia. In the study, different input error descriptions are used in both single and multi-objective Bayesian calibration case studies aimed at simulating streamflow and Leaf Area Index (LAI). The emphasis on each objective is represented as different prior distributions defined for error parameters for multi-objective cases. Results show better parameter estimates and predictions for the cases including input error. Comparing the results from the cases in which different input error descriptions are used, a simple bias term works well for both streamflow and LAI estimations. Although a more complex rainfall multiplier approach to represent input error performs best for streamflow predictions, increasing the dimensionality of the input error model is not always justified given the information content of the data. In addition, some of the rainfall multipliers values are not meaningful in the real case, reflecting an overfitting problem.

1. Introduction

One of the major sources of uncertainty in hydrological modeling is input error from precipitation observations. Several recent studies have demonstrated that model simulations and parameter estimates can be largely impaired by uncertainty in the input rainfall observations (Huard and Mailhot, 2006; Bardossy and Das, 2008). Gauge-based rainfall errors may arise from the problem that single point gauges cannot adequately measure the rainfall amount across an entire catchment, or due to inaccurate spatial interpolation between rain gauge networks. Other factors include the effect of winds, evaporation and other local issues (McMillan et al., 2011). Satellite-based precipitation error is more significant and complex. Research has found that rainfall measurements from single point gauges are more likely to be underestimated and actual amount of precipitation is usually higher than gauge measurements. However, the satellite-based precipitation estimates do not contain this type of systematic error and the precipitation errors are generally not Gaussian (Gebremichael and Krajewski, 2005).

A variety of approaches have been developed to characterize input error in hydrologic model inference. With the development of Markov chain Monte Carlo (MCMC) methods and advanced computing power, Bayesian inference has become a popular tool for uncertainty analysis in hydrological modeling (e.g. Bates and Campbell, 2001; Kuczera, 1983; Marshall et al., 2004; Smith and Marshall, 2008; Vrugt et al., 2003; Tang et al., 2016; Marshall et al., 2006; Jeremiah et al., 2011), and extensions to traditional Bayesian inference allow for explicit representation of input errors. Bayesian inference combines prior information and observations into probability distributions on the model parameters, which allows the modeler to statistically quantify model uncertainties. One of the popular approaches to incorporate precipitation uncertainty using the Bayesian methods is to consider input error as multipliers on the input rainfall time series (Kavetski et al., 2006; Vrugt et al., 2008b). The multipliers could be dependent on the magnitude of each storm
event and inferred as latent variables (variables that are not directly measured but are inferred from the model) using a Bayesian hierarchical approach. Such a method is the basis of the popular Bayesian Total Error Analysis (BATEA) methodology (Kuczera et al., 2006; Thyer et al., 2009; McMillan et al., 2011; Vrugt et al., 2008b). While this method is widely used in rainfall-driven hydrological modeling, it has drawbacks in that it may not provide a realistic description of rainfall uncertainty especially when input errors are strongly dynamic (Del Giudice et al., 2016). Another disadvantage of the multiplier method is that the high dimensionality increases the calibration difficulty and computational expense. Furthermore, limited prior information on input errors may lead to an ill-posed problem (Sun and Bertrand-Krajewski, 2013). While the use of storm dependent rainfall multipliers addresses the need to capture the dynamic nature of input uncertainty, it comes at the cost of added computational burden with increased estimation uncertainty given the higher dimensionality of the problem.

Recently, studies have been conducted on developing different error models that may represent input error more realistically. Reichert and Mieletin (2009) developed an approach for identifying time varying stochastic model parameters to reduce uncertainties for dynamic models. Del Giudice et al. (2016) introduced a Stochastic Input Process (SIP) method in which ‘rainfall potential’ is used. This method does not require the true rainfall to be proportional to the observations, yet can be inferred from prior and observations, which makes it possible to deal with time-varying errors in observations. However, both of these methods are computationally expensive and challenges remain in finding proper input error terms for different catchments and models.

In recent years, the important role of vegetation dynamics in the water cycle of catchments has gained increasing attention (Porporato et al., 2002; Asbjørnsen et al., 2011). The complex interrelationship between vegetation dynamics and water flows can be investigated via ecohydrological models that integrate hydrological and ecological processes in the catchment. Ecohydrological models expand on traditional hydrologic model formulations and usually incorporate mathematical descriptions of vegetation dynamics such as photosynthesis and respiration processes (Aroza, 2002; Montaldo et al., 2005). Some of the more commonly used ecohydrological models include SWAT (Arnold et al., 1993; Arnold et al., 1998) which is widely used for water quantity analysis as well as water quality assessment (Aouissi et al., 2014; Bieger et al., 2014; Jeong et al., 2014), and climate change impact assessment in catchments with tightly coupled water-vegetation relationship (Ficklin et al., 2009; Eckhardt and Ulbrich, 2003); SWIM (Krysanova et al., 1998; Krysanova et al., 2000) which is widely used for simulating nitrogen dynamics, crop yield and erosion rates in mesoscale basins (Krysanova et al., 2005); RHESSys (Band et al., 1993; Tague and Band, 2004) for simulating water, carbon and nitrogen dynamics in forested and mountainous catchments, and Tethys-Chloris (Faticchi et al., 2012) which is the first ecohydrological model to include the effects of snow dynamics and can be used for simulating ecohydrological processes in both water and energy limited catchments.

Ecohydrological models are generally more complex than hydrological models as more physical processes are coupled to describe the interactions between hydrologic and ecosystemic processes. This integration of hydrologic and ecologic processes increases model dimensionality and multiple input data are used for model conditioning. In this case, the analysis of input uncertainties becomes increasingly important as it impacts accuracy of multiple outputs (i.e. stream flow and leaf area index) along with model state variables and ecohydrologic parameters. Most recently, the Bayesian multi-objective approaches have been introduced as an alternative formal framework to traditional Pareto-based multi-objective optimization methods for ecohydrologic models (Tang et al., 2018) to characterize ecohydrologic parameter uncertainties. However, to date the impact of precipitation uncertainty on ecohydrologic model parameters and outputs has not been widely explored (Fuentes et al., 2006; Shields and Tague, 2012). Rainfall is one of the main drivers that controls vegetation growth, therefore it is essential to investigate the impact of precipitation on ecohydrological modeling and how parameter estimations for both hydrological and ecological model components would be affected by input error from rainfall observations. In addition, it is not clear how ecohydrologic models might benefit from a more explicit treatment of input errors by using methods such as the BATEA approach due to the model complexity, while multiple types of observations might help to inform input error parameters.

In our study, we implement a multi-objective Bayesian calibration for a conceptual ecohydrological model considering different cases to characterize input error. We aim to investigate the impact of precipitation uncertainty as rainfall is the main driving input for both runoff and biomass production. We further compare the results using different types of input error descriptions, from the simplest single bias term to rainfall multipliers, for both single and multi-objective calibrations. The overall goal is to investigate the appropriate representation of input error to improve parameter estimations for multiple objectives (stream flow and leaf area index) while minimizing the computational cost.

2. Methodology

We assess the impacts of rainfall error on ecohydrologic response by calibrating a conceptual vegetation-hydrology model in a forested catchment in Australia. Using a Bayesian approach, we compare simulations for both single and multi-objective calibrations with different input error descriptions, focusing on differences in derived posterior parameter estimates and streamflow simulations. Our overall approach is detailed below.

2.1. Model description

A conceptual ecohydrological model which combines a hydrologic model (HYMOD) (Boyle, 2001; Wagener et al., 2001) and a modified bucket grassland model (BGM) (Istanbulluoglu et al., 2012) is used in our study (Fig. 1). HYMOD is a lumped conceptual hydrological model which has a nonlinear soil moisture tank connected with quick flow tanks (to route surface runoff) and a slow flow tank (to represent subsurface runoff) in parallel. BGM is a vertically lumped bucket model which simulates biomass dynamics, and is coupled to the soil moisture tank in HYMOD. Model inputs are daily rainfall (R) and potential evapotranspiration (PET), and major model outputs are streamflow (Q) and leaf area index (LAI). Two-way feedbacks between hydrologic and eco-logic components are explicitly incorporated. Seven model parameters are calibrated in the simulations. Detailed model descriptions, model parameters and equations can be found in (Tang et al., 2018).

2.2. Catchment and data

The Bloomfield River catchment at China Camp (108003A), located in Daintree Basin, Queensland, Australia is selected in this study (Fig. 2). The catchment area is 263km² with average annual rainfall of 2170 mm and average annual runoff of 1551 mm. 98% of its area is covered by trees and the maximum LAI value is 5.8 (leaf area m²/ground area m²). Daily gridded Precipitation and PET data are from the Australian Water Availability Project (AWAP) with spatial resolution of (0.05°×0.05°) generated from gauge-based daily rainfall measurements (Beeley et al., 2009). Daily stream flow data from the Australian Bureau of Meteorology Hydrologic Reference Stations (HRS) and 8-day LAI data from the Moderate Resolution Imaging Spectroradiometer (MODIS) are used for model calibration. Data from 2001 to 2005 are used for model calibrations.

The quality of MODIS LAI data is assessed by checking the MODIS landcover data over the study period. The land cover classification based on the International Geosphere-Biosphere Programme (IGBP) classification is checked and found to be consistent with catchment vegetation types. The MODIS LAI quality flags are also checked for each pixel to assess the quality of LAI data over the study catchment. The LAI
Fig. 1. Model description. A conceptual rainfall-runoff model (HYMOD) and a dynamic vegetation model (DVM) are coupled. The precipitation and potential evapotranspiration are adjusted by interception and vegetation dynamics represented by changes in LAI in the model respectively. These forcings further influence the soil moisture tank in HYMOD and streamflow. The leaf area index (LAI) in the vegetation dynamics model is calculated based on total biomass at every time step (daily). The rate of biomass production depends on net primary productivity (NPP) which is computed based on simulated ET and water use efficiency parameter, and biomass allocation and decay coefficients. Seven parameters (Huz: soil moisture tank; B: Distribution function shape; Alp: Quick-slow split; Kq: Quick-flow routing rate; Ks: Slowflow routing rate; WUE: Water use efficiency; and Ksg: Natural decay factor for live/green biomass) are calibrated. Further model information can be found in (Tang et al., 2018).

Fig. 2. Catchment location of Bloomfield River at China Camp (108003A), located in Daintree Basin, Queensland, Australia.

data are then filtered at every pixel (1 km resolution) using the Svitzky-Golay method in the TIMESAT software (Jonsson and Eklundh, 2002; Jonsson and Eklundh, 2004). Catchment averaged MODIS LAI values are calculated and used for model calibration where simulated green LAI is compared with MODIS LAI.

2.3. Multi-objective Bayesian approach

In the multi-objective Bayesian approach, we calibrate on both streamflow (Q) and leaf area index (LAI) (Tang et al., 2018). Assuming the error terms of each objective are independent, the log multi-objective likelihood function can be expressed as:

$$L_{multi} = L(Q|\theta) + L(LAI|\theta)$$

where \(\theta\) represents model parameters. The homoscedasticity/ heteroscedasticity of the model errors were first assessed by using the approach described by Smith et al. (2015). In this approach, residual scatterplot and quantile-quantile plot were used to check the likelihood function for each of the objectives. From the results, the error term for Q is assumed to be Gaussian, heteroscedastic and independent. The log likelihood is defined by applying a Box Cox transformation
(Box and Cox, 1964):

\[
L(Q_i|\theta) = -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma_R^2) - \frac{1}{2\sigma_R^2} \sum_{i=1}^{N} [Q_i - f_R(x_i; \theta)]^2 \frac{1}{2\sigma_R^2} \sum_{i=1}^{N} [Q_i - f_R(x_i; \theta)]^2
\]

(2)

where \( \lambda \) is the transformation parameter fixed as 0.3 based on the work of Thyer et al. (2002) for catchments in Queensland, Australia. The error term for LAI is Gaussian, homoscedastic and independent:

\[
L(LAI_i|\theta) = -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma_{LAI}^2) - \frac{1}{2\sigma_{LAI}^2} \sum_{i=1}^{N} [LAI_i - f_{LAI}(x_i; \theta)]^2
\]

(3)

So the log multi-objective likelihood becomes:

\[
L_{multi} = -N \log(2\pi) - \frac{N}{2} \log(\sigma_R^2) - \frac{N}{2} \log(\sigma_{LAI}^2)
\]

\[
- \frac{1}{2\sigma_R^2} \sum_{i=1}^{N} [Q_i - f_R(x_i; \theta)]^2
\]

\[
- \frac{1}{2\sigma_{LAI}^2} \sum_{i=1}^{N} [LAI_i - f_{LAI}(x_i; \theta)]^2 - \log(Q_i + \lambda)
\]

(4)

The prior distributions for model parameters are defined as bounded uniform distributions. The error parameters \( \sigma_R^2 \) and \( \sigma_{LAI}^2 \) in the likelihood function reflect the ‘weight’ of each objective (Reichert and Schuwirth, 2012; Minet et al., 2015). Therefore, the prior distributions for these variances should be defined carefully based on modeler’s confidence level in the ecohydrologic observations. Tang et al. (2018) presented a detailed discussion about defining appropriate prior distributions for error parameters in multi-objective Bayesian calibrations. Following Tang et al. (2018) and based on the level of confidence about the observations, the prior distributions for the output error parameters are defined as:

\[\sigma_R^2 \sim T(N(0,0.2),\sigma^2) \quad \sigma_{LAI}^2 \sim T(N(0,0.8))\]

(5)

where \( T \) refers to a normal distribution truncated at 0. The prior distributions are truncated at 0 to avoid negative values. The mean of the prior is set to 0 which indicates that the best optimized objectives (the lowest errors) are preferable. The specified prior variances reflect different emphases that put on each of the objectives, where a smaller variance means more emphasis on a particular objective. This approach follows the work by Reichert and Schuwirth (2012). The prior distributions of error parameters are maintained the same across all the multi-objective cases.

2.4. Input error descriptions

We consider several input error models of increasing dimensionality to evaluate the efficacy of the models and their efficiency in calibration of multivariate model outputs. We compare a traditional model calibration, assuming no input errors, to several input (rainfall) error cases described as: bias, clustering biases and rainfall multipliers. In these approaches, the rainfall errors are all defined based on the multiplicative factor(s), as multiplicative rainfall error approaches are well-supported by research characterizing precipitation error (Tian et al., 2013), although the multiplicative error means that rainfall depths of zero are not corrected (Vrugt et al., 2008a). These approaches are commonly used in rainfall data driven hydrologic model calibrations (Kavetski et al., 2006; Ajami et al., 2007). We include these approaches in the Bayesian multi-objective calibrations of the ecohydrologic model to investigate the impact of input error on stream flow and leaf area index simulations. A brief discussion of the different input error characterization on ecohydrologic model simulations is given below.

2.4.1. Simulations with no input error

In this case, rainfall observations are assumed to be the same as the true rainfall values. As a result, input error is ignored and the total error of the model is lumped into a single residual error term for each objective: \( \epsilon_Q \) and \( \epsilon_{LAI} \), which are assumed to be independent and Gaussian with constant variance for LAI (\( \epsilon_{LAI} \)) and heteroscedastic variance for stream flow (\( \epsilon_Q \)). This approach is also known as the standard least squares (SLS) approach, and has been criticized for its simplicity and lack of realism for hydrologic case studies (Schoups and Vrugt, 2010). While it has been reported that lumping of input error effects into a single residual term can lead to biased parameters, we still perform this scenario as a benchmark to compare with cases that include an input error model for rainfall.

2.4.2. Rainfall input with constant multiplying bias

The simplest model of rainfall errors assumes a single multiplicative term for a calibration data set in order to describe constant bias in the rainfall observations from the true rainfall:

\[ R_{true} = R_{obs} \gamma \]

(6)

Where \( R_{true} \) denotes true rainfall, \( R_{obs} \) is the observed rainfall, and \( \gamma \) is the bias. In our study, we assume the bias parameter has a normal prior distribution truncated at zero:

\[ \gamma \sim T(N(\mu,\sigma^2)) \]

(7)

Where the mean of the prior is 1 (\( \mu = 1 \)) and the variance of prior is 0.1 (\( \sigma^2 = 0.1 \)) as defined in this work. The bias term can be considered as an additional model parameter that corrects the rainfall observations. Such an approach assumes the input error does not vary in time (or cannot be modeled as dynamic). Ajami et al. (2007) presented an approach in which a random multiplier at each time step is drawn from the same distribution with unknown mean and variance. The two unknown variables are treated as additional parameters in the model. Renard et al. (2009) commented that such an approach is able to overcome the problem of unknown true rainfall values and high dimensionality. However, the prior distribution used in Ajami et al., (2007) led to a relatively low rainfall uncertainty which may not be realistic. Later, Ajami et al., (2009) considered this limitation and relaxed the mean and the variance of the prior. Nevertheless, the constant multiplier approach is computationally efficient and based on the preceding studies, our bias approach assumes that the prior distribution of the bias term is known with a variance which is relatively relaxed, but the multiplier is treated as a model parameter.

2.4.3. Rainfall input with clustering error

While a simple bias term is easily specified, it may not appropriately represent the true rainfall error structure. This is the case when the errors vary over time as a function of the storm magnitude, as different uncertainties can be expected for different magnitudes. In the second proposed error model, rainfall observations are assumed to have different bias terms depending on the magnitude of rainfall observations. In this case, the observations are firstly classified using a hierarchical clustering method (Johnson, 1967) to identify natural groupings of the observed daily rainfall time series. The Euclidean distances are used to separate the clusters of daily observations and based on the distances, 5 clusters are defined in this case study. For each cluster, one single bias error value is defined:

\[ R_{true} = R_{obs} \cdot y_i \]

(8)

where \( y_i \) is the clustering bias term for cluster \( i \) with prior distribution:

\[ y_i \sim T(N(\mu, \sigma^2)) \]

(9)

Similar to the constant bias approach, \( \mu = 1 \) and \( \sigma^2 = 0.1 \). The clustering approach increases the dimensionality of the inference scheme while retaining a simple representation of time varying errors.
2.4.4. Rainfall multipliers

The rainfall multiplier approach assumes the input error to be multiplicative and independent (Kavetski et al., 2006; McMillan et al., 2011). In this method, the hyperparameter (the prior distribution parameter) and the hyperprior (the prior distribution for a hyperparameter) need to be defined (Bernardo and Smith, 2001). In the Bayesian Total Error Analysis (BATEA) approach (Kavetski et al., 2006), the input error is defined as:

\[ R_{\text{true}} = f(\varphi, R_{\text{obs}}) \]

\[ \log \varphi_{\text{true}} \sim N(\mu, \sigma^2) \] (10)

where \( \varphi \) is a vector of storm dependent rainfall multipliers. In this approach, \( \varphi \) is treated as a latent variable following a lognormal distribution with hyperparameters \( (\mu, \sigma^2) \) that are inferred together with the model parameters. Therefore, an additional likelihood function describing the input error model is included in the inference:

\[ L(\mu, \sigma^2|\varphi) = \prod_{i=1}^{n} \left( 2\pi \sigma^2 \right)^{-1/2} \varphi_i^{-1} \exp \left( -\frac{(\log(\varphi_i) - \mu)^2}{2\sigma^2} \right) \] (11)

Then the total log likelihood function for a multi-objective calibration case can be written as:

\[ L_{\text{multi}} = -N \log(2\pi) - \frac{N}{2} \log(\sigma_{Q}^2) - \frac{N}{2} \log(\sigma_{LAI}^2) \]

\[ -\frac{1}{2\sigma_{Q}^2} \sum_{i=1}^{N} [Q_i - f(Q_i; \lambda)]^2 - \log(Q_i + \lambda) \]

\[ -\frac{1}{2\sigma_{LAI}^2} \sum_{i=1}^{N} [LAI_i - f(LAI_i; x_i; \theta)]^2 - \frac{n}{2} \log(2\pi \sigma^2) \]

\[ -\sum_{i=1}^{n} \log(\varphi_i) - \frac{\sum_{i=1}^{n} (\log(\varphi_i) - \mu)^2}{2\sigma^2} \] (12)

In applying the rainfall multiplier approach, the number of inferred rainfall multipliers \( n \) must be determined. The desire is to ensure the rainfall errors appropriately represent the true rainfall uncertainty, while recognizing that with a different rainfall multiplier applied to each day the calibration becomes high-dimensional. Kavetski et al. (2006) addressed this by having rainfall multipliers vary between storms, where a storm is defined as continuous daily rainfall without break. In our work, we include two rainfall multipliers cases with 73 and 122 rainfall multipliers for the storms in the calibration year respectively. These varying numbers of multipliers are counted by defining the storm events with rainfall exceeding 10 mm and 0.5 mm respectively. The value of the hyperparameter \( \mu \) is set to be zero across all the cases, assuming the mean of the multipliers is 1. Kavetski et al. (2006) mentions that when the hyperprior distribution for hyperparameter \( \sigma^2 \to \infty \), the calibration will lead to an ill posed problem. However, when the prior is too constrained (i.e. \( \sigma^2 \to 0 \)), BATEA will converge towards a SLS approach.

In our study, the hyperprior distribution for \( \sigma^2 \) is defined as a uniform distribution varying from 0 to 0.1 (standard deviation around 0.32) to reflect our belief of the level that true rainfall values are impared by the error. The hyperprior distribution remains the same for all the case studies:

\[ \sigma^2 \sim \text{unif}(0, 0.1) \] (13)

2.5. Setup

A Bayesian calibration framework with the Adaptive Metropolis (AM) algorithm (Haario et al., 2001) is set up for each case representing different formulations of input error for the rainfall. For each simulation, the first 2 years of data were used as a warm-up or spin-up period to minimize the effect of initial condition assumptions regarding the initial LAI and catchment storage, and the third year was used for calibrations. For each simulation, 100,000 iterations are carried out. The first 20,000 iterations are discarded and the remaining 80,000 iterations are used to analyze the posterior distributions. Convergence is diagnosed via visualizing diagnostic plots of multiple MCMC runs to reduce computational demand (Marshall et al., 2004). The calibration runs are performed for single objective (Q or LAI) and multiple objectives (Q and LAI) formulations of the Bayesian approach. We compare the results from cases with no rainfall error, constant bias, cluster and multiplier descriptions of rainfall error in our work.

3. Results

3.1. Single objective calibrations

Single objective calibrations on streamflow (Q) and leaf area index (LAI) are firstly developed. Results of 90% confidence limits from cases with different input error descriptions are compared in Table 1A and 1B. Reliability (the percentage of observations captured) and sharpness (the mean width) from 90% confidence limits for each case are compared (Smith et al., 2015). The Mean square error (MSE) from optimized predictions based on log transformed data are calculated and shown in Table 1A and 1B. The case studies incorporating input errors (bias, cluster and multiplier cases) show improved model sharpness and reduced MSE while maintaining similar reliability to the case with no input error. The 90% confidence band (sharpness) in the no error case is 1.40 log(mm/day) for log-transformed streamflow and 3.04 for LAI, while results from the bias case are 1.23 log(mm/day) and 2.88 respectively. MSE results are largely improved for both streamflow and LAI, from 0.18 log(mm/day) in the no error case to 0.14 log(mm/day) in the bias case, and from 0.85 in the no error case to 0.77 in the bias case respectively. Calibrations with rainfall multipliers work best for the streamflow-only cases. However, for the LAI-only calibration, results from the bias and multipliers are similar. There is no significant improvement comparing the multiplier cases with the bias case.

Next, to evaluate the model complexity and diagnose overfitting problems, the Bayesian Information Criteria (BIC) (Schwarz, 1978) is calculated for each of the single calibration cases and shown in the last column in Table 1A and Table 1B. The BIC is an asymptotic approximation to the likelihood which is widely used for model selection in Bayesian studies (Marshall et al., 2005). The BIC is calculated as:

\[ \text{BIC} = \log(n) k - 2 \log(\hat{L}) \] (14)

Where \( n \) is the number of observations, \( k \) is the number of parameters and \( \hat{L} \) is the maximized likelihood function (from Eq. (2) for Q (3) for LAI). Smaller BIC values suggest a better model. For both streamflow-only and LAI-only cases, the BIC results were the best for the bias cases (2885 and 453 respectively), while significantly deteriorated for multipler cases, especially for LAI.

3.2. Multi-objective calibrations

Table 2 shows the statistical results of 90% confidence limits for multi-objective calibrations of the ecohydrologic model from the cases with different input error descriptions. Predictions for both streamflow and LAI are largely improved when rainfall input error is included. Overall, results from the 73 rainfall multipliers case have the best statistics for both streamflow and LAI. Predictions from input error with constant bias are better than the cluster case for streamflow, but worse for LAI. Predicted LAI from the 122 multiplier cases is worse than the other input error approaches, although the predicted streamflow is better.

The posterior distributions for model residual error parameters for (a) streamflow and (b) LAI in the multi-objective calibration cases are compared in Fig. 3. The mean residual for streamflow is about 0.17 and for LAI about 2.50 from the case assuming no input error (black lines). Results from the bias (blue line) and the cluster (light blue line) cases are similar for streamflow. The residual errors reduced to around 0.14 for streamflow and 1.1 for LAI for the bias case. Results from the two multiplier cases (green and red lines) are similar and the model
the calibration residuals for the bias case has a mean of about 1.38. The mean of the posterior distributions for the cluster input errors vary from the smallest value of 0.8 (green line, c5 case) to about 1.4 (light blue line, c3 case). The bottom panel of Fig. 4 shows the best estimations of all multiplier values for each storm. It can be observed that although most of the multipliers are located around 0.8 and 1.5 for the 73 multipliers case and 1.0 for 122 multipliers case, there are many extreme values (i.e. multipliers greater than 2 or more) for both multiplier cases.

The optimized input errors parameters for the streamflow only cases are shown in Fig. 5. Comparing the results from Fig. 5 with the results from Fig. 4, it can be seen that the input error errors are less informed in the single objective cases than the multi-objective cases. For example, the input error posterior distribution for the bias case (a) is more diffuse with a lower optimized value (about 1.35), comparing with the results of bias case for the multi-objective calibration (Fig. 4(a)). The differences between the posterior distributions for the cluster case (b) between single and multi-objective cases are less significant, but it is easy to see that the variations among the cluster values are smaller for the single-objective case than the multi-objective case. The optimized rainfall multipliers are also very different from the multi-objective cases (Fig. 5, bottom panel).

Next, 90% confidence limits for 150 days of streamflow for the no input error and 73 multipliers cases for both single and multi-objective scenarios are compared in Fig. 6. In each plot, the red dots are log transformed observations while the black line is the corresponding prediction. Results from the no input error (left panel) for (a) single and (c) multi-objective cases are very similar with larger sharpness, compared to the 73 multipliers cases (right panel). Nevertheless, it is evident that predicted streamflow is better from (d) the 73 multipliers multi-objective case with narrower sharpness (1.08) than (b) the 73 multipliers single objective case with sharpness of 1.24.

Lastly, the posterior distributions of the model hydrologic parameters Huz (height of soil moisture tank), B (soil moisture distribution function shape) and vegetation parameters Ksg (natural decay factor for live biomass), WUE (water use efficiency) from different cases are compared in Fig. 7. In each plot, the red cross is the optimized parameter value. It can be seen that the posterior distributions of Huz and B (top two panels) are considerably different for each case. The posterior distributions for Huz moved towards the upper limit boundary with a much narrower shape, for the multipliers cases compared to the no input error case. However, the optimized values from each case are similar. The optimized parameter B value dramatically changed from about 1.8 in the no error case to about 0.35 in the bias case, 1.3 in the cluster case, and dropped to about 0.5 in the multiplier cases. For Ksg and WUE (bot-
tom two panels), the optimized values for cases including input error (b-e) are similar, except that the optimized Ksg values from the cluster and 73 multiplier cases are different from other cases. However, differences among the results can be clearly seen in terms of the parameter uncertainty. The posterior distributions become much more diffuse with lower density in the case assuming no input error in comparison to cases that the input error for rainfall is considered.

4. Discussion

4.1. Incorporating input error in ecohydrologic modeling

The results of our case study show a substantial improvement in streamflow simulations for each of the cases that rainfall input error is considered compared to no input error. The 90% confidence limits for the rainfall bias, cluster and multiplier cases are considerably narrower than the no input error cases for both single and multi-objective
calibrations (Table 1). The mean square error (MSE) largely reduced by including input error as well. Including input error is more important for multi-objective calibrations than the single objective calibrations as residual errors become larger when the trade-offs in objective functions are considered in the multi-objective scenarios (comparing Table 2 with Table 1). As a result, including input error can significantly improve the predictions in the multi-objective cases. For example, the MSE from the bias case for LAI predictions reduced more than half of the result from no input error case (2.50 to 1.18), and the sharpness dropped from 4.88 to 3.43. In addition, larger improvements for LAI predictions than streamflow predictions are observed in multi-objective cases. This means that there is a larger amount of residual error in simulating LAI than streamflow. It can be seen that, with no input error, the MSE for LAI in the multi-objective case is much larger than the single objective case, while the MSE for Q is similar in both cases. We suspect that this outcome indicates that the LAI observations do not have sufficient information to improve model calibration. We used five years of data for calibration due to the computational burden caused by the large number of multipliers that increased the dimensionality of the model. However, our general experience indicates that LAI observations do not reduce uncertainty in model simulations to the same degree as streamflow observations (Tang et al., 2018). To address the problem of poor calibration on LAI, more weight should be placed on LAI and LAI observational error should be investigated in future calibrations.

Due to the absence of the rainfall error information in the validation period, these studies cannot be effectively tested in the validation. This shortcoming is discussed in Thyer et al. (2009). Nevertheless, our results from an ecohydrological model calibration suggest that incorporating input error can improve model predictions via improved calibration.

4.2. Impact of input error model parameters

Theoretically, including input uncertainty should improve model predictions and more accurate parameter estimations should be obtained. In our work, four of the total 7 calibrated model parameters differed considerably among the multi-objective case studies (Fig. 7). The remaining parameters (the quick-slow split parameter and quick/slow tank rates) did not show as significant shifts for both optimized values and shapes of their posterior distributions among the different cases characterizing input uncertainty. This suggests that the parameters shown in Fig. 7 compensated the most for precipitation errors. For parameter Huz, although the distributions become much narrower, the optimized values from each case are similar. For parameter B, the posterior distributions are quite different. As B becomes smaller the soil moisture tank will become more uniform and respond faster to rainfall, which leads to a completely different catchment representation. This would indicate that the multipliers method largely helped to avoid the impact of the input errors. However, this might also be because it was forced to shift in order to fit the specified objectives.

Interestingly, comparing the shifts of the posterior distributions of model parameters among different cases, it indicates that the rainfall input errors biased the hydrologic parameter distributions (especially for parameter B), while increasing the uncertainty of vegetation parameters (diffuse shape of parameter distribution). This result suggests input
error affects the bias of the hydrologic parameters, but not the bias of the vegetation parameters. Instead, the vegetation parameters become more uncertain with increasing model dimensionality. This means that the ecologic components of the model are less impacted by rainfall input errors. However, it also might suggest the ecologic component is too simple or not informative enough for modeling the input error.

4.3. Comparing input error descriptions

One of the critical issues of input error analysis is determining an appropriate expression of input error. There is a need to include an input error model that appropriately represents the known dynamic nature of input errors, while acknowledging that the input error dimensionality must be justified by the information content of the data. An overly complex input error model may lead to model overfitting and a model performance that is poor in extrapolation. To address this, we compare model performance using an information criterion that balances model fit with complexity. From the results of the estimated BIC in Table 1, the BIC value for the bias case is the smallest for both streamflow and LAI. It is therefore suggested that from a model parsimony point of view, bias cases performed better. From the resulting 90% confidence limit and MSE, it seems that the rainfall multipliers method is superior for streamflow prediction. The sharpness is the narrowest and MSE is the smallest. However, there are a few outliers seen from the histograms of estimated rainfall multipliers in Fig. 4. Some multiplier values are greater than 2, which is unlikely to occur given the reliability of the observations. These values are forced to adjust in the input error model to get better estimations of streamflow, and may instead be highlighting a model structural deficiency.

In the cluster case, different error values are estimated for each cluster. This means the errors are defined based on the amount of daily rainfall. As can be seen in the figure, rainfall observations in cluster 3 (smallest rainfall depths) and 5 (largest rainfall depths) have the most significant bias (with a posterior mean of 1.4 and 0.75 respectively). This indicates that the model underestimated the smallest rainfall observations grouped in cluster 3, while overestimated the most significant rainfall observations in cluster 5. However, the posterior distribution for cluster 2 shows that rainfall observations in this cluster have relatively little bias. This could be caused by the number of observations grouped in this cluster is not enough and therefore there is insufficient information for the model to infer the bias parameter for this cluster. On the other hand, it is interesting to see that the simplest bias description works well for both streamflow and LAI predictions. Only one bias parameter was used to describe input error in bias case. The posterior distribution shows that the estimated bias for the entire time series is about 1.38, which highlights the general tendency for rain gauges to underestimate rainfall. The bias method can offer an efficient alternative and appropriately simulate the input uncertainty. However, it cannot summarize the dynamics of errors in the rainfall observation. We have applied our approach to another catchment which is shown in the supplementary document. The bias of the rainfall observations was significantly smaller than the case study shown above (approximately 1.08). Nevertheless, the results from different input error cases showed good agreement in both catchments, suggesting the amount of uncertainty varies across different catchments.

It is suggested in the scientific literature that the rainfall multipliers method might be a better approach than a single bias method as rainfall multipliers are able to capture heteroscedastic error in the observations (Thyer et al., 2009). Our results show that rainfall multipliers method is preferable for modeling streamflow. This requires, however, sufficient information from the observations to inform the parameters. The posterior distributions for model parameters did not converged when rainfall
multiplicators are included for LAI-only calibrations. One reason for this is that it takes a much longer time for vegetation to respond to rainfall than streamflow particularly in forested catchments. This means that the LAI response to the variations in rainfall error is delayed as well. In addition, it is impossible to simulate the model when increasing the dimensionality of model parameters as there is limited information in the LAI observations. Therefore, for improved simulation of vegetation dynamics, a simple bias error term is preferable. Another issue of rainfall multipliers method is that the outliers in the estimated multipliers lead to an overfitting problem (as estimated by the BIC). Overall, our study confirms that the assumptions for the multipliers need to be checked and meaningful and constrained hyperprior distributions defined. While this does not imply that the outcomes from rainfall multipliers method are worse than the bias method, the simplicity and robustness of the calibration process makes the bias method recommendable, particularly when simulating vegetation dynamics.

4.4. The value of multiple observations types for estimating input error parameters

One of our research aims is to investigate whether multiple types of observations can help to inform input error parameters. Alternately, we aim to know whether including LAI information in a multi-objective calibration scenario can help to improve the input error estimation and consequently help improve streamflow predictions, compared to a single objective calibration on streamflow, which is similar to the original BATEA approach. First of all, comparing the input error parameters in the multi-objective case (Fig. 4) with the single objective case (Fig. 5), it is evident that the posterior distribution for the bias case became sharper. More error can be found in $\epsilon_5$ for the cluster case, and the optimized multipliers results are completely different. These results emphasize that LAI observations did impact the input error estimations. In addition, for the 73 multipliers cases, the predictions of streamflow for the multi-objective case are markedly improved over the single objective case (Fig. 6): the confidence limits became narrower and the predictions fitted the observations better. This demonstrates that the ecological component of the model and the additional information from LAI observation can help to improve the streamflow estimations. The benefit of this approach to ecohydrological modeling is thus evident.

5. Conclusion

Different input error descriptors are incorporated in a multi-objective Bayesian ecohydrological modeling framework and compared with the case that does not consider input error. Our approaches can be tested in catchments of different types and sizes. However, this study was aimed at looking at the efficacy of different error models with very high computational demand that require implementing a rather exhaustive MCMC sampling algorithm to ensure proper characterization of the input error parameters. From the results, we summarize the following conclusions and outline future work:

1. Incorporating input error can improve model simulations. The rainfall multipliers method worked well for simulating streamflow in our case study. However, it did not work for simulating LAI because the limited information in LAI observations makes it impossible to update such a high dimensional parameterization. For the considerations of ecohydrological model complexity and associated computational demand, we recommend a simple and practical bias approach to include input error in the simulations. The bias approach emphasizes the importance of model parsimony and is applicable to a more generalized model and calibration approach when the information in the forcing data is limited. However, the bias approach cannot stochastically describe the uncertainty.

2. Our results highlight that development of an appropriate approach for realistic and practical description of model input error particularly for highly dynamic and heterocedastic systems are required. It is important to investigate whether the input error model might actually be accounting for some type of model structural deficit. While we used gridded precipitation products interpolated from the gauge-based precipitation observations, this approach can potentially be used to diagnose satellite-based precipitation uncertainties. In addition, observation errors (such as streamflow observational uncertainties) need to be considered, especially for ecohydrological modeling when multiple observations are used. In particular, the observation errors for MODIS LAI are associated with a large amount of uncertainty as satellite products are known to be less reliable than equivalent ground observations.

Acknowledgement

This research is supported by Australian Research Council (ARC) award FT201002609 to Dr. Marshall and Australian Research Council (ARC) Discovery Award DP170103959. Data used in this study is available by contacting the authors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.advwatres.2018.10.015.

Reference


