AUTOMATIC CALIBRATION AND PREDICTIVE UNCERTAINTY ANALYSIS OF A SEMI-DISTRIBUTED WATERSHED MODEL

Z. Lin[1/] and D.E. Radcliffe[2/]

AUTHORS: [1/] Postdoctoral Research Associate, [2] Professor, Department of Crop and Soil Sciences, The University of Georgia, Athens, GA 30602-7274
REFERENCES: Proceedings of the 2005 Georgia Water Resources Conference, held April 25-27, 2005, at The University of Georgia. Kathryn J. Hatcher, editor, Institute of Ecology, The University of Georgia, Athens, Georgia.

Abstract. A two-stage routine has been developed for automatic calibration of the Soil Water Assessment Tool (SWAT, a semi-distributed watershed model) that finds the best values for the model parameters, preserves the spatial variability in essential parameters, and leads to a measure of the model prediction uncertainty. We calibrated the stream flow in the Etowah River measured at Canton, GA (a watershed area of 1,580 km^2) for the years 1983-1992 and used the years 1993-2001 for validation. Calibration for daily and monthly flow produced a very good fit to the measured data. Nash-Sutcliffe coefficients for daily and monthly flow over the calibration period were 0.60 and 0.86, respectively; they were 0.61 and 0.87 respectively over the validation period. Regardless of the level of model-to-measurement fit, non-uniqueness of the optimal parameter values necessitates uncertainty analysis for model prediction. The nonlinear prediction uncertainty analysis showed that caution must be exercised when using the SWAT model to predict short-term (7-day average) flows, especially under low and high flow conditions.

INTRODUCTION

Commonly, modelers only have a set of observations of stream discharge and/or water quality parameters at a watershed outlet to calibrate (semi-)distributed models. Neither a sampling-based global (such as Monte Carlo) nor a gradient-based local method (such as Levenberg-Marquardt) is suitable for such a situation. On one hand, the cost in terms of model runs required for sampling-based method to simultaneously calibrate the large numbers of parameters of a (semi-)distributed watershed model may be too high to be realistic; on the other hand, the highly correlated relationship among parameters and the susceptibility to initial conditions of local methods make it unavoidable for gradient-based local methods to fail in finding an unique optimal set of parameter values. A few attempts have been reported on combining both global and local methods to automatically calibrate the watershed models. In such a multi-step approach, the global methods served to find good starting points for a subsequent local search, and the local methods were used to fine-tune the parameter values. However, to our knowledge, the issues of preservation of the heterogeneous design of the (semi-)distributed watershed models and the stabilization of the ill-posed problems resulting from over-parameterization remained unaddressed. In this paper, it is our goal to present a two-stage routine for automatically calibrating the semi-distributed SWAT watershed model that will find the optimal values for the model parameters, preserve the spatial variability in essential parameters and leads to a measure of the model prediction uncertainty.

In the context of environmental modeling, where the model is usually complex and highly over-parameterized, the model can be useful in predicting the behavior of natural systems under the similar situations where the model has been calibrated against the historical data. Due to the uncertainties associated with the model structure, parameter values, initial conditions, and measured data, model predictions are inevitably fraught with uncertainty. A number of methods are available to assess the effect of the parameter uncertainty (initial conditions can be treated as model parameters) and measurement errors to the model prediction uncertainty (Beck, 1987; and many others). Among these methods, Monte-Carlo based methods hold many attractions, such as providing probability distribution for model predictions, but their main disadvantage is that the cost in model runs is extremely high, especially when parameter number is more than just a few in a (semi-)distributed watershed model. In this research, a nonlinear calibration-constrained method based on the theory presented by Vecchia and Cooley (1987) has been used to provide an estimate of the flow prediction uncertainty of the SWAT model on various flow conditions.

WATERSHED MODEL AND STUDY SITE

The SWAT model was developed by the USDA-ARS to predict the impact of land management practices on water, sediment, and agricultural chemicals in large basins (Arnold *et al.*, 1998). In SWAT, a watershed is partitioned into a number of sub-basins, and with additional subdivisions within such a sub-basin to represent different soils and land use combinations. Each of these subdivisions is referred to as a hydrologic response unit (HRU) and is assumed to be spatially uniform in terms of soil, land use, topographic and climatic data. A water budget is computed for each HRU based on precipitation, runoff, ET, percolation, and return flow from subsurface and ground water flow. Subdivision of a watershed into HRUs allows the model to reflect heterogeneity of the watershed. Thus, the SWAT model is considered as a semi-distributed watershed model.

The Etowah River is the main tributary of Lake Allatoona, which has a Total Maximum Daily Load (TMDL) limit for phosphorus (P). Our long-range objective is to use SWAT to determine the uncertainty in non-point sources of P loading to Lake Allatoona and develop a framework for trading P credits between point and non-point sources in the Etowah River basin. From the U.S. Geological Survey (USGS) gauging station at Canton, GA, we delineated an area we called the Upper Etowah River watershed which covered approximately 1,580 square kilometers. The watershed was subdivided into 6 subbasins (Figure 1) and 48 HRUs. The primary land covers were forest (89.7%), grassland/pasture (7.9%), and agriculture (1.9%); urban cover was less than 0.5%. The soils consisted primarily of mapping units with hydrologic group categories of B and C, having slow to moderate infiltration rates. The land use data were obtained from the National Land Cover Dataset (NLCD) from 1991 to 1992 and the soils data were obtained from the State Soil Geographic Data Base (STATSGO). The hydrologic phase of this model was calibrated against the daily streamflow records from USGS gauging station at Canton, GA (available online at http://waterdata.usgs.gov/nwis). The rainfall data include the precipitations gauged at six different weather stations within the watershed (see Figure 1). The minimum and maximum air temperature measurements were taken at two of those weather stations.

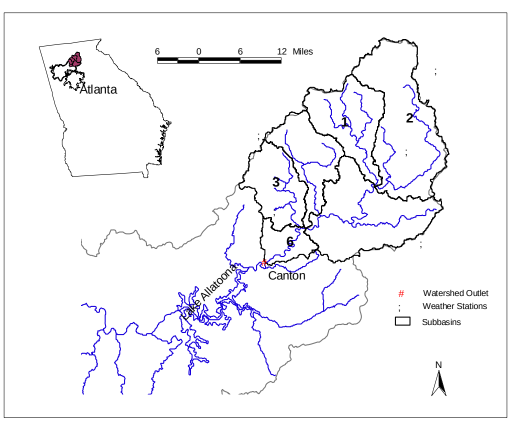

Figure 1. Upper Etowah River watershed

NUMERICAL METHODS AND SOFTWARE

Two-Stage Calibration Strategy

The proposed automatic calibration scheme of the semi-distributed SWAT model consists of two stages. In the first stage, we simplified the SWAT model to a lumped model by eliminating the spatial variability of parameters that could not be calculated from available GIS measurements and needed to be determined through calibration. By this simplification, the dimension of the adjustable parameter space was reduced so that *global* search algorithms such as SCE-UA (Shuffled Complex Evolution – University of Arizona, Duan *et al.*, 1992) could be used in searching for the "best" parameter sets. However, our primary goal of the first stage is not to try to find the unique global minimum of the objective function (if it exists), but to find a reasonable set of parameter values that would serve as starting points of the parameters to be estimated in the second stage. Therefore, we used less stringent convergence criterion for this global searching method, which dramatically reduced the number of model runs. Subsequently, in the second stage, the spatial variability of the original model parameters was restored and the number of the calibrated parameters was sharply increased. Hence, a *local* search method (Levenberg-Marquardt method) was preferred to find the more distributed set of parameters using the results of the previous stage as starting values. Furthermore, in order to prevent numerical instabilities and parameters taking extreme values, a strategy called "regularization" (Doherty, 2003) was adopted in the second stage, by which the distributed parameters were constrained to vary *as little as possible* from the initial values of the lumped parameters.

Nonlinear Calibration-Constrained Predictive Uncertainty Analysis Method

In this nonlinear calibration-constrained method, we first specify a prediction of interest (for example, annual flow volume or 7-day average flow) whose uncertainty requires exploration; then we find a parameter set that maximized or minimized that prediction while still maintaining the model in a *calibrated state*. The calibrated state was defined by an upper objective function limit, a number that was slightly larger (for example, 15%) than the minimum objective function found in the model calibration stage. Note that it is the maximum and the minimum values of the specified model prediction, instead of the parameter sets generating these extreme values, that is our major interest. Therefore, the prediction interval limited by the maximum and minimum values represents the prediction uncertainty of the watershed model on the specified prediction. Compared to the Monte Carlo based method, the nonlinear calibration-constrained method requires much less model runs to give an estimated interval for a model prediction. But it is unable to provide a corresponding probability distribution of the model prediction. The PEST (Parameter Estimation, http://www.sspa.com/pest/) free software was used to conduct the two-stage automatic calibration and prediction uncertainty analysis of the SWAT model.

RESULTS AND DISCUSSION

Model Calibration and Validation

During the calibration, a multi-objective function, which is a weighted sum of squares of the mismatches between the measured and predicted daily flow, monthly volume, and the fraction of flow exceedance, was minimized. The calibration period included 10 years from January 1983 to December 1992. During this period, the spring of 1990 was considered the wettest season and the years of 1985 and 1986 were the driest years. The following nine years from January 1993 to December 2001 were selected as the validation period.

Sixteen SWAT parameters, including curve numbers (CN), soil available water content (AWC), soil evaporation compensation factor (ESCO), Manning coefficients, etc., which govern the surface and subsurface hydrological processes and stream routing were adjusted in the first stage. It is important to note that each of the 16 adjustable parameters was treated identical across different sub-basins or HRU's. At the second stage, instead of assigning the same curve number (CN) for one type of land cover, we specified a unique curve number (CN) for each HRU and assigned different soil evaporation compensation factors (ESCO) and different the available water capacities (AWC) to different soil types, such that the adjustable parameter number increased from 16 to 69.

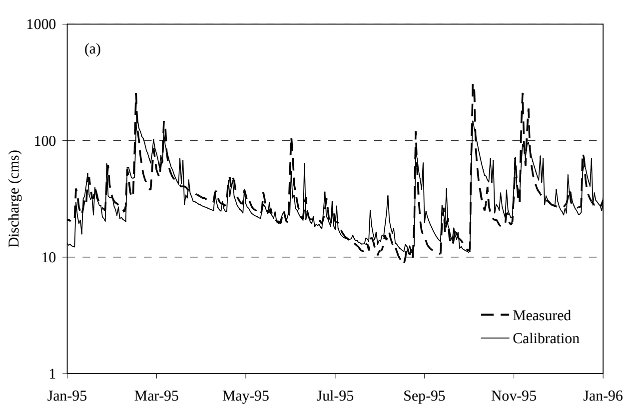

Figure 2. Observed and modeled stream discharge at Canton, GA during one year (1995) of the validation period.

Modeled and measured daily stream discharge through one year (1995) of the validation period are shown in Figure 2. The restriction of graphed flows only a part of the validation period is done for the sake of clarity. Graphs over the remainder of the calibration and validation periods are similar. Note also that the flow axis is logarithmic to allow a better comparison between model outputs and field measurements under both high and low flow conditions.

The calibration and validation results obtained from our proposed two-stage automatic calibration procedure were also compared with results from an optimal use of the SCE-UA method that is favored by many hydrologists. Nash-Sutcliffe coefficient is usually used to quantify the closeness of fit between modeled and observed time series (Nash and Sutcliffe, 1970). The Nash-Sutcliffe coefficients for our two-stage auto-calibration procedure were similar to the SCE-UA method during both calibration and validation periods, but the number of model evaluations required by our routine was much less than that required by the SCE-UA method (Table 1).

Predictive Uncertainty Analysis

The previous calibration process found that the minimum objective function value was 1450 (shown in Table 1). For the predictive uncertainty analysis, we assumed that the model was in a calibrated state when the corresponding objective function is less than 1675 (1450 plus 15%). The predictive uncertainty analysis was conducted for long-term (annual) and short-term (daily) stream discharge prediction. Table 2 shows the minimum and maximum predictions of SWAT, along with the observed stream discharge at Canton over the validation period. The results for annual flow volume prediction are listed in the left panel and those for 7-day average flow prediction in the right panel. Most of the observed annual (long-term) flows fell inside of the SWAT prediction intervals, while most of the observed daily (short-term) flows fell outside of the intervals. It should be noted that the prediction interval given by the predictive uncertainty analysis may be smaller than it should be because the underlying nonlinear calibration-constrained method uses a local method (e.g., Levenberg-Marquardt method) to search for the minimum or maximum value of the model prediction. Regardless, Figure 3 shows that all model-generated stream flows resulting from the uncertainty analysis are very close to the observations over the calibration period and they are indistinguishable from each other.

Table 1. Comparison of the two-stage routine and the optimal use of the SCE-UA method

	Calibration Period		Validation Period	
	Two-stage	SCE-UA	Two-stage	SCE-UA
Approximate model runs	3000	8000	—	—
Objective function value	1450	1903	1404	1400
Nash-Sutcliffe Coefficients	0.61 / 0.86 *	0.61 / 0.83	0.62 / 0.89	0.62 / 0.85

* The numbers in the left side of the back slash (/) are measurements of the "goodness-of-fit" between the modeled and observed daily flow time series over the entire calibration/validation period; while the numbers in the right side of the back slash (/) are measurements of the corresponding monthly volume time series.

Table 2. Uncertainty analysis of SWAT long- and short-term stream flow prediction

Annual flow prediction ($\times 10^9$ m^3)				7-day average flow prediction (m^3/s)			
Year	Observed	Minimum	Maximum	Flow scenario	Observed	Minimum	Maximum
1993	1.103	1.082	1.176	Low (1993)	7.551	6.502	11.62
1994	1.083	1.051	1.124	Low (1996)	8.941	10.56	12.02
1995	1.079	1.057	1.091	Low (1999)	7.624	8.053	10.96
1996	1.291	1.113	1.138	Medium (1993)	10.15	11.37	14.70
1997	1.074	1.054	1.136	Medium (1996)	11.23	13.79	15.33
1998	1.308	1.308	1.411	Medium (1999)	17.98	16.05	19.84
1999	0.686	0.697	0.766	High (1993)	98.21	67.74	68.86
2000	0.609	0.615	0.703	High (1996)	192.5	80.36	92.96
2001	0.760	0.751	0.857	High (2000)	86.13	54.72	60.25

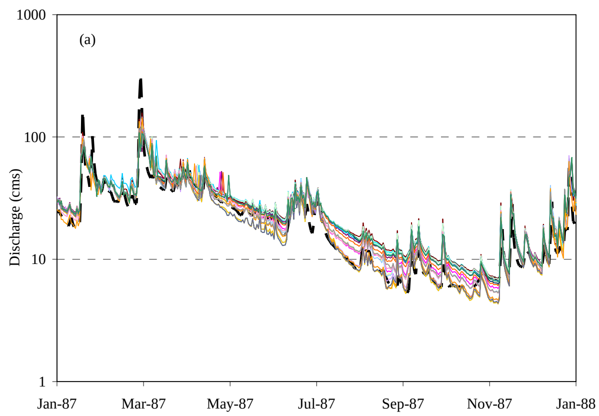

Figure 3. Observations (thick dashed line) and model-generated daily stream flows from uncertainty analysis (thin solid lines) during one year (1987) of the calibration period.

CONCLUSIONS

The two-stage automatic calibration strategy, which is bound to exploit the merits of both global and local optimization methods while avoiding their faults, has been successfully applied to SWAT (a semi-distributed watershed model) calibration. It may also be extended to the automatic calibration of other (semi-)distributed watershed models without difficulty. This two-stage automatic model calibration not only produced a very good fit of a model against observations, but also preserved the heterogeneity of watershed, which is essential in the analysis of the impact of the land uses and soils to the water quality in streams. Based on the automatic calibration results, predictive uncertainty analysis of SWAT has been conducted as well. The results of uncertainty analysis indicated that SWAT model is more reliable in long-term flow prediction than short-term flow prediction.

LITERATURE CITED

Arnold, J.G., Srinivasan, S., Muttiah, R.S., and Williams, J.R., 1998. Large area hydrologic modeling and assessment, Part I: model development. *Journal of American Water Resources Association (JAWRA)*, **34**(1), 73-89.

Beck, M.B, 1987. Water Quality Modeling: A Review of the Analysis of Uncertainty. *Water Resources Research*, **23**(8):1393-1442.

Doherty, J., 2003. Ground water model calibration using pilot points and regularization. *Ground Water*, **41**(2), 170-177.

Duan, Q.Y., Sorooshian, S., and Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall runoff models. *Water Resources Research*, **28**(4), 1015-1031.

Nash, J.E., and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. I. A discussion of principles. *Journal of Hydrology*, **10,** 282-290.

Vecchia, A.V., and Cooley, R.L., 1987. Simultaneous confidence and prediction intervals for nonlinear regression models with application to a groundwater flow model. *Water Resources Research*, **23**(7), 1237-1250.