Discussion of *Comparison of three global optimization algorithms for calibration of the Xinanjiang model parameters* by Dong-mei Xu, Wen-chuan Wang, Kwok-wing Chau, Chun-tian Cheng and Shou-yu Chen, 2013 *Journal of Hydroinformatics* 15 (1), 174–193, doi: 10.2166/hydro.2012.053

K. S. Kasiviswanathan and K. P. Sudheer

K. S. Kasiviswanathan
K. P. Sudheer (corresponding author)
Department of Civil Engineering,
Indian Institute of Technology Madras,
Chennai – 600036,
India
E-mail: *sudheer@iitm.ac.in*

Calibration of hydrological models is a complex optimization task as parameters induce a large amount of uncertainty in model simulations. In this context, the work presented by the authors to compare the performance of three different global optimization algorithms in calibrating the Xinanjiang model is important. The authors should be commended for their extensive effort in bringing out the effectiveness of three different optimization algorithms in calibrating the model. The discussers would like to comment some related issues, and feel that there are certain considerations that need further clarifications, which would help readers in better understanding and extending the authors' work.

The results presented in Table 6 (Xu *et al.* 2013) suggest that all the three algorithms work in a similar way, while both SCE-UA and SCEM-UA are found to be slightly better compared to GA in 3-hour forecast analysis. The SCEM-UA is found to be better than the other two algorithms in 1-hour forecasts (Table 10, Xu *et al.* 2013). One of the major concerns, the discussers feel, is about the selection of data for calibration and validation of 3-hour and 1-hour lead forecasts by the model (Tables 4 and 5 for 3-hour forecasts and Tables 8 and 9 for 1-hour forecasts, Xu *et al.* 2013). It is noted from these tables (3, 4, 8 and 9, Xu *et al.* 2013) that different data sets are used for demonstrating the model performance at different lead time forecasts (flood data pertaining to the period 1984–2000 is used for 3-hour forecast analysis, and flood data corresponding to 2000–2006 is used for 1-hour forecast analysis, without any common data among them). The discussers fail to understand the rationale behind the selection of such data. This raises a major concern about the validity in comparing the performance of the algorithms at different lead times (Table 6 and Table 10 together, Xu *et al.* 2013), since the uncertainty induced by the input information to the model might be different in both the data sets. The discussers feel that this might be one of the reasons for observing a higher predictive uncertainty in 1-hour lead forecasts compared to 3-hour lead forecasts. The discussers wonder if the authors have analyzed the results by interchanging the data used in different lead time analysis. It would be very informative if such information is present in the manuscript.

The figures (Figures 3 and 4, and also Figures 5 and 6, Xu *et al.* 2013) that depict the simulations from the three algorithms are not very confirmatory about the effectiveness of the algorithms (in fact these figures are not very clear). It would have been good if the authors presented the converged value of the objective function (OF) for each algorithm. If the values of OF were of similar order, the uncertainty in the model predictions could be used for comparing the optimization algorithms. While the authors have presented the uncertainty region from model as well as the parameters (in Figures 3 and 4, and also in Figures 5 and 6, Xu *et al.* 2013), the authors have not provided any

doi: 10.2166/hydro.2014.016

details about the procedure they employed to quantify these uncertainties. The discussers do not understand the way in which they have separated the model uncertainty and parameter uncertainty. Since the quantification of predictive uncertainty is a complex task, which is faced by the modeling community in general, the discussers feel that the method of constructing the uncertainty region (presented in these figures), if discussed, should have been very useful to the readers.

Further, the authors vaguely mention model and measurement uncertainty while discussing the results presented in Figures 5 and 6 (Xu *et al.* 2013), but the discussers fail to understand how the authors have correlated these uncertainties with the prediction uncertainty. There is no mention about the amount of measurement uncertainty present in the observed flows, and also no discussion is found about the difference between parameter and model uncertainty. It is well known that the model error is mismatched between the observed and the simulated values due to inherent uncertainty in the process (Shrestha & Solomatine 2008). These uncertainties mainly arise from input, parameter and model structure. The input uncertainty is mainly due to measurement and sampling error. The parametric uncertainty lies in the inability to identify a unique set of best parameters. The simplification, inadequacy and ambiguity in description of real world process through mathematical equation leads a model structure uncertainty.

A major concern in hydrologic modeling is the equifinality of the model parameters where multiple combinations of parameter values may yield the same model output (Wagener & Kollat 2007). Consequently, the identifiability of optimal combinations of parameters that result in a truly calibrated model is a major challenge (Cibin *et al.* 2010). In this context, the discussers wonder if the authors have made any effort to verify the variability between parameter combinations resulted from calibrating the model using different optimization algorithms. As a conclusion, the authors mention that SCEM-UA generates useful information about the nature of the response surface in the vicinity of the optimum. We did not find any results that substantiate this conclusion; also, does it mean that the other two algorithms did not converge to the optimum?

The coverage probability (also known as percentage of coverage) and width of the prediction interval are the two major indices reported in literature for evaluating prediction interval (Xiong *et al.* 2009; Zhang *et al.* 2009; Alvisi & Franchini 2011; Kasiviswanathan & Sudheer 2013). The coverage probability measures the percentage of observed values that fall within the prediction band. In theory, if the width of prediction band is wider, it covers most of the observed values. However, in order to include more observed values in the prediction band, compromising on the width of the prediction band is not desirable. Since these measures are conflicting, a desired solution is to have maximum coverage with a narrow prediction band. In this premise, the discussion of authors that total prediction uncertainty range brackets the observed discharge during almost the entire 1-hour time step does not seem to be appropriate. The discussers feel that merely bracketing the entire observed values in a band does not ensure that the model predictions are less uncertain, rather one should expect a narrow prediction band that contains most of the observed values. In fact, it is very difficult to compare across uncertainty bands unless the band is numerically evaluated using certain indices (Zhang *et al.* 2009; Kasiviswanathan & Sudheer 2013).

Finally, the discussers would like to bring out the following issues related to the authors' work, which if clarified/answered, would better enlighten the readers.

1. The statement of the authors, '… greater uncertainty for 1 hour forecasting than for 3 hour forecasting under the current model structure and field data'. Generally, one expects higher uncertainty in higher lead time forecasts as one is trying to forecast a value with information available only at many time steps behind. This may plausibly be due to the difference in data used for the analysis. The discussers feel that the appropriate justification for such an observation in their results should have been presented.
2. The authors present the basic statistics of the parameters and the correlation between parameters that are resulted from SCEM-UA with highest posterior probability. While they have mentioned that the coefficient of variation of parameters D_m, I_m is very high, there is no mention about the uncertainty induced by these parameters. The basic question that arises is that whether the parametric uncertainty presented in Figures 3 and 4 explains this.

3. The authors also mention that the cross correlation between the parameters are high for some parameters in 1-hour ahead forecast compared to 3-hour ahead forecast, and based on this observation they infer about the complexity of optimization in two different lead hour forecast. The discussers fail to understand the connection between the 'high correlation' and the 'complexity of optimization' without any additional information.
4. It would have been nice if the authors had presented the prior and posterior probability distribution of parameters. In addition, no information is provided about the number of iterations required for convergence of the algorithm and also about the sample size.
5. In the model description, the authors mentioned that runoff is first transformed into discharge by linear system calculated from the runoff generating component. The discussers are concerned about the difference between runoff and discharge.

Despite the issues discussed above, the research work presented by the authors is an important contribution in improved flood forecasts with high confidence. A response from the authors to clarify certain issues and queries raised in this discussion would not only help in better understanding the authors' work but also increase the dissemination of the authors' work.

REFERENCES

Alvisi, S. & Franchini, M. 2011 Fuzzy neural networks for water level and discharge forecasting with uncertainty. *Environ. Model. Softw.* **26** (4), 523–537.

Cibin, R., Sudheer, K. P. & Chaubey, I. 2010 Sensitivity and identifiability of stream flow generation parameters of the SWAT model. *Hydrol. Processes*. 10.1002/hyp.7568.

Kasiviswanathan, K. S. & Sudheer, K. P. 2013 Quantification of the predictive uncertainty of artificial neural network based river flow forecast models. *Stoch Environ. Res. Risk Assess.* **27**, 137–146.

Shrestha, D. L. & Solomatine, D. P. 2008 Data-driven approaches for estimating uncertainty in rainfall runoff modelling. *J. River Basin Manage.* **6** (2), 109–122.

Wagener, T. & Kollat, J. 2007 Visual and numerical evaluation of hydrologic and environmental models using the Monte Carlo Analysis Toolbox (MCAT). *Environ. Model. Softw.* **22**, 1021–1033.

Xiong, L., Wan, M., Wei, X. & O'Connor, K. 2009 Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation. *Hydrol. Sci. J.* **54** (5), 852–871.

Xu, D.-M., Wang, W.-C., Chau, K.-W., Cheng, C.-T. & Chen, S.-Y. 2013 Comparison of three global optimization algorithms for calibration of the Xinanjiang model parameters. *J. Hydroinf.* **15** (1), 174–193.

Zhang, X., Liang, F., Srinivasan, R. & Van Liew, M. 2009 Estimating uncertainty of streamflow simulation using Bayesian neural networks. *Water Res. Res.* **45**, W02403.