Title
Reply to Comment by B. Renard et al. on “An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction”

Journal
Water Resources Research, 45(3)

ISSN
0043-1397

Authors
Ajami, Newsha K
Duan, Qingyun
Sorooshian, Soroosh

Publication Date
2009-03-01

DOI
10.1029/2008wr007215

Reply to Comment by B. Renard et al. on "An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction"

Newsha K. Ajami,[1] Qingyun Duan,[2] and Soroosh Sorooshian[3]

Received 11 June 2008; revised 16 September 2008; accepted 3 November 2008; published 10 March 2009.

Citation: Ajami, N. K., Q. Duan, and S. Sorooshian (2009), Reply to Comment by B. Renard et al. on ''An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction,'' *Water Resour. Res.*, *45*, W03604, doi:10.1029/2008WR007215.

1. Introduction

[1] We would like to thank *Renard et al.* [2009] (hereafter referred as RKK2009) for their commentary on how hydrologic input errors are treated in Integrated Bayesian Uncertainty Estimator (IBUNE) by *Ajami et al.* [2007] (hereafter referred as ADS2007). To our knowledge, two of the coauthors of this commentary were the first to develop a statistically principled approach to explicitly address the effects of rainfall input errors on the identification of a rainfall-runoff model. As a matter of fact, our approach for dealing with rainfall input errors was inspired by the work presented by *Kavetski et al.* [2003], who developed the Bayesian Total Error Analysis (BATEA) to account for uncertainties not only in model parameters and streamflow observations, but also in rainfall inputs. Both BATEA and IBUNE explicitly consider input errors by applying multipliers to the rainfall data values. As pointed out by RKK2009, there are differences in these two approaches. In BATEA, the multiplier for each individual storm is treated as an independent parameter and needs to be estimated along with other model parameters. In IBUNE, the rainfall multipliers are random realizations of a Gaussian distribution, whose mean and variance need to be estimated along with other model parameters. *Kavetski et al.* [2003] demonstrated the appealing theoretical properties of BATEA and tested it on a simple rainfall-runoff model using synthetic data. As pointed out by RKK2009, however, BATEA is computationally expensive because the dimensionality of the model identification problem is equal to $n + N_t$, where n is the number of rainfall-runoff model parameters and N_t is the total number of individual storms. Depending on the length of rainfall data, N_t can go as high as several hundred or more, making practical implementation of BATEA very difficult. In contrast, IBUNE introduces only two additional parameters, the input error mean (m) and variance (σ_m^2); that is, the dimensionality of the IBUNE as presented in ADS2007 is equal to $n + 2$.

[1]Berkeley Water Center, University of California, Berkeley, California, USA.

[2]Atmospheric, Earth, and Energy Department, Lawrence Livermore National Laboratory, Livermore, California, USA.

[3]Department of Civil and Environmental Engineering, University of California, Irvine, California, USA.

0043-1397/09/2008WR007215

[2] In RKK2009, a statistical account is first provided to explain how input uncertainty is treated in a Bayesian framework. Then, an extensive mathematical description of two alternative interpretations of the IBUNE, based on the commentators' understanding of the IBUNE is presented. In IBUNE-A, a new set of random deviates are sampled from a Gaussian distribution, N(0,1), whenever the two hyperparameters related to rainfall errors, m and σ_m^2 are sampled along with other rainfall-runoff model parameters. The Gaussian deviates are then modified by m and σ_m^2 to obtain normally distributed rainfall multipliers. In IBUNE-B, a Markov Chain Monte Carlo (MCMC) sampler is used to sample the full Bayesian posterior at each time step, including both the rainfall-runoff model parameters as well as m and σ_m^2. They proceed to provide the results of their comparison of the two alternative IBUNE algorithms with the BATEA algorithm. RKK2009 concluded that the IBUNE had various drawbacks: IBUNE-A has convergence problems because the likelihood function is a random variable, and IBUNE-B has no efficiency advantage over BATEA because the full Bayesian posterior of the rainfall multipliers needs to be calculated. Additionally, RKK2009 contend that IBUNE is limited because it requires the variance of input errors be constrained to a small range. Unfortunately, most of RKK2009's interpretations of IBUNE reflect neither the original intent of IBUNE formulation nor the actual IBUNE algorithmic implementation. The commentators had obviously spent an inordinate amount of efforts (nearly 20+ pages) focusing on their own interpretations, instead of addressing our original formulation (of which we had shared with them the original MATLAB code). In IBUNE-A, RKK2009 interpreted "multipliers drawn from the same Gaussian distribution" as using a completely new set of Gaussian deviates whenever the likelihood function is evaluated. This implies that the multiplier at each time step is conditioned on a newly sampled random value, in addition to being conditioned on the randomly sampled m and σ_m^2. This naturally results in the value of the likelihood function being a random variable and, consequently, leads to nonconvergence of the algorithm.

[3] In their interpretation B of IBUNE (IBUNE-B), RKK2009 adhere to the original rationale behind BATEA formulation, i.e., IBUNE-B would sample the full Bayesian posterior probability space, just as in BATEA. It is not surprising that IBUNE-B would not achieve the desired efficiency. Meanwhile, IBUNE-B provides a relatively poor

approximation of the target posterior distribution compared to BATEA. The reason is that in their IBUNE-B interpretation, the input errors' mean and variance are held constant over the entire data period, while those parameters are storm-dependent for BATEA. What RKK2009 essentially did was to make IBUNE-B retain the computational inefficiency of BATEA while keeping a simpler input error model structure than BATEA.

[4] In the next sections, we clarify the correct interpretation of IBUNE and answer some of the specific points raised by RKK2009. We also elaborate on the limitations of IBUNE.

2. Correct Implementation of IBUNE and Computational Efficiency of IBUNE

[5] The goal of original ADS2007 IBUNE has been to account for total uncertainty in rainfall-runoff modeling, including rainfall data, model parameters, and model structure, in addition to uncertainty in streamflow data. In treating the rainfall data uncertainty, we aimed to keep the essence of BATEA's treatment of input uncertainty while improving its computational efficiency for practical use. Toward this goal, our original formulation is similar to IBUNE-B formulation with a fixed pool of presampled standard Gaussian deviate vectors as described in section 5.5 of RKK2009. ADS2007 IBUNE treats the multiplier at a given time step as a normally distributed random variable with unknown m and σ_m^2. The presampled multipliers are generated and used in the following manner: (1) Presample M standard Gaussian (N(0,1)) deviate vectors, $(\Omega^{(k)}) = (\omega^{(k)})_t$, $k = 1,\ldots, M$, $t = 1,\ldots, T$, where M is the number of MCMC samples and T represents total number of time steps (M should be selected large enough so it approximates a standard normal distribution well); (2) Sample a set of $m^{*(i)}$ and $\sigma_m^2{}^{*(\mathrm{i})}$ values, along with other rainfall-runoff model parameters θ^*; (3) Apply $m^{(i)}$ and $\sigma_m^{2(\mathrm{i})}$ to the ith Gaussian deviate vector $(\omega^{(i)})_{t=1,\ldots,T}$ to obtain the multiplier $(\Phi^{(i)}) = m^{*(i)} + (\omega^{(i)})_{t=1,\ldots,T} \cdot \sigma_m^{*(i)}$; (4) Calculate Bayesian likelihood and posterior; (5) Accept or reject the sampled parameter set (including $m^{*(i)}$, $\sigma_m^2{}^{*(\mathrm{i})}$ and other rainfall-run off model parameters) based on the MCMC algorithm.

[6] Motivated by the main goal of developing a procedure useful for practical applications, the algorithm described above is much more computationally efficient than the BATEA algorithm because of its reduced dimensionality ($n + 2$ versus $n + N_t$). The efficiency is achieved at the expense of assuming that the presampled fixed pool of Gaussian deviates are adequate approximation of the Gaussian distribution, and m and σ_m^2 are constant over the entire data period.

3. Constraint on Input Error Variance

[7] RKK2009 pointed out that for IBUNE to work properly, the input error variance needs to be constrained to a relatively small range, i.e., $m \in [0.9, 1.1]$ and $\sigma_m^2 \in [10^{-5}, 10^{-3}]$ (RKK2009, section 6). Our own study suggested that the range for m and σ_m^2 can be relaxed a little more, i.e., $m \in [0.8, 1.2]$ and $\sigma_m^2 \in [10^{-5}, 10^{-2}]$. Note that if we assume $m = 1$, and $\sigma_m^2 = 0.01$, it implies that there is a $\sim$33% chance that the rainfall multiplier is outside [0.9, 1.1], and there is a $\sim$5% chance that the multiplier is outside [0.8, 1.2]. Therefore, the range of potential rainfall multipliers is not trivially small. Nevertheless, we do realize and acknowledge this potential limitation. In fact, we conducted a synthetic study in order to test IBUNE's capability to recover the "true" input error parameters and rainfall-runoff model parameters. We found that when the rainfall input was perturbed with a relatively small variance (i.e., $\sigma_m^2 \leq 0.01$), IBUNE converges nicely to the "true" parameters even with a short calibration period (500 days). However, imposing a rainfall error with a relatively large variance would result in biased estimates of the "true" parameters. Therefore the tradeoff facing us is whether IBUNE with a constrained input error variance is better than the alternative which does not consider input error at all. Our study shows that the former is better. Furthermore, since IBUNE multipliers are not storm-dependent and can be calibrated in advance using historical data, it gives IBUNE an advantage over BATEA in applications to real-time forecasts to account for the input error uncertainty.

4. Caveat of the IBUNE Approach

[8] We stress again that ADS2007 IBUNE was developed with the intention to address explicitly the total uncertainties in the rainfall-runoff modeling process. It is not our intention to convey the message that IBUNE has no drawbacks or limitations. As a matter of fact, IBUNE is ultimately limited by the additive error assumption used in its formulation; that is, all errors are assumed additive and are manifested in streamflow observations. This is in contrast to BATEA, which is not limited by this assumption, at least with respect to input errors. In making this concession, however, IBUNE has gained computational efficiency over BATEA. Even as computational resources continue to advance, there is still a tradeoff to be made by hydrologic modelers between complexity and efficiency.

References

Ajami, N. K., Q. Y. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour. Res.*, *43*, W01403, doi:10.1029/2005WR004745.

Kavetski, D., S. Franks, and G. Kuczera (2003), Confronting input uncertainties in environmental modeling, in *Calibration of Watershed Models*, *Water Sci. Appl. Ser.*, vol. 6, edited by Q. Duan et al., pp. 49–68, AGU, Washington, D. C.

Renard, B., D. Kavetski, and G. Kuczera (2009), Comment on "An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction," *Water Resour. Res.*, *45*, W03603, doi:10.1029/2007WR006538.

N. K. Ajami, Berkeley Water Center, University of California, 413 O'Brien Hall 1718, Berkeley, CA 94720-1718, USA. (newshaajami@berkeley.edu)

Q. Duan, Atmospheric, Earth, and Energy Department, Lawrence Livermore National Laboratory, P.O. Box 808, 7000 East Avenue, Livermore, CA 94551, USA.

S. Sorooshian, Department of Civil and Environmental Engineering, University of California, E-4130 Engineering Gateway, Irvine, CA 92697-2175, USA.