PROOF COPY [HYENG-6796] 018912QHY

Discussion of "Automatic Calibration of the U.S. EPA SWMM Model for a Large Urban Catchment" by J. Barco, K. M. Wong, and M. K. Stenstrom

April 2008, Vol. 134, No. 4, pp. 466–474.
DOI: 10.1061/(ASCE)0733-9429(2008)134:4(466)

James E. Ball, M.ASCE[1]
[1]School of Civil and Environmental Engineering, Univ. of Technology-Sydney, Sydney, Australia

The authors have contributed to and illustrated the need for continued research into calibration methodologies for complex catchment modeling systems. The search for suitable approaches for parameter evaluation has resulted in the development of many new techniques and concepts. The authors have continued the traditional approach of identifying a unique optimal parameter set or near optimal parameter set that is assumed to represent the generic catchment characteristics. There are concerns with this approach and outlined herein is one of these concerns, namely the identification of a globally optimal set of parameters that does not represent the generic catchment characteristics due to the following:

- Data errors in the input to the model and that used to assess the performance of the parameter values;
- Uniformity in the performance of alternative parameter sets.

Many studies have demonstrated the difficulties, if not the impossibility, of finding a unique optimal parameter set due to uncertainty of model structure, errors associated with input and observed data, and interactions between parameters (Kuczera 1983; Sorooshian et al. 1983; Beven and Binley 1992; Gan and Biftu 1996). As a result of these sources of error, an optimal parameter set for one set of events may not be optimal for other events. Searching for a unique optimal value may lead to the parameter evaluation being based on the best "curve fitting" rather than the best representation of the catchment processes. This was explored by Choi and Ball (2002) who proposed the use of monitoring data to define the point where further parameter modification does not result in additional information being extracted from the available data. The conceptual basis of this approach is based on dividing the available data into calibration, monitoring, and validation data sets and using the monitoring data at each (or a predefined number of iterations) iteration of the parameter modification as the optimal set is obtained to ensure that the objective function used to measure the improvement in performance decreases for events other than those used for the calibration. As shown in Fig. 1, the process ends when further iterations result in a decrease in the performance of the monitoring data even though the performance of the calibration data continues to improve. Results from an application of this approach with SWMM to a catchment in Sydney, NSW (Choi and Ball 2002) are shown in Table 1; for this application of the approach 81.25% of the calibrations were concluded prior to reaching the maximum number of iterations. Choi and Ball (2002) found that continuing the search beyond the "early stop point" resulted in a decrease in the performance of the parameter values when applied to different events to those being used for the calibration and, hence, they postulated that the perceived improvement in performance past the "early stop point" was due to the model performing as a "curve fitting" transformation rather than one where the model was replicating the major catchment processes.

The authors have not tested the parameter values developed during their optimization process with storm events not used as part of the calibration process. If the authors applied a monitoring approach to their data set, it would be interesting to see if the same parameter values were obtained.

This problem of identifying the point where further modifications to parameter values does not result in the extraction of additional information from the available data suggests that there are many alternative combinations of parameter values that result in similar performance. This has lead to development of techniques for estimating the parameter uncertainty for simple catchment modeling systems. Simple modeling systems can be categorized as those systems where evaluation of only a few parameters is required for application. Examples of these approaches are:

- Bayesian methodology first explored by Kuczera (1983), whereby parameter uncertainty is described by the posterior distribution, which expresses the probability of the parameter values given the observed data. Marshall et al. (2004), however, claim that while the Bayesian frameworks are widely used, the implementation of Bayesian procedures has been hindered due to difficulties in summarizing and exploring the posterior distribution of parameters for complex catchment modeling systems.
- Markov Chain Monte Carlo (MCMC) approaches as presented by Kuczera and Parent (1998), Bates and Campbell (2001), Marshall et al. (2004), and Gallagher and Doherty (2007). While these approaches provide computationally feasible implementations of Bayesian inference with the aim of generating samples of parameter values from the posterior distribution with reasonable efficiency, a priori knowledge about the proposal distribution of parameters is crucial for effective implementation of a MCMC algorithm.
- The generalized likelihood uncertainty estimation (GLUE) method as presented by Beven and Binley (1992). Application of a GLUE methodology usually involves making a large number of Monte Carlo (MC) simulations with different sets of parameter values, generated randomly from uniform distributions within the feasible parameter space. While the GLUE methodology is capable of exploring the whole search space, it is computationally inefficient when very large numbers of initial parameter sets are required (Spear et al. 1994; Bates and Campbell 2001). To mitigate this problem, a number of studies have investigated methods for improving the efficiency of MC-based techniques. An approach commonly adopted (Helton and Davis 2003) has been to use a more efficient sampling

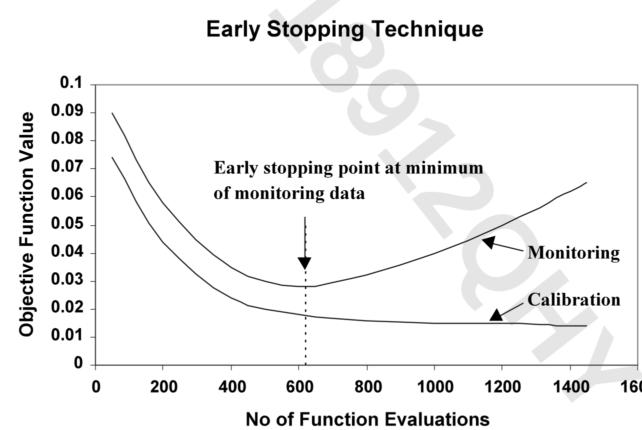

Fig. 1. Early stopping technique (adapted from Choi and Ball 2002)

PROOF COPY [HYENG-6796] 018912QHY

PROOF COPY [HYENG-6796] 018912QHY

Table 1. Occurrence Rate of the Minimum Function Value (Adapted from Choi and Ball 2002)

Event	SP (%)	ESP (%)	EP (%)
Nov. 01, 1994	0	87.5	12.5
Dec. 22, 1994	0	87.5	12.5
Jan. 04, 1995	0	87.5	12.5
Feb. 28, 1995	25	62.5	12.5
Mean	6.25	81.25	12.5

Note: SP=start point; ESP=early stop point; and EP=end point

Table 2. Performance of Behavioral Sets of Control Parameter Values (Adapted from Fang and Ball 2007)

Storm event	Jan. 5, 1998	Dec. 14, 1998	Feb. 24, 1999
Average RMSE	0.0783	0.0880	0.0715
Standard deviation of RMSE	0.0012	0.0010	0.0014

algorithm, such as Latin hypercube sampling. Another alternative was presented by Khu and Werner (2003) who used a hybrid genetic algorithm and artificial neural network, known as GAANN to improve the efficiency of a GLUE approach.

Building on these studies into the parameter uncertainty, Fang and Ball (2007) used a genetic algorithm (GA) within a GLUE framework to investigate the parameter uncertainty associated with the application of SWMM for flow prediction in an urban catchment. In this case the approach was not limited to a simple catchment modeling system but rather to a complex catchment modeling system with a significant number of spatially variable interrelated parameters. Defining a behavioral set as being a set of parameters where the RMSE in discharge was less than 0.1 m^3/s, Fang and Ball (2007) found, after 50 generations with 1,000-parameter sets per generation, approximately 900 alternative sets of parameter values meeting the criterion. Shown in Table 2 are the mean and standard deviation of the RMSE for these behavioral parameter sets. Using these values gives a coefficient of variation of approximately 1.5%, which can be interpreted as suggesting that there is minimal difference in the performance of any one of the approx. 900 behavioral parameter sets highlighting the difficulty of selecting one set of parameter values as the most desirable.

Using the concept that there are multiple alternative sets of parameter values that result in similar performance, it would be interesting if the authors could provide information about the variability in the predicted flows of the best-performing sets of parameter values. Inclusion of the concept of monitoring the calibration process in determining the best performing set of parameter values would be useful also.

References

Bates, B. C., and Campbell, E. P. (2001). "A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling." *Water Resour. Res.*, 37(4), 937–947.

Beven, K., and Binley, A. (1992). "The future of distributed models: Model calibration and uncertainty prediction." *Hydrolog. Process.*, 6(■), 279–298.

Choi, K. S., and Ball, J. E. (2002). "A generic calibration approach: Monitoring the calibration." *Proc., 2002 Hydrology and Water Resources Symp.*, I. E. Aust, ■.

Fang, T., and Ball, J. E. (2007). "Evaluation of spatially variable parameters in a complex system: An application of a genetic algorithm." *J. Hydroinform.*, 9(3), 163–173.

Gallagher, M., and Doherty, J. (2007). "Parameter estimation and uncertainty analysis for a watershed model." *Environ. Modell. Software*, 22(7), 1000–1020.

Gan, T. Y., and Biftu, G. F. (1996). "Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure." *Water Resour. Res.*, 32(12), 3513–3524.

Helton, J. C., and Davis, F. J. (2003). "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems." *Reliab. Eng. Syst. Saf.*, 81, 23–69.

Khu, S. T., and Werner, M. G. F. (2003). "Reduction of Monte-Carlo simulation runs for uncertainty estimation in hydrological modeling." *Hydrology Earth Syst. Sci.*, 7(5), 680–692.

Kuczera, G. (1983). "Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty." *Water Resour. Res.*, 19(5), 1151–1162.

Marshall, L., Nott, D., and Sharma, A. (2004). "A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling." *Water Resour. Res.* ■(■), ■-■.

Sorooshian, S., Gupta, V. K., and Fulton, J. L. (1983). "Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: Influence of calibration data variability and length on model credibility." *Water Resour. Res.*, 19(■), 251–259.

Spear, R. C., Grieb, T. M., and Shang, N. (1994). "Parameter uncertainty and interaction in complex environmental models." *Water Resour. Res.*, 30(11), 3159–3169.

PROOF COPY [HYENG-6796] 018912QHY

PROOF COPY [HYENG-6796] 018912QHY