WILEY
InterScience®
DISCOVER SOMETHING GREAT

Comment

Comment on Cao W, Bowden BW, Davie T, Fenemor A. 2006. 'Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability.' *Hydrological Processes* 20(5): 1057–1073

Kati White Migliaccio[1]* and Indrajeet Chaubey[2]

[1] *Tropical Research and Education Center, University of Florida, 18905 SW 280th St, Homestead, FL 33031, USA*
[2] *Purdue University, 225 S. University St, West Lafayette, IN 47907, USA*

INTRODUCTION

Recently, a manuscript was published with the objective to complete a multivariable and multisite calibration and validation for the Soil and Water Assessment Tool (SWAT) model on the Motueka River basin by Cao *et al.* (2006). Physically based distributed parameter models, such as SWAT, have become a preferred tool to predict watershed response not only at the watershed outlet, but also at intermediate points within the watershed. Calibrating and validating such models before they can be applied to make watershed decisions is always a challenge. A methodology to perform multisite and multiparameter calibration should be of interest to a wide rage of stakeholders working in the area of watershed management. However, upon reviewing this paper, we felt two topics warranted additional discussion.

MULTISITE CALIBRATION

Multisite calibrations are generally described using a multiobjective function. The purpose of this function is to provide optimization criteria for the multiple modeling objectives in a mathematical function (Gupta *et al.*, 1998; Yapo *et al.*, 1998). Multiobjective functions are often presented in hydrologic modeling efforts that consider multiple sites and/or multiple variables or multiple modeling objectives. Previously referred to as multiple objective programming (Mendoza *et al.*, 1986), this method was further developed by Yan and Haan (1991) to include

optimization or minimization criteria. Since then, generalized multiobjective functions have been defined in literature to mathematically represent the optimization of multiobjectives identified in a modeling effort. One example of a general multiobjective function was presented by Yapo *et al.* (1998):

$$\underset{w.r.t.\theta}{minimize}\, F(\theta) = \langle f_1(\theta), \dots, f_m(\theta) \rangle \quad (1)$$

where F is the multiobjective function and $f_1(\theta)$, $\dots f_m(\theta)$ are the m noncommensurable objective functions to be simultaneously minimized with respect to the parameters θ of the model. The general format of the multiobjective function presented in Equation 1 provides a functional basis for devising specific multiobjective functions. A similar approach was taken by White and Chaubey (2005), who identified a multiobjective function that considered different variables, time steps, and site locations for calibrating and validating a distributed parameter model for watershed flow and water quality response predictions:

$$F(O, P) =$$
$$\left\{ \begin{array}{l} minimize \sum_{i=1}^{y} \left(\sum_{j=1}^{c} \left[\sum_{k=1}^{v} f_1(O,P) \right] \right) \\ optimize \sum_{l=1}^{m} \left(\sum_{j=1}^{c} \left[\sum_{n=1}^{w} f_2(O,P), f_3(O,P) \right] \right) \end{array} \right\} (2)$$

where F is the multiobjective function, O is observed, P is predicted, f_1 is the relative error, f_2 is the coefficient of determination, and f_3 is the Nash–Sutcliffe coefficient, y is the number of years, c is the number of calibration sites, v is the number of variables evaluated annually, m is the number of months, and w is the number of variables evaluated monthly. Inclusion of

* Correspondence to: Kati White Migliaccio, Tropical Research and Education Center, University of Florida, 18905 SW 280th St, Homestead, FL 33031, USA. E-mail: klwhite@ufl.edu

multiobjective functions (such as Equation 1 and 2) to express a complex calibration procedure provides a clear indication of how the calibration procedure will be performed. Hence, multiobjective functions provide a concise expression of the calibration procedure and minimize misinterpretation of how the calibration process was performed.

In addition, multisite calibration of watershed models is more appropriately applied to sites that are not hydrologically connected. We illustrate this point using the two hypothetical watersheds presented in Figure 1. Each watershed contains three sampling sites. In Watershed A, Site 1 is independent of Sites 2 and 3; however, Site 2 is hydrologically connected to Site 1, and Site 3 is hydrologically connected to Sites 1 and 2 due to their nested characteristics. If a multisite calibration is performed on this watershed, the catchment area contributing to Site 1 will be included in three different calibration procedures. Another issue presented by nested calibration sites is that bias may occur in the calibration process. For example, if Site 1 is calibrated to achieve a specific simulated R^2 goal (as is the case described in *Cao et al.*, 2006), Sites 2 and 3 may require no additional calibration due to the minimization in error achieved with Site 1 calibration. Another approach to using nested watershed sampling sites in model calibration is to not use multiple
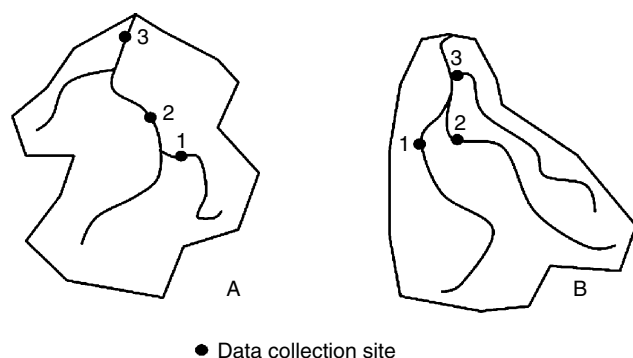

Figure 1. Two examples of watersheds are presented. Watershed A contains three data collection sites that are nested so that two sites encompass upstream site catchment areas; and Watershed B contains three data collection sites that are mutually exclusive and do not encompass upstream surface water catchment areas
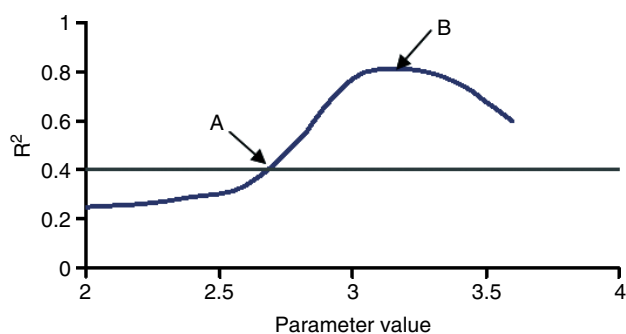

Figure 2. An example of a parameter is plotted with the model response R^2 as a function of the parameter value. The 'A' designates the selected value if the decision to stop calibration is $R^2 \geq 0.4$ and the 'B' designates the selected value if the decision to stop calibration is to minimize the error

sites for calibrating. Instead, use sites that are not nested for calibration and the remaining, nested sites for model validation.

Although nested data sites may be commonly found due to monitoring strategies, nonnested or mutually independent sites provide a better format for multisite calibration. Example Watershed B represents an alternative scenario where the sampling sites are mutually independent. Watershed B is an ideal candidate for multisite calibration since each site contributes to a larger area but there are no catchment area overlaps.

Furthermore, the flow calibration depicted in Figure 3 of the Cao *et al.* (2006) manuscript does not represent a multisite calibration. The flow calibration matrix provided by Cao *et al.* (2006) implies that each site (or catchment area) was calibrated separately. This was further indicated by text, which stated that '... multisite calibration started with the subcatchments farthest upstream, followed by the next downstream subcatchment ...'. This suggested that each site was calibrated independently from the previous. Calibration that is completed one site at a time deviates in the parameterization process from those that calibrate multiple sites simultaneously. This is primarily due to the parameterization structure in a particular model. For the SWAT model, parameters are spatially designated at three levels: watershed level, subbasin level, and Hydrologic Response Unit (HRU) level. Watershed level is the largest spatial unit and includes the entire modeled area; subbasins are hydrologically connected subunits of the watershed; and HRUs are unique combinations of soil and land cover within each subbasin. Hence, if calibration sites are parameterized separately and not simultaneously, parameter estimation may not be appropriately completed for parameters that are specified on the watershed level. For example, the watershed level parameter set selected for one calibration site might be different from the watershed level parameter set selected for the next calibration site, which would result in modification of the selected parameters from the previous site as well as modification of the calibration results.

CALIBRATION AND PARAMETER SELECTION

Watershed model calibration is part science and part art. The science of calibration relies on the physical relationships mathematically expressed in the model and the inputs measured in the watershed to simulate specific outputs. The art of watershed calibration includes the calibration methods used, particularly for parameter estimation. Hence, the replication of a modeling effort can only be completed if these two aspects of the modeling process are clearly defined.

The watershed model selected (as is the case of the SWAT model) has been well documented (Arnold *et al.*, 1998; Srinivasan *et al.*, 1998) and hence a detailed review of its mechanics was not necessary. The manuscript by Cao *et al.* (2006) included detailed information on modifications to the modeling framework that were

incorporated into their model simulations. However, little information was presented regarding the parameter estimation portion of the modeling process.

On the basis of information provided in the manuscript, the parameters selected by Cao *et al*. (2006) for calibration were: CSC, CN, GW, and SW. Typically, parameter selection for model calibration is based on sensitivity analyses of the model for the outputs of interest. The authors provide no explanation on how these three parameters were selected for model calibration. CN and CSC are described in the manuscript as representing curve number and canopy storage capacity, respectively. No definitions of GW or SW are given and their identities are not deducible using the information presented. The SWAT user's manual and SWAT theoretical documentation (Neitsch *et al*., 2002a,b) does not contain a parameter identified as GW or SW. There are several parameters that affect base flow predictions by the SWAT model, e.g. delay time for aquifer recharge (GW_Delay), threshold water level in shallow aquifer for base flow, base flow recession constant (GWQMN), revap coefficient (GW_REVAP) and many of these parameters have been used by modelers to calibrate the base flow (White and Chaubey, 2005 for a summary of these studies). Similarly, SW is defined as water content of soil profile on a given day (mm H_2O), and thus, is not a parameter, but an output calculated by the model.

In addition, the calibration procedure is identified as 'adjust'. This implies that the parameters were modified in some way and once the set criteria were met, calibration ended. Hence, parameter values that minimized errors were not selected, but rather, parameter values that met an R^2 designation were selected. This calibration approach is limiting (Yan and Haan, 1991). We graphically illustrated this limitation in Figure 2. A more robust calibration process is one that requires some degree of adjustment in a linear direction (either negative or positive), followed by an iterative process to determine the optimum parameter value. If merely an adjustment is made to the parameter until a model prediction R^2 numerical calculation is reached, error has not been minimized, but merely reduced (Figure 2, 'A'). A sound

calibration should involve optimization or minimization of a statistical measure (such as R^2 or Nash–Sutcliffe coefficient). The concept of optimizing R^2 is illustrated as 'B' in Figure 2. Hence, different parameter values would be estimated depending on the technique chosen (A = 2·7, B = 3·2).

The two topics we discussed, multisite calibration and calibration and parameter selection, are directly related to the appropriate application of a watershed model. Our goal in this response was to address the weaknesses that were present in the Cao *et al*. (2006) article regarding their model calibration so that future modeling endeavors might be completed with regard to our comments.

REFERENCES

Arnold JG, Srinivasan R, Muttiah RS, Williams JR. 1998. Large area hydrologic modeling and assessment. Part I: model development. *Journal of the American Water Resources Association* **34**: 73–89.

Cao W, Bowden WB, Davie T, Fenemor A. 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrological Processes* **20**(5): 1057–1073.

Gupta HV, Sorooshian S, Yapo PO. 1998. Towards improved calibration of hydrologic models: multiple and non-commensurable measures of information. *Water Resources Research* **34**(4): 751–763.

Mendoza GA, Campbell GE, Rolfe GL. 1986. Multiple objective programming: An approach to planning and evaluation of agroforestry systems—Part 1: model description and development. *Agricultural Systems* **22**: 243–253.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR. 2002a. *Soil and Water Assessment Tool User's Manual*, Version 2000. Available at http://www.brc.tamus.edu/swat/downloads/doc/swatuserman.pdf.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR, King KW. 2002b. *Soil and Water Assessment Tool Theoretical Documentation*, Version 2000. Available at http://www.brc.tamus.edu/swat/downloads/doc/swat2000theory.pdf.

Srinivasan R, Ramanarayanan TS, Arnold JG, Bednarz ST. 1998. Large area hydrologic modeling and assessment. Part II: model application. *Journal of the American Water Resources Association* **34**: 91–101.

White KL, Chaubey I. 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. *Journal of the American Water Resources Association* **41**(5): 1077–1089.

Yan J, Haan CT. 1991. Multiobjective parameter estimation for hydrologic models—multiobjective programming. *Transactions of the Asae* **34**: 848–856.

Yapo PO, Gupta HV, Sorooshian S. 1998. Multi-objective global optimization for hydrologic models. *Journal of Hydrology* **204**: 83–97.