SWAT: Model Use, Calibration, and Validation

J. G. Arnold, D. N. Moriasi, P. W. Gassman, K. C. Abbaspour, M. J. White, R. Srinivasan, C. Santhi, R. D. Harmel, A. van Griensven, M. W. Van Liew, N. Kannan, M. K. Jha

Abstract. *SWAT (Soil and Water Assessment Tool) is a comprehensive, semi-distributed river basin model that requires a large number of input parameters, which complicates model parameterization and calibration. Several calibration techniques have been developed for SWAT, including manual calibration procedures and automated procedures using the shuffled complex evolution method and other common methods. In addition, SWAT-CUP was recently developed and provides a decision-making framework that incorporates a semi-automated approach (SUFI2) using both manual and automated calibration and incorporating sensitivity and uncertainty analysis. In SWAT-CUP, users can manually adjust parameters and ranges iteratively between autocalibration runs. Parameter sensitivity analysis helps focus the calibration and uncertainty analysis and is used to provide statistics for goodness-of-fit. The user interaction or manual component of the SWAT-CUP calibration forces the user to obtain a better understanding of the overall hydrologic processes (e.g., baseflow ratios, ET, sediment sources and sinks, crop yields, and nutrient balances) and of parameter sensitivity. It is important for future calibration developments to spatially account for hydrologic processes; improve model run time efficiency; include the impact of uncertainty in the conceptual model, model parameters, and measured variables used in calibration; and assist users in checking for model errors. When calibrating a physically based model like SWAT, it is important to remember that all model input parameters must be kept within a realistic uncertainty range and that no automatic procedure can substitute for actual physical knowledge of the watershed.*

Keywords. *Autocalibration, Hydrologic model, SWAT, Validation.*

The SWAT (Soil and Water Assessment Tool) model is a continuous-time, semi-distributed, process-based river basin model. It was developed to evaluate the effects of alternative management decisions on water resources and nonpoint-source pollution in large river basins. The first version of SWAT was developed in the early 1990s and released as version 94.2. Engel et al. (1993) reported the first application of SWAT in the peer-reviewed literature; Srinivasan and Arnold (1994) and Arnold et al. (1998) later published the first peer-reviewed description of a geographic information system (GIS) interface for SWAT and overview describing the key components of SWAT, respectively. Arnold and Forher (2005) described the expanding global use of SWAT as well as several subsequent releases of the model: versions 96.1, 98.2, 99.2, and 2000. Gassman et al. (2007) provided further description of SWAT, including SWAT version 2005, and also presented an in-depth overview of over 250 SWAT-related applications that were performed worldwide. Krysanova and Arnold (2008), Douglas-Mankin et al. (2010), and Tuppad et al. (2011) provide further updates on SWAT application and development trends, and the latter two articles provide further description of SWAT version 2009, the latest release of the model.

The development of SWAT is a continuation of USDA Agricultural Research Service (ARS) modeling experience that spans a period of over 30 years (Gassman et al., 2007; Williams et al., 2008). The current SWAT model includes key components contributed from USDA-ARS models as well as from other models (fig. 1). Core pesticide transport, hydrology, and crop growth models that have been incorporated into SWAT can be traced to earlier USDA-ARS field-

Submitted for review in November 2011 as manuscript number SW 9517; approved for publication by the Soil & Water Division of ASABE in May 2012.

The authors are **Jeffrey G. Arnold, ASABE Fellow,** Agricultural Engineer, USDA-ARS Grassland, Soil and Water Research Laboratory, Temple, Texas; **Daniel N. Moriasi, ASABE Member,** Research Hydrologist, USDA-ARS Great Plains Agroclimate and Natural Resources Research Unit, El Reno, Oklahoma; **Philip W. Gassman, ASABE Member,** Associate Scientist, Center for Agricultural and Rural Development, Iowa State University, Ames, Iowa; **Karim C. Abbaspour,** Senior Scientist, Eawag: Swiss Federal Institute for Aquatic Science and Technology, Dübendorf, Switzerland; **Michael J. White, ASABE Member,** Agricultural Engineer, USDA-ARS Grassland, Soil and Water Research Laboratory, Temple, Texas; **Raghavan Srinivasan, ASABE Member,** Director, Spatial Sciences Laboratory, Texas A&M University, College Station, Texas; **Chinnasamy Santhi,** Associate Research Scientist, Texas Research and Extension Center, Temple, Texas; **Daren Harmel, ASABE Member,** Research Leader, USDA-ARS Grassland, Soil and Water Research Laboratory, Temple, Texas; **Ann van Griensven,** Associate Professor, UNESCO-IHE Institute for Water Education, Delft, The Netherlands; **Michael W. Van Liew,** Hydrologic Computer Simulation Modeler, Department of Biological Systems Engineering, University of Nebraska, Lincoln, Nebraska; **Narayanan Kannan, ASABE Member,** Assistant Research Scientist, Texas Research and Extension Center, Temple, Texas; and **Manoj K. Jha, ASABE Member,** Assistant Professor, College of Engineering, North Carolina A&T, Greensboro, North Carolina. **Corresponding author:** Jeffrey G. Arnold, USDA-ARS Grassland, Soil and Water Research Laboratory, 808 East Blackland Road, Temple, TX 76502; phone: 254-770-6502; e-mail: jeff.arnold@ars.usda.gov.

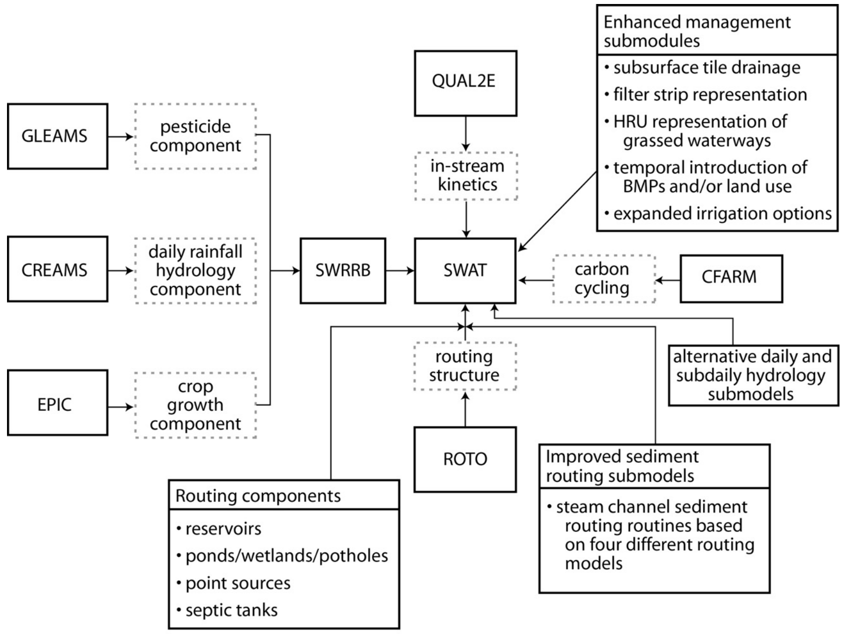

Figure 1. Schematic of SWAT development history and model adaptations (adapted from Gassman et al., 2007).

scale models (fig. 1): the Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) model (Leonard et al., 1987), the Chemicals, Runoff, and Erosion from Agricultural Management Systems (CREAMS) model (Knisel, 1980), and the Environmental Policy Integrated Climate model (Williams et al., 2008; Wang et al., 2011). These components were first grafted into the Simulator for Water Resources in Rural Basins (SWRRB) model (fig. 1; Arnold and Williams, 1987), along with other key components including a weather generator, sediment routing routine, and groundwater submodel (Arnold and Allen, 1999). The initial version of SWAT was then created by interfacing SWRRB with the routing structure in the Routing Outputs to Outlet (ROTO) model (fig. 1; Arnold et al., 1995b). Expanded routing and pollutant transport capabilities have since been incorporated into the model (fig. 1), including reservoir, pond, wetland, point source, and septic tank effects as well as enhanced sediment routing routines (Arnold et al., 2010b) and in-stream kinetic routines from the QUAL2E model (Brown and Barnwell, 1987). Additional modifications that have been incorporated into SWAT (fig. 1) include an improved carbon cycling routine based on the CFARM model (Kemanian, 2011), alternative daily and subdaily hydrology routines including the Green-Ampt infiltration method (Green and Ampt, 1911), temporal accounting of management practice and land use changes and enhanced subsurface tile drainage, filter strips, grassed waterways, irrigation, and other improved representations of conservation and management practices (fig. 1; Arnold et al., 2010b). The temporal accounting routine allows users to introduce the adoption of different selected management practices or account for changes in land use part way through a SWAT simulation run, such as the hydrologic and pollutant impacts simulated by Chiang et al. (2010) in response to temporal changes in pasture use for a 32 km^2 watershed in northwest Arkansas.

The current SWAT2009 code incorporates all of the components shown in figure 1 as well as other routines, and also features several pre- and post-processing software tools, including the widely used ArcGIS SWAT (ArcSWAT) GIS interface (Olivera et al., 2006). Extensive SWAT2009 documentation can also be accessed at the SWAT website (http://swatmodel.tamu.edu), including theoretical documentation describing all equations, a user's manual describing model inputs and outputs, ArcSWAT and Map Window interface manuals, and a developer's manual. In addition to the model documentation, access is also provided at the website to all supporting software, selected journal articles and other publications, a SWAT literature database, previous and forthcoming conference information, forthcoming workshops, SWAT-related job openings, and an email newsletter called *SWATbytes*. The core SWAT development and user support team is located at the USDA-ARS Grassland, Soil and Water Research Laboratory and the Texas AgriLife Blackland Research Center in Temple, Texas. SWAT development is also occurring at other research sites in North America and in other regions (Gassman et al., 2010), and multiple user groups have developed worldwide, including SWAT, ArcSWAT, VizSWAT, SWAT-CUP, Latin American, southeast Asia, Africa, Iran, and Brazil user groups.

Many of the previous studies published in the extensive body of peer-reviewed and other SWAT literature describe calibration and validation approaches used for verifying the accuracy of the model for the simulated conditions. These testing procedures have been reported at varying levels of detail for a wide range of watershed scales, environmental conditions, and application goals worldwide (e.g., Gassman et al., 2007, 2010). More in-depth procedures have also been reported for specific aspects of the calibration and validation process, such as the guidelines proposed by Moriasi et al. (2007) regarding specific statistical criteria to judge the success of SWAT (and other model) testing results. However, a comprehensive overview of all key facets required for an ideal SWAT calibration and validation process is currently lacking in the literature. Thus, the objectives of this study are as follows: (1) to provide a brief description of the key SWAT components, (2) present a general overview of a logical calibration and validation sequence, (3) describe calibration options and parameters in more detail, (4) show how the calibration and validation process is applied for two case studies, and (5) discuss weaknesses and future research needs regarding calibration and validation approaches with SWAT.

Model Description

SWAT operates on a daily time step and is designed to predict the impact of land use and management on water, sediment, and agricultural chemical yields in ungauged watersheds. The model is process based, computationally efficient, and capable of continuous simulation over long time periods. Major model components include weather, hydrology, soil temperature and properties, plant growth, nutrients, pesticides, bacteria and pathogens, and land management. In SWAT, a watershed is divided into multiple subwatersheds, which are then further subdivided into hydrologic response units (HRUs) that consist of homogeneous land use, management, topographical, and soil characteristics. The HRUs are represented as a percentage of the subwatershed area and may not be contiguous or spatially identified within a SWAT simulation. Alternatively, a watershed can be subdivided into only subwatersheds that are characterized by dominant land use, soil type, and management.

Water balance is the driving force behind all the processes in SWAT because it impacts plant growth and the movement of sediments, nutrients, pesticides, and pathogens. Simulation of watershed hydrology is separated into the land phase, which controls the amount of water, sediment, nutrient, and pesticide loadings to the main channel in each subbasin, and the in-stream or routing phase, which is the movement of water, sediments, etc., through the channel network of the watershed to the outlet. Below is a brief description of the processes simulated by SWAT. Details of these processes are given in the SWAT theoretical documentation (http://swatmodel.tamu.edu).

The hydrologic cycle is climate driven and provides moisture and energy inputs, such as daily precipitation, maximum/minimum air temperature, solar radiation, wind speed, and relative humidity, that control the water balance. SWAT can read these observed data directly from files or generate simulated data at runtime from observed monthly statistics. Snow is computed when temperatures are below freezing, and soil temperature is computed because it impacts water movement and the decay rate of residue in the soil. Hydrologic processes simulated by SWAT include canopy storage, surface runoff, infiltration, evapotranspiration, lateral flow, tile drainage, redistribution of water within the soil profile, consumptive use through pumping (if any), return flow, and recharge by seepage from surface water bodies, ponds, and tributary channels. SWAT uses a single plant growth model to simulate all types of land cover and differentiates between annual and perennial plants. The plant growth model is used to assess removal of water and nutrients from the root zone, transpiration, and biomass/yield production. SWAT uses the Modified Universal Soil Loss Equation (MUSLE) (Williams and Berndt, 1977) to predict sediment yield from the landscape. In addition, SWAT models the movement and transformation of several forms of nitrogen and phosphorus, pesticides, and sediment in the watershed. SWAT allows the user to define management practices taking place in every HRU.

Once the loadings of water, sediment, nutrients, and pesticides from the land phase to the main channel have been determined, the loadings are routed through the streams and reservoirs within the watershed. The water balance for reservoirs includes inflow, outflow, rainfall on the surface, evaporation, seepage from the reservoir bottom, and diversions.

Model equations are given in the SWAT theoretical documentation (http://swatmodel.tamu.edu) and in Arnold et al. (1998). Gassman et al. (2007) presents an overview of: (1) climatic inputs and HRU hydrologic balance; (2) cropping, management inputs, and HRU-level pollutant losses; and (3) flow and pollutant routing. Arnold et al. (2010b) describe current research on enhancements to SWAT to route water across discretized landscape units that simulate the impacts of spatial land use changes and land management on the hillslope-valley continuum.

SWAT Calibration and Validation

SWAT input parameters are process based and must be held within a realistic uncertainty range. The first step in the calibration and validation process in SWAT is the determination of the most sensitive parameters for a given watershed or subwatershed. The user determines which variables to adjust based on expert judgment or on sensitivity analysis. Sensitivity analysis is the process of determining the rate of change in model output with respect to changes in model inputs (parameters). It is necessary to identify key parameters and the parameter precision required for calibration (Ma et al., 2000). In a practical sense, this first step helps determine the predominant processes for the component of interest. Two types of sensitivity analysis are generally performed: local, by changing values one at a time, and global, by allowing all parameter values to change. The two analyses, however, may yield different results. Sensitivity of one parameter often depends on the value of other related parameters; hence, the problem with one-at-a-time analysis is that the correct values of other parameters that are fixed are never known. The disadvantage of the global sensitivity analysis is that it needs a large number of simulations. Both procedures, however, provide insight into the sensitivity of the parameters and are necessary steps in model calibration.

The second step is the calibration process. Calibration is an effort to better parameterize a model to a given set of local conditions, thereby reducing the prediction uncertainty. Model calibration is performed by carefully selecting values for model input parameters (within their respective uncertainty ranges) by comparing model predictions (output) for a given set of assumed conditions with observed data for the same conditions. The final step is validation for the component of interest (streamflow, sediment yields, etc.). Model validation is the process of demonstrating that a given site-specific model is capable of making sufficiently accurate simulations, although "sufficiently accurate" can vary based on project goals (Refsgaard, 1997). Validation involves running a model using parameters that were determined during the calibration process, and comparing the predictions to observed data not used in the calibration. In general, a good model calibration and validation should involve: (1) observed data that include wet, average, and dry years (Gan et al., 1997); (2) multiple evaluation techniques (ASCE, 1993; Legates and McCabe, 1999; Boyle et al., 2000); (3) calibrating all constituents to be evaluated; and (4) verification that other important model outputs are reasonable. In general, graphical and statistical methods with some form of objective statistical criteria are used to determine when the model has been calibrated and validated. Calibration can be accomplished manually or using autocalibration tools in SWAT (van Griensven and Bauwens, 2003; Van Liew et al. (2005) or SWAT-CUP (Abbaspour et al., 2007).

Ideally, calibration and validation should be process and spatially based, while taking into account input, model, and parameter uncertainties. A good example of process-based calibration involves streamflow. Streamflow processes are comprised of the water balance in the land phase of the hydrology, including ET, lateral flow, surface runoff, return flow, tile flow (if present), channel transmission losses, and deep aquifer recharge. If data are available for each of these processes, they should be calibrated individually. For sediments, nutrients, pesticides, and bacteria, sources and sinks should be considered. If a longer time period is available for hydrology than water quality data, it is important to use all the hydrology data available for calibration and validation to capture long-term trends. This process-based calibration should be done at the subwatershed or landscape level to ensure that variability in the predominant processes for each of the subwatersheds is captured instead of determining global (watershed-wide) processes. There are, however, generally insufficient observed data to enable a full spatial calibration and validation at the watershed scale. The metrics and methods used to compare observed data to model predictions are also important. Multiple graphical and statistical methods could be used, such as time-series plots, Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970), and percent bias.

A general calibration flowchart for flow, sediment, and nutrients is shown in figure 2 to aid with the manual model calibration process. Users should check the water balance components (ET, surface/baseflow ratios, tile flow proportions, plant yield, and biomass) during the calibration process to make sure the predictions are reasonable for the study region or watershed. Because plant growth and biomass production can have an effect on the water balance, erosion, and nutrient yields, reasonable local/regional plant growth days and biomass production should be verified during model calibration to the extent possible. Examples of SWAT crop growth and/or yield calibrations are reported by Hu et al. (2007), Ng et al. (2010a), Andersson et al. (2011), and Nair et al. (2011). Several studies (e.g., Santhi et al., 2001; Engel et al., 2007) also recommend that streamflow, sediment, and nutrient transport be calibrated sequentially (in that order) because of interdependencies between constituents due to shared transport processes. Even though a complete set of hydrologic and water quality data are rarely available, all available data should be considered. We recommend that baseflow and surface runoff be separated from the observed total daily streamflow using a baseflow filter. The baseflow filter developed by Arnold et al. (1995a) and modified by Arnold and Allen (1999) is available at http://swatmodel.tamu.edu/software/baseflow-filter-program. Baseflow and recharge data from this procedure have shown good correlation with those produced by SWAT (Arnold et al., 2000). To help with the recommended SWAT calibration and validation process, a program was recently developed by White et al. (2012). This program warns users if selected model outputs vary outside typical ranges to ensure that processes are realistically simulated (http://swatmodel.tamu.edu/software/swat-check). Following these simple recommendations can avoid major errors in scenario analysis, which is the primary objective of most SWAT-related projects.

Calibration and validation are typically performed by splitting the available observed data into two datasets: one for calibration, and another for validation. Data are most frequently split by time periods, carefully ensuring that the climate data used for both calibration and validation are not substantially different, i.e., wet, moderate, and dry years occur in both periods (Gan et al., 1997). Data may also be split spatially, with all available data at a given monitoring location assigned to the calibration phase and correspondingly performing the validation at one or more other gauges within the watershed. This approach can be necessary when users are faced with data-limited situations that preclude performing a split-time calibration and validation using a single gauge. SWAT users have also used calibrated parameters from a watershed with approximately similar climatic, soils, and land use conditions for validation in their study watershed, or vice versa. Split-location calibration and validation approaches have been performed in some previous SWAT studies (e.g., Arnold et al., 2001; Van Liew and Garbrecht, 2003; Cao et al., 2006; Parajuli et al., 2009).

Although these are the recommended calibration and validation approaches, they are not enforced, and thus there are several ways in which SWAT has been calibrated and validated. Most published SWAT applications report both graphical and statistical hydrologic calibration results, especially for streamflow, and hydrologic validation results are also reported for a large percentage of the studies. Similar pollutant testing results are also reported for many SWAT studies, although not nearly as many as are reported for streamflow results. An extensive array of statistical techniques can be used to evaluate SWAT hydrologic and pollutant predictions; for example, Coffey et al. (2004) describe nearly 20 potential statistical tests that can be used to judge SWAT predictions, including coefficient of determination (r^2), NSE, root mean square error (RMSE), nonpar-

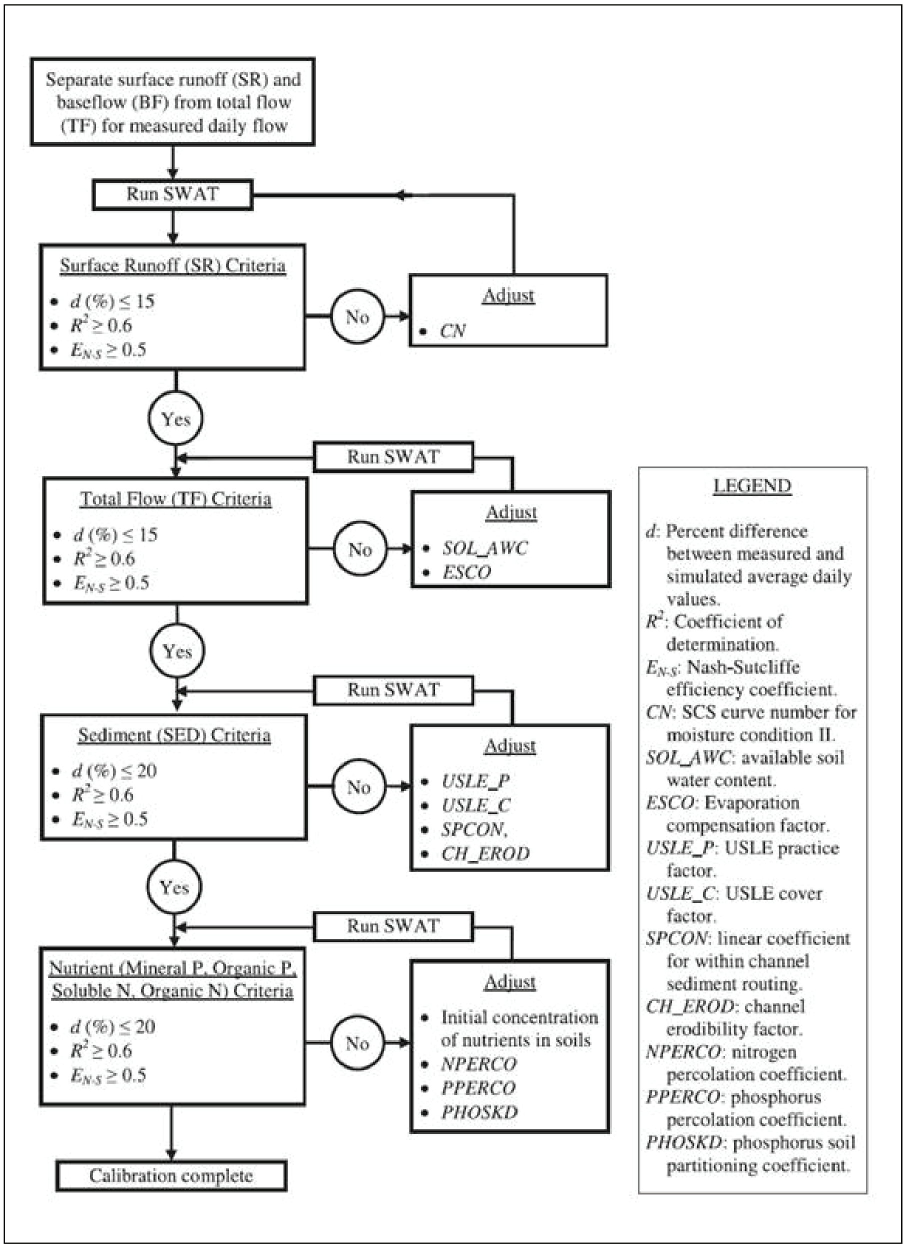

Figure 2. Example SWAT manual calibration flowchart (from Engel et al., 2007; adapted from Santhi et al., 2001).

ametric tests, t-test, objective functions, autocorrelation, and cross-correlation. By far, the most widely used statistics reported for calibration and validation are r^2 and NSE. The r^2 statistic can range from 0 to 1, where 0 indicates no correlation and 1 represents perfect correlation, and it provides an estimate of how well the variance of observed values are replicated by the model predictions(Krause et al., 2005). A perfect fit also requires that the regression slope and intercept are equal to 1 and 0, respectively; however, the slope and intercept have typically not been reported in published SWAT studies. If r^2 is the primary statistical measure, it should always be used with slope and intercept to ensure that means are reasonable (slope = 1) and bias is low. NSE values can range between -∞ to 1 and provide a measure how well the simulated output matches the observed data along a 1:1 line (regression line with slope equal to 1). A perfect fit between the simulated and observed data is indicated by an NSE value of 1. NSE values ≤0 indicate that the observed data mean is a more accurate predictor than the simulated output. Both NSE and r^2 are biased toward high flows. To minimize this bias, some researchers have taken the log of flows for statistical comparison or have developed statistics for low and high flow seasons (Krause et al., 2005). Krause et al. (2005) provide further discussion regarding the strengths and weaknesses of using r^2, NSE, and other efficiency criteria measures.

An extensive list of r^2 and NSE calibration and/or validation statistics is provided by Gassman et al. (2007) for 115 SWAT studies that reported hydrologic results as well as 37 SWAT studies that reported pollutant results. Similar r^2 and NSE statistical compilations for an additional 20 SWAT studies are reported by Douglas-Mankin et al. (2010), and 23 SWAT studies are reviewed by Tuppad et al. (2011). These statistics provides valuable insight regarding the hydrologic performance of the model across a wide spectrum of conditions. To date, no absolute criteria for judging model performance have been firmly established in the literature, and for good reason: the criteria for judgment of model performance should be tied to the intended use of the model (Engel et al., 2007). SWAT has been used for a variety of applications, ranging from simple hydrologic assessments for watershed planning to the assignment of blame and damages in a court of law (White et al., 2011). The risk of adverse impacts arising from model prediction uncertainty or error for a particular application should be a consideration during the calibration. However, for a more typical application, Moriasi et al. (2007) proposed that NSE values should exceed 0.5 in order for model results to be judged satisfactory for hydrologic and pollutant loss evaluations performed on a monthly time step (and that appropriate relaxing and tightening of the standard be performed for daily and annual time step evaluations, respectively). Assuming this criterion for both the NSE and r^2 values at all time steps, the majority of the calibration and validation statistics listed by Gassman et al. (2007) were judged satisfactory, or adequately replicating observed streamflows and other hydrologic indicators. However, it is clear that poor test statistics occurred for parts or all of some studies. The poorest results generally occurred for daily predictions, although this was not universal (e.g., Grizzetti et al., 2005). Interestingly, Douglas-Mankin et al. (2010) found that all of the daily flow calibration statistics reported among the 20 SWAT studies they reviewed were satisfactory or better, based on the 0.5 criterion described above. Tuppad et al. (2011) further found that 85% of the daily flow statistics compiled from their review of 23 SWAT studies also met this criterion. When they combined their overall compiled statistics with those reported by Gassman et al. (2007) and Douglas-Mankin et al. (2010), Tuppad et al. (2011) also reported that 72% of 134 calibration results and 58% of 113 validation results were rated as satisfactory or better, and 21% of calibration results and 12% of validation results were rated as very good, where "very good" meant NSE (and r^2) >0.75, again based on monthly flow criteria proposed by Moriasi et al., 2007.

Some of the poorer testing results reported in previous SWAT studies can be partially attributed to inadequate spatial coverage of precipitation inputs, which can occur because of an inadequate number of rain gauges in the simulated watershed or coarse subwatershed configurations that failed to capture the spatial detail of available rainfall data (e.g., Cao et al., 2006; Conan et al., 2003; Bouraoui et al., 2002, 2005). Inadequate model calibration (Bosch et al., 2004), measurement uncertainty in streamflow data (Harmel et al., 2006a, 2009), and short streamflow records (Muleta and Nicklow, 2005) can also result in weak SWAT hydrologic predictions. Most reported SWAT studies contain both calibration and validation, while others performed only calibration due to a lack of observed data. In a few cases, calibration of SWAT was not performed. For example, Srinivasan et al. (2010) describe an uncalibrated application of SWAT for the Upper Mississippi River basin in the north-central U.S., which was conducted with the goal of determining how the default parameters represented crop yield and streamflow components of interest in the region.

Accounting for Uncertainty Bands

The above statistical indices only apply to the comparison of two signals and are not adequate when outputs are expressed as uncertainty bands. In this case, as the simulation results are usually expressed by the 95% prediction uncertainties (95PPU), they cannot be compared with the observation signals using the traditional r^2 and NSE statistics. For this reason, Abbaspour et al. (2004, 2007) suggest using two measures, referred to as the P-factor and the R-factor. The P-factor is the percentage of the measured data bracketed by the 95PPU. This index provides a measure of the model's ability to capture uncertainties. As all the "true" processes are reflected in the measurements, the degree to which the 95PPU does not bracket the measured data indicates the prediction error. Ideally, the P-factor should have a value of 1, indicating 100% bracketing of the measured data, hence capturing or accounting for all the correct processes. The R-factor, on the other hand, is a measure of the quality of the calibration and indicates the thickness of the 95PPU. Its value should ideally be near zero, hence coinciding with the measured data. The combination of P-factor and R-factor together indicate the strength of the model calibration and uncertainty assessment, as these are intimately linked.

Calibration Approaches

Conventionally, calibration is performed manually and consists of changing model input parameter values to produce simulated values that are within a certain range of the measured data (Balascio et al., 1998). However, when the number of parameters used in the manual calibration is large, especially for complex hydrologic models, manual calibration can become labor-intensive (Balascio et al., 1998) and automated calibration methods are preferred. Both manual algorithms and automated methods have been developed for calibration of SWAT simulations. An iterative approach is usually used for manual calibration involving the following steps: (1) perform the simulation; (2) compare measured and simulated values; (3) assess if reasonable results have been obtained; (4) if not, adjust input parameters based on expert judgment and other guidance within reasonable parameter value ranges; and (5) repeat the process until it is determined that the best results have been obtained. Several studies present systematic strategies for performing streamflow and/or pollutant calibration and validation. Coffey et al. (2004) recommend using NSE and r^2 for analyzing monthly output and median objective functions, sign test, autocorrelation, and cross-correlation for assessing daily output based on comparisons of SWAT results with measured streamflow. Santhi et al. (2001) propose a manual calibration approach (including a flowchart) that was a function of sensitive input parameters (15 were selected), realistic uncertainty ranges, and satisfactory r^2 and NSE statistical results. Cao et al. (2006) also present a flowchart of their SWAT manual calibration approach, which was based on multiple hydrologic outputs and multiple gauge sites. Nair et al. (2011) present another systematic SWAT calibration approach, including a schematic of the methodology that incorporates crop growth/yield calibration and validation along with streamflow and pollutant loss calibration and validation. Many studies also report the use of automated techniques for calibrating SWAT, which typically rely on Monte Carlo or other sampling schemes to estimate the best choice of values for multiple input parameters, without violating practical or theoretical boundaries for each specific input parameter. The input values are usually determined over the course of iterative SWAT simulations, which sometimes number in the thousands. Several optimization schemes have been used in SWAT autocalibration applications, including generalized likelihood uncertainty estimation (GLUE), shuffled complex evolution (SCE), and the Parameter Estimation (PEST) program (Doherty, 2004). Govender and Everson (2005) and Wang and Melesse (2005) used PEST to calibrate key hydrology-related parameters for SWAT applications in South Africa and northwest Minnesota, respectively. Wang and Melesse (2005) also found that manual calibration resulted in more accurate predictions than the automated PEST approach. Ng et al. (2010b) described advantages and disadvantages of using PEST versus the GLUE method for calibrating SWAT for a watershed in central Illinois. Setegn et al. (2009) and Razavi et al. (2010) described additional SWAT calibration studies that relied partially on GLUE methodology for watersheds located in Ethiopia and south-central New York. To determine optimum input parameters based on the global objective criteria for a simulation of a river basin in Belgium, van Griensven and Bauwens (2003, 2005) incorporated an SCE module directly into the SWAT code. Calibration parameters, corresponding parameter ranges, and measured daily streamflow and pollutant data were input for the application, which required several thousand SWAT simulations for completion. Eckhardt and Arnold (2001) and Eckhardt et al. (2005) used similar SCE-based automatic calibration methods for SWAT simulations of German watersheds. Other SCE-based SWAT automatic calibration applications are reported by Di Luzio and Arnold (2004) and Van Liew et al. (2005, 2007).

Automatic calibration and uncertainty analysis capability is now directly incorporated in SWAT2009 (Gassman et al., 2010) via the SWAT-CUP software developed by Eawag (2009). A number of previous SWAT application projects report automated calibration/validation and uncertainty analysis using SWAT-CUP. Abbaspour et al. (2007) performed a multi-objective calibration and validation of the Thur watershed in Switzerland using discharge, sediment, nitrate, and phosphate in the objective function with uncertainty analysis. Schuol et al. (2008a, 2008b) calibrated with uncertainty analysis and validated models of west Africa and the entire continent of Africa. Yang et al. (2008) compared five different optimization algorithms in SWAT-CUP and calibrated a watershed in China (2007) using the MCMC algorithm. Faramarzi et al. (2009) used SWAT to build a hydrologic model of Iran and calibrated and validated it with the SUFI2 algorithm accounting for prediction uncertainty. Akhavan et al. (2010) calibrated a model of nitrate leaching for a watershed in Iran, and Andersson et al. (2009) used SWAT-CUP to calibrate a hydrologic model of the Thukela River basin in South Africa. In the above applications, the goodness of fit criteria is provided by P-factor and R-factor. For the objective function, however, a weighted version of r^2 (Krause et al., 2005) was selected as the efficiency criterion:

$$\Phi = \begin{cases} |b|\ R^2 & \text{if } b \le 1 \\ |b|^{-1}\ R^2 & \text{if } b > 1 \end{cases} \quad (1)$$

where b is the slope of the regression line between measured and simulated signals. A major advantage of this efficiency criterion is that it ranges from 0 to 1, which compared to NSE with a range of -∞ to 1, ensures that in a multisite calibration the objective function is not governed by a single or a few badly simulated stations.

Calibration Parameters

Numerous studies have reported input parameters used in SWAT model calibration. Table 1 summarizes the parameters used in 64 studies and in studies previously summarized by Douglas-Mankin et al. (2010) and Tuppad et al. (2011). All of these studies include detailed reporting of model parameterization and calibration procedures, including tables with parameter ranges and/or final values. Many publications in the literature (https://www.card.iastate.edu/swat_articles) do not adequately report changes in parameters. Model parameters used in calibration studies and even in the selected publications exhibited gaps. Tuppad et al. (2011) re-emphasizes an important point made by Douglas-Mankin et al. (2010): "Improved reporting of calibration and validation procedures and results, perhaps guided by a set of standard reporting guidelines, is essential for adequate interpretation of each study and comparison among studies in the future. This increased information would also form the basis for assigning typical parameters and ranges for use in both manual or automatic calibration and uncertainty processes."

Table 1 categorizes parameters by process. Since SWAT is a comprehensive model that simulates process interactions, many parameters will impact multiple processes. For example, CN directly impacts surface runoff; however, as surface runoff changes, all components of hydrology balance change. Soil erosion and nutrient transport are also directly impacted by surface runoff, as are plant growth and nutrient cycling. This is the primary reason why most manual calibration methods start with the hydrology balance and streamflow, then move to sediment, and finally calibrate nutrients and pesticides, as shown in figure 2 (Santhi et al., 2001). It is evident from table 1 that hydrology is calibrated in most studies, with CN2, AWC, ESCO, and SURLAG used routinely. The baseflow process is also often calibrated with the baseflow recession parameters used in many studies.

Parameters for sediment calibration are used less often due to inadequate reporting or studies that did not focus on sediment. It is interesting to note that of the studies in which sediment was calibrated, channel parameters were used more often than parameters affecting sediment

Table 1. Calibration parameters reported in 64 selected SWAT watershed studies.[a] Numbers in parentheses are the number of times the parameter was used in calibration. Definitions of variables are found in the SWAT user manual (http://swatmodel.tamu.edu/documentation).

Process	Input Parameters									
Surface runoff	CN2 (36)	AWC (28)	ESCO (23)	EPCO (10)	SURLAG (22)	OV_N (8)				
Baseflow	GW_ALPHA (28)	GW_REVAP (18)	GW_DELAP (21)	GW_QWN (12)	REVAPMN (13)	RCHARG_DP (14)				
Snow	SFTMP (11)	SMFMN (14)	SMFMX (18)	SMTMP (13)	TIMP (7)	SNO50COV (4)	SNOCOVMX (3)			
Sediment from channels	PRF (10)	APM (7)	SPEXP (10)	SPCON (11)	CH_EROD (6)	CH_COV (7)				
Sediment from landscape	USLE_P (7)	USLE_C (7)	USLE_K (7)	LAT_SED (1)	SLSOIL (2)	SLOPE (8)				
N from landscape	RCN (1)	UBN (3)	GWNO3 (2)	ERORGN (5)	NPERCO (11)	ANION_EXCL (2)				
P from landscape	PSP (5)	PHOSKD (6)	UBP (5)	PPERCO (8)	GWQMINP (1)	ERORGP (5)				
Pesticides	KOC (1)	HL_SOIL (1)	HL_FOL (1)	WSOL (1)	WOFFW (2)					
Subsurface tile	TDRAIN (1)	GDRAIN (2)	DEP_IMP (1)							
N and P from channels	BC1 (2)	BC2 (2)	BC3 (2)	BC4 (2)	RS4 (2)	RS5 (1)				
Plant growth	GSI (3)	HI (1)	BLAI (3)	PHU (1)	CN_YLD (1)					
Bacteria	BACTRDQ (1)	BACTMIX (1)	BCNST (1)	CFRT_KG (1)	WDPRCH (1)	WDPQ (1)				
Other	BIOMIX (4)	SOL_ROCK (1)	MSK_COL (1)	MSK_CO2 (2)	CBNINT (1)	SOL_BD (3)	ALPHA_BNR (1)	EVRCH (1)	SOL_ALB (2)	LAT_TTIME (1)

[a] Abbaspour et al. (2007), Ahl et al. (2008), Alibuyog et al. (2009), Behera and Panda (2006), Bekele and Nicklow (2007), Benaman and Shoemaker (2004), Benaman et al. (2005) , Bekele and Knapp (2010), Cheng et al. (2007), Chin et al. (2009), Chu et al. (2004), Coffey et al. (2010), Debele et al.(2008), Di Luzio and Arnold (2004), Douglas-Mankin et al.(2010), Du et al.(2006), Easton et al.(2008), Eckhardt et al.(2002), Eckhardt et al. (2005), Engel et al. (2007), Ghaffari et al. (2010), Gikas et al. (2006), Gitau et al. (2008), Green et al. (2006, 2007), Green and van Griensven (2008), Heuvelmans et al. (2004, 2006), Hu et al. (2007), Inamdar and Naumov (2006), Jha et al. (2010), Lemonds and McCray (2007), Maski et al. (2008), Meng et al. (2010), Mukundan et al. (2010), Muleta and Nicklow (2005), Narasimhan et al. (2010), Santhi et al. (2008), Shoemaker et al. (2007), Starks and Moriasi (2009), Sui and Frankenberger (2008), Sudheer et al. (2007), Tuppad et al. (2011),van Griensven et al. (2008), Van Liew et al. (2007), White and Chubey (2005), Zhang et al. (2008a), Zhang et al. (2008b).

transport from the landscape. A potential explanation is that there is more uncertainty in the channel sediment routing parameters, and thus users feel more comfortable using them in calibration. These parameters can also be very sensitive, making adjustment very effective during the calibration process. However, it is critical to ensure that sources and sinks of sediment and that the ratio of upland sources versus channel sources and deposition are realistic, even though measured data are often relatively scarce. Sediment measurement is very difficult and involves considerable error. Furthermore, the modified Universal Soil Loss Equation in SWAT is inadequate in many cases, such as in accounting for the "second storm effect" reported by Abbaspour et al. (2007). Therefore, adjustment of the parameters is actually compensating for the lack of precision in the measurement or errors in the conceptual model.

Based on the literature review, it is also evident that many processes are not as rigorously calibrated as hydrology and streamflow. Only a few studies adequately reported calibration parameters for N, P, pesticides, bacteria, tile flow, and plant growth. Standard reporting guidelines would help form the basis for assigning parameters and ranges for these processes.

CASE STUDIES

Two case studies were chosen as examples of SWAT validation. The first study by Van Liew et al. (2005) highlights the advantages of manual and automated calibration techniques, and the second study by Rouholahnejad et al. (2012a) uses the semi-automated SUFI2 program. Both case studies emphasize that no automatic calibration procedure can substitute for actual physical knowledge of watershed processes.

MANUAL AND AUTOMATED CALIBRATION

The calibration study by Van Liew et al. (2005) compared and discussed both manual and automated calibration techniques for five watersheds at the ARS Little River experimental watersheds at Tifton, Georgia (fig. 3) and at the Little Washita experimental watersheds operated by ARS scientists at El Reno, Oklahoma (fig. 4). The locations represent a wide range of climate, soils, and land use (table 2).

Manual Calibration

SWAT was calibrated manually by following a multistep procedure recommended by Neitsch et al. (2002). For the Little Washita, the upper watershed (subwatershed 526) was calibrated first, and the parameters in that subwatershed were then held constant while the larger watershed (subwatershed 550) was calibrated on that portion of the subwatershed below the outlet of 526. Although it was recognized that computational differences between measured and simulated streamflow at the outlet of subwatershed 526

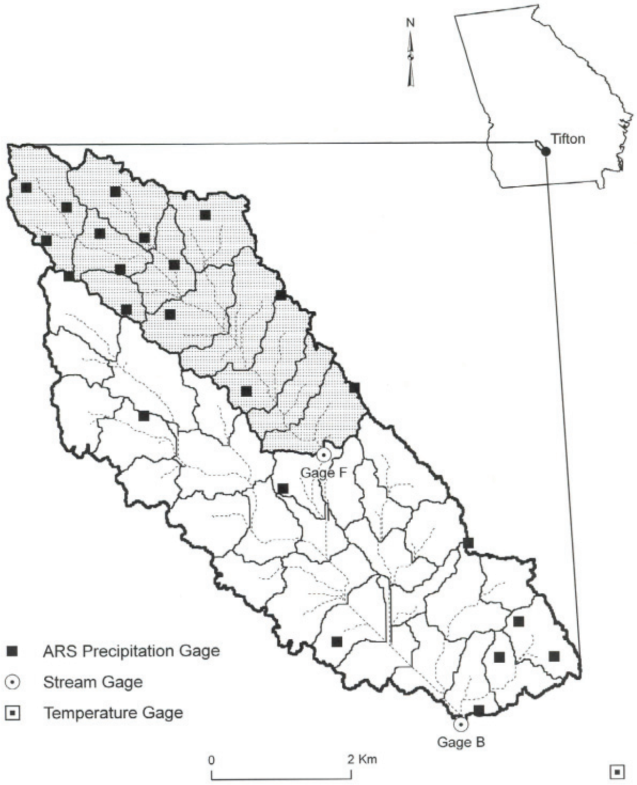

Figure 3. Location of the Little River Experimental Watershed in Georgia (from Van Liew et al., 2005).

were passed on to subwatershed 550, this approach to calibration was the most reasonable option that could be exercised, based on the availability and quality of existing datasets in the watershed (Van Liew et al., 2005). A similar approach was taken in calibrating the Little River. The upper portion of the watershed (subwatershed F) was calibrated first, which was then followed by a calibration of subwatershed B. Manual calibration attempted to minimize total flow (minimized average annual percent bias), accompanied by visual inspection of daily hydrographs and duration of daily flow curves. The sum of squares of residuals objective function could have been used in the manual calibration, but preliminary testing showed that the total mass balance method in combination with the inspection of duration of daily flow curves gave a better representation of the range of simulated flows (Van Liew et al., 2005). Detailed calibration of SWAT on the Little Washita was previously reported by Van Liew and Garbrecht (2003). A preliminary calibration was conducted on a monthly basis to identify the order of magnitude of all parameters to reproduce proper runoff volumes and seasonal characteristics. The runoff curve number (CN2) that governs the surface runoff response was first calibrated. Second, the groundwater "revap" coefficient (GW_REVAP), the threshold depth of water in the shallow aquifer required for return flow (REVAPMN), and the deep aquifer percolation fraction (RCHRG_DP), which governs the fraction of percolation from the root zone to the deep aquifer, were calibrated. Third, the baseflow recession factor (ALPHA_BF) and the groundwater delay (GW_DELAY) parameters were calibrated so that the monthly measured versus simulated hydrographs agreed well (Van Liew et al., 2005). This preliminary calibration was followed by a fine-tuning at the daily time scale so that the predicted versus measured peak flows and recession curves on a daily time step matched as closely as possible. This same approach was taken in the manual calibration of the Little River.

Autocalibration

The autocalibration procedure described by Van Liew et al. (2005) was developed by van Griensven and Bauwens (2003) and is based on the shuffled complex evolution algorithm (SCE-UA; Duan et al., 1992) that allows for the

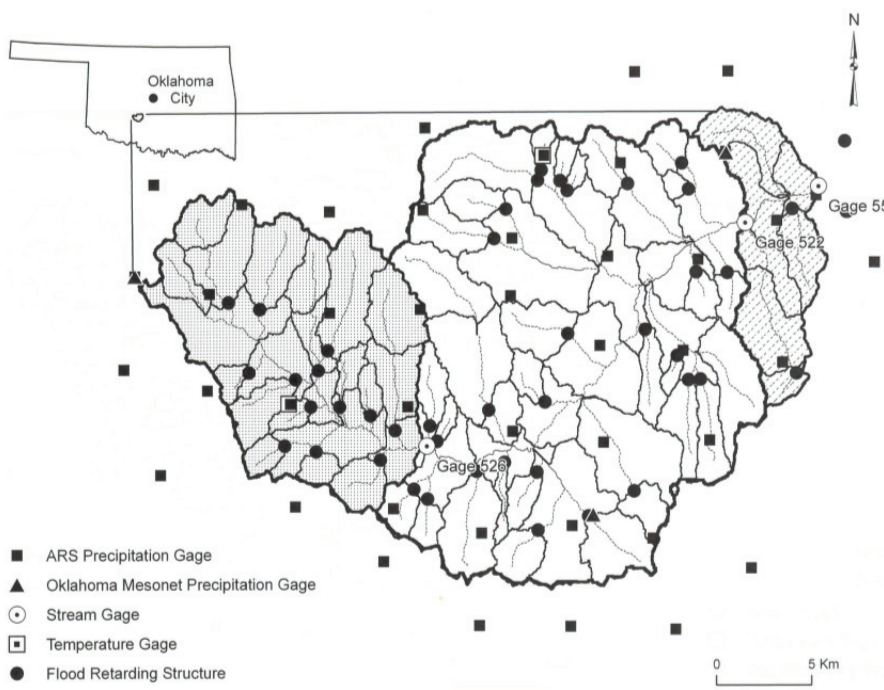

Figure 4. Location of the Little Washita River Experimental Watershed in Oklahoma (from Van Liew et al., 2005).

Table 2. Number of subbasins, number of hydrologic response units, drainage areas, and land use types for the two USDA ARS experimental watersheds.

Watershed	No. of Subbasins	No. of HRUs	Drainage Area (km^2)	Land Use Type (%)				
				Range/Pasture	Crop	Forest	Wetland	Misc.
Little River F	12	51	114	19	45	26	9	1
Little River B	40	161	330	10	42	45	2	1
Little Washita 526	22	138	160	59	28	6	0	7
Little Washita 550	73	486	600	66	19	9	0	6
Little Washita 522	66	413	538	66	18	9	0	7

calibration of model parameters based on a single objective function. The SCE-UA has been widely used in watershed model calibration and other areas of hydrology, such as soil erosion, subsurface hydrology, remote sensing, and land surface modeling, and has generally been found to be robust, effective, and efficient (Duan, 2003).

Parameters in SWAT were calibrated at the daily time scale in a distributed fashion using the automated calibration procedure, in which observed and simulated outputs were compared at the same outlet points as the manual calibration. With the completion of a given optimization, two sets of calibrated parameters were computed for the Little River that corresponded to subwatersheds F and B, and two sets were computed for the Little Washita that corresponded to subwatersheds 526 and 550. Default values suggested by van Griensven (2002) were selected as initial upper and lower ranges for the respective model parameters. Minimizing the sum of squares residuals was used as the objective function in the autocalibration procedure.

The Van Liew et al. (2005) study highlighted an important difference that must be realized in comparing the manual versus autocalibration approaches. The autocalibration approach was strictly a quantitative comparison that involved minimizing the difference between measured and simulated values. The manual approach involved both quantitative and qualitative comparisons, since it involved using the total mass controller in conjunction with graphical comparisons of monthly and daily hydrographs and duration of daily flow curves to calibrate the model against measured data. Use of the manual calibration accentuates the tradeoffs that exist in achieving total mass balance, reasonable hydrograph responses, and adequate representation of the range in flows. Van Liew et al. (2005) suggest that the strengths of both the manual and autocalibration approaches can be used to facilitate the calibration process. With proper selection of the upper and lower ranges for model parameter values, autocalibration can provide an initial parameter set with minimal labor on the part of the user. Depending on the particular modeling needs, a manual approach can then be taken to refine the calibration, so that an appropriate balance is achieved regarding the amount, timing, and distribution of the output variable.

Results of the Van Liew et al. (2005) suggest that the autocalibration option in SWAT provides a powerful, labor-saving tool that can be used to substantially reduce the frustration and subjectivity that often characterize manual calibrations. If used in combination with a manual approach, the autocalibration tool shows promising results in providing initial estimates for model parameters. To maintain mass balance and adequately represent the range in magnitude of output variables, manual adjustments may be necessary after autocalibration. Caution must also be exercised in using the autocalibration tool so that the selection of initial lower and upper ranges in the parameters results in calibrated values that are representative of watershed conditions (Van Liew et al., 2005).

Semi-Automated SUFI2

For the second case study, an example calibration of the Danube project (Rouholahnejad et al., 2012b) was selected using SWAT-CUP. Rouholahnejad et al. (2012a) referred to the process of parameter assignment as parameterization. Correct parameterization is an important step in model calibration and must be based on the knowledge of the hydrologic processes and variability in soil, land use, slope, and location as defined by the subbasin number. Parameterization, therefore, could be defined as "the process of imparting the analyst's knowledge of the physical processes of the watershed to the model." No automatic calibration procedure can substitute for actual physical knowledge of the watershed, which can translate into correct parameter ranges for different parts of the watershed. These ranges can effectively guide the optimization routine. Hence, correct parameterization can result in faster and more accurate model calibration with smaller prediction uncertainty. SWAT-CUP includes automated as well as semi-automatic procedures for model calibration. The following steps are suggested in a calibration exercise with the semi-automated program SUFI2:

1. Develop initial or default SWAT input parameters (as created by ArcSWAT or other GIS interfaces) and prepare the input files for SWAT-CUP.
2. Run the model with initial parameters and plot the simulated and observed variables at each gauging station for the entire period of record.
3. Based on step 2, divide the entire period into calibration and validation periods while attempting to ensure that both periods have a similar number of wet and dry years and similar average water balances.
4. Determine the most sensitive parameters for the observed values of interest. This information can usually be deduced from the literature (see table 1).
5. Assign an initial uncertain range (typically 20% to 30%) to each parameter globally, meaning scaling the parameters identically for each HRU.
6. Run the SWAT-CUP-SUFI2 model 300 to 500 times and view the results for each gauged outlet, as shown in figure 5.
7. Perform the global sensitivity analysis and view the results. At this stage, the P-factor and p-value

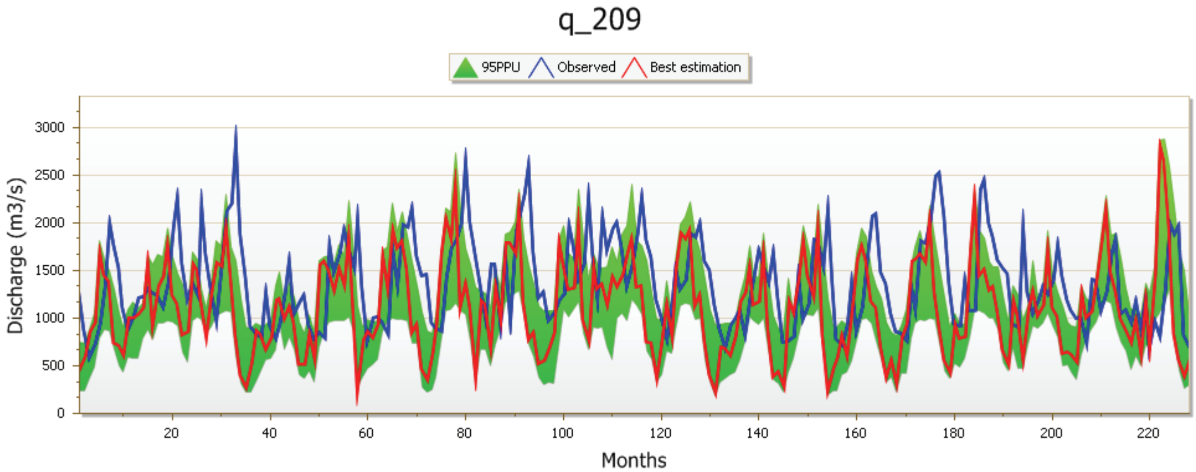

Figure 5. Observed flow, 95% model uncertainty, and best estimation at gauging station 209 before calibration.

t-statistic can be used to eliminate non-sensitive parameters from the calibration process.

8. After observing model performance in step 6, regionalize the respective parameters. For example, as shown in figure 5, the model systematically underestimated baseflow at outlet q_209 (in subbasin 209), and there is an early shift in the flow peak. To increase baseflow, decrease deep percolation (GWQMN), decrease the groundwater revap coefficient (GW_REVAP), and increase the threshold depth of water in shallow aquifer (REVAPMN). To correct the early shift, decrease the slope (HRU_SLP), increase Manning's roughness coefficient (OV_N), increase the value of overland flow rate (SLSUBBSN), and increase snow melt parameters (SMTMP).

To increase baseflow and delay peaks, identify the subbasins that contribute to the outlet at subbasin 209 and implement the changes in the respective parameters. For example, make changes in the parameters only in subbasins draining into 209, and set new ranges for the parameters using one-at-a-time sensitivity analysis implemented in SWAT-CUP. The parameters must always be kept within realistic ranges, as influenced by the uncertainty in defining the parameter.

In the manner described above, the parameters of each observational gauge can be used to spatially calibrate the model in the drainage area between the gauges. At this point, the analyst's knowledge of processes in the watershed could also be implemented in the optimization.

Figure 6 shows the results after implementing the above changes in the parameters and running the model, where NS increased from -1.5 to 0.2. Additional iterations can further improve the results. Details on parameterization and results can be found in Rouholahnejad et al. (2012a, 2012b). After calibration, the model should be run for the validation period to assess its performance.

Discussion

Gassman et al. (2010) discussed trends in SWAT use and the technical and networking factors that are regarded as strengths of the model, which include: web-based documentation, user support groups, SWAT literature database, GIS interface tools, pre- and post-processing tools, open source code, regional and international conferences, and

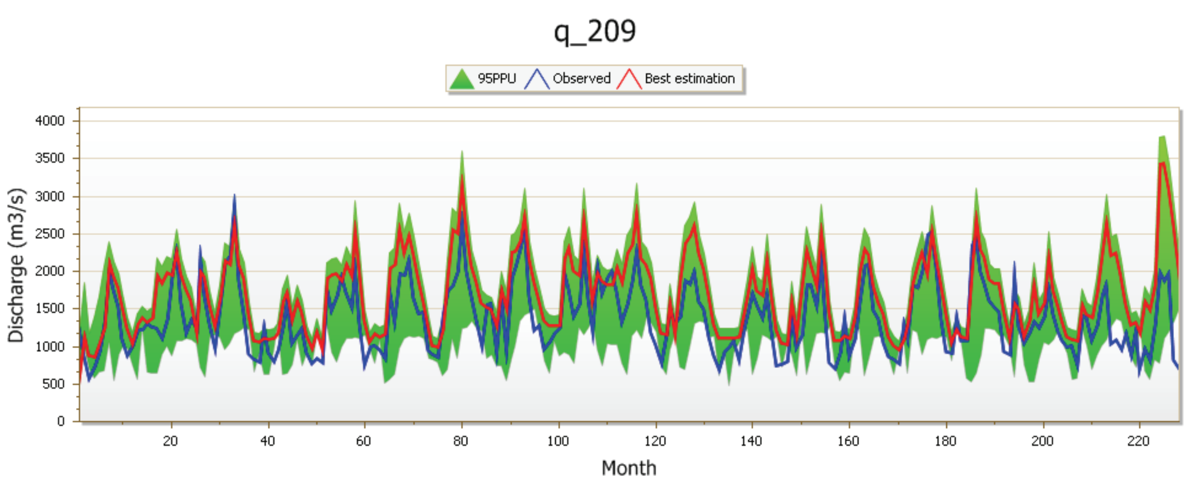

Figure 6. Observed flow, 95% model uncertainty, and best estimation at gauging station 209 after calibration.

model training workshops. The fundamental strengths of SWAT are flexibility in combining upland and channel processes and simulation of land management. As noted by Gassman et al. (2007), each process is a simplification of reality and could be improved. Gassman et al. (2007) also discussed several weaknesses that include: simplified representation of HRUs, simulation of certain management practices, pathogen fate and transport, in-stream sediment routing and kinetic functions, static soil carbon, subsurface tile flow and nitrate losses, and routines for automated sensitivity, calibration, and input uncertainty analysis. Considerable progress has been made on many of these weaknesses since Gassman et al. (2007); however, some processes are difficult to characterize accurately due to insufficient monitoring data, inadequate data to parameterize inputs, or insufficient understanding of the processes themselves.

Arnold et al. (2010a) developed routines to route flow across the landscape between HRUs, which allows for more process-based simulation of riparian and floodplain processes. Documentation and interfaces are being developed to guide users in parameterizing management scenarios. White and Arnold (2009) developed improved routines for vegetative filter strips, and Arabi et al. (2006) suggest appropriate input parameterization for several structural management practices. A web-based tool for spatial management scenario parameterization for SWAT has been developed within the eRAMS (Environmental Risk Assessment and Management System) interface (www.eramsinfo.com/erams_beta). Progress has also been made on a dynamic soil carbon model (Kemanian et al., 2011) and on improving the tile flow and nitrate submodels (Moriasi et al., 2011, 2012). Several other new model components are in final development, such as modeling different types of septic systems (Jeong et al., 2011), simulation of urban processes at shorter time intervals, and urban best management practices. These new developments will complement SWAT modeling processes and calibration in urban and septic system dominant watersheds.

As previously noted, SWAT is a comprehensive, semi-distributed model that uses readily available inputs. The weakness in a comprehensive watershed model is the high number of parameters, which complicates model parameterization and calibration. Van Griensven and Bauwens (2003) overcame some of these problems by developing an autocalibration method that reduced multiple objective functions into a single global criterion in an objective way, thus solving the weighting problem. Abbaspour et al. (2007) developed autocalibration and uncertainty software for SWAT, called SWAT-CUP, which includes the method of van Griensven and Bauwens (2003) and other methods, including a multi-site, semi-automated inverse modeling routine (SUFI-2) for calibration and uncertainty analysis. Schuol et al. (2008b) and Abbaspour et al. (2009) applied SWAT-CUP for a blue-green water analysis of the continent of Africa and the country of Iran. In SWAT-CUP, all SWAT parameters can be included in the calibration process, including all water quality parameters, crop parameters, crop rotation and management parameters, and weather generator parameters. Furthermore, rainfall and temperature can also be treated as random variables and fitted in the calibration process. Fitting rainfall data, however, should be exercised with caution, as rainfall is a driving variable and fitting it may mask the importance of other parameters.

FUTURE DEVELOPMENTS

The international modeling community has made significant strides over the last decade on model parameterization, calibration, and uncertainty analysis. As each of the remaining issues is addressed, users will build greater confidence in model results and improve conservation and environmental policy development.

IMPROVED ACCOUNTING FOR HYDROLOGIC PROCESSES

Current autocalibration tools optimize the accuracy at one or more stream gauges, without regard for predictions at locations without measured data. Sites lacking measured data have no weight in the autocalibration procedure. It is important that models accurately predict hydrographs at selected points in a watershed; however, it is equally important that the model simulates the processes realistically at all locations. For example, if surface runoff is overestimated, it is likely that ET and/or subsurface and tile flow are underestimated, resulting in overestimation of sediment yields and underestimation of subsurface nitrate yields. This will cause errors when parameterizing variables related to sediment and nutrient transport and result in unrealistic policy recommendations when running scenarios that target erosion and fertilizer management. Similarly, it is important to realistically simulate sediment sources and sinks within a watershed in addition to sediment loads at a gauge. If upland erosion is overpredicted, and thus channel erosion is underpredicted to match measured gauge loads, then management practices designed to reduce erosion from the landscape may show significant impact on total sediment yields, while in reality the practices would have little impact at the basin outlet. Before calibrating time-series of nutrient loadings (N and P) at gauging stations, the overall nutrient balance of the watershed should be examined. This step will ensure that proper processes and sources are realistically simulated, such as amount of fertilizer applied, nutrient uptake by plants, denitrification, fixation, volatilization, nitrification, and organic versus soluble nutrient loadings. Nutrient calibration should focus on calibrating the major constituents rather than calibrating total N and total P. SWAT has a dynamic nutrient simulation routine that considers transformation and movement of all constituents at multiple levels. Major constituents for N loadings are organic N and mineral N. Similarly, organic P and mineral P constitute the total P. If monitoring data are not directly available for nutrient constituents, their proportions should at least be verified. For example, total N may be calibrated to match the observed data, but the relative contribution of organic and mineral N should also be checked for the specific region. In addition, if plant growth is not properly simulated, the model may not be properly parameterized or calibrated, which may result in errors with cropping systems and fertilizer management scenarios. Nair et al. (2011) suggest that crop yield comparison be added to the calibration procedure. Compared to traditional approaches that do not include crop yield calibration, Nair et al. (2011) produced improved prediction efficiencies, especially for the nutrient balance. Faramarzi et al. (2009) found that inclusion of irrigation made a significant difference in the simulation of hydrology and calculation of correct evapotranspiration.

SPATIAL CALIBRATION

Even when models are calibrated at multiple gauge sites within a watershed, further spatial calibration would improve accuracy. In large river basins, rainfall, runoff, and water yield can vary widely across the basin. When spatial data are available for runoff, water yield, or ET, spatial calibration at the subbasin level can be used to calibrate the local water balance better, which significantly improves the temporal (time series) calibration of streamflow at the gauges in the basin (Santhi et al., 2008). When such data are not available, calibration at multiple gauges can be used to capture the spatial variation in flow, as reported by several authors (White and Chaubey, 2005; Qi and Grunwald, 2005; Santhi et al., 2001). Remotely sensed estimates of ET, leaf area index, residue cover, and soil moisture have the potential to improve spatial calibration. Remotely sensed data could also be used for calibrating landscape processes. Land use export coefficients and point-source loads should also be verified. Data are usually limited or unavailable; however, databases from research plots and small watersheds (e.g., MANAGE database; Harmel et al., 2006b, 2008) have been assembled and are useful to ensure reasonable load estimates from different land uses (HRUs) within a watershed.

RUN TIME EFFICIENCY

Many of the autocalibration techniques require hundreds or thousands of simulations to find the optimal solution. SWAT-CUP, through a parallel processing scheme developed for the Windows platform, allows individual runs to be sent to different processors, thus taking advantage of multiprocessor PCs, supercomputers, and clusters (Whittaker et al., 2004). Yalew et al. (2010) split individual SWAT simulations into several submodels, ran the submodels in parallel, collected the subbasin outputs at a central computer, and then performed the routing. This technique allows individual simulations to be parallelized and requires minor modification to the source code. Parallelizing the SWAT code by sending sections of code to different processors is also being examined.

IMPACT OF UNCERTAINTY ON CALIBRATION AND DECISION MAKING

Because models are used to develop and evaluate water resource policy, several recent pleas have been made to consider inherent uncertainties in model development and application (e.g., Beven, 2006; Bende-Michl et al., 2011). Definition and quantification of calibration uncertainty in distributed hydrological modeling has become the subject of much research in recent years (Abbaspour, 2005). Three sources of uncertainty or error must be considered: (1) the uncertainty or error in the measured input data (e.g., rainfall and temperature), (2) the uncertainty or error in the measured data used in model calibration (e.g., river discharges and sediment load), and (3) the uncertainty or error in the conceptual model and model parameters (e.g., hydrologic processes). Abbaspour (2005) states that there is an intimate relationship between calibration and uncertainty analysis and that they must be performed simultaneously. In other words, calibration must always be accompanied by an assessment of the goodness of the calibration, taking into account all modeling errors.

The uncertainties in the conceptual model and model parameters, as well as the uncertainty in measured data used in calibration, all affect simulation quality and appropriateness; therefore, Harmel et al. (2010) developed a simple model evaluation matrix to incorporate data and simulation uncertainty in model evaluation and reporting. In addition, the modified goodness-of-fit indicator calculations of Harmel et al. (2010), which are based on Haan et al. (1995), are currently being incorporated into the Abasspour et al. (2007) SWAT-CUP software.

GUIDELINES FOR CALIBRATION AND VALIDATION PERIODS

Since it is impossible to replicate watersheds and river basins, common practice in hydrologic studies is to divide the measured data either temporally or spatially for calibration and validation (Engel et al., 2007). One view suggests that both wet and dry periods be included in both the calibration and validation periods (Gan et al., 1997), ensuring that both periods reflect the range of conditions under which a model is expected to perform. This is often not feasible due to limitations in the length of monitoring data available for calibration and validation. Previous studies (Kannan et al., 2007; Van Liew and Garbrecht, 2003) recommend including a wet period with high runoff events in the calibration period. A contrasting view from Reckhow (1994) contends that validation conditions should be different in the sense that the important processes and forcing functions or responses differ from the calibrated conditions, as the purpose of validation is to provide an independent assessment of model performance. There remains some confusion in the literature about what validation is and what it means to validate a model (Rykiel 1996), and there are currently no guidelines for separating measured data for calibration and validation. Research to determine the impact of selection of calibration and validation periods on model parameterization would benefit the modeling community and advance the science of modeling. Guidelines for selection of the periods should consider recommending a minimum length of period required for calibration and validation. Such guidelines could be developed by conducting additional model runs and analysis in watersheds with extensive observed data.

AUTOMATED ERROR CHECKING

Model interfaces and automated calibration routines have simplified SWAT calibration and validation such that the effort required is a fraction of that needed a decade ago. These advances allow SWAT application by lessexperienced users and those without sufficient background in hydrology, sedimentology, soil science, and nutrient dynamics. In particular, the use of automated calibration software may produce simulated values that appear appropriate because they adequately mimic the measured data used in calibration and validation, but the model may contain input data errors and/or inappropriate parameter adjustments not readily identified by the user or the autocalibration software. SWAT Check (White et al., 2012) is a stand-alone program that examines model output relative to typical ranges, creates process-based figures for visualization of output values, and detects common model application problems. The program examines 56 model outputs and summaries and prompts users if unusual values are encountered. SWAT Check is currently in beta release, with updates pending based on user feedback. This software should assist the SWAT community (especially new users) in developing better model applications.

Manual and Automated Calibration

Manual calibration of distributed watershed models like SWAT is difficult and almost infeasible in many large-scale applications. However, manual calibration forces the user to better understand the model, the important processes in the watershed, and parameter sensitivity. Tools like VizSWAT (http://swatmodel.tamu.edu/software/vizswat.aspx) can be used to visualize complex spatial data from HRUs, subbasins, and reaches. Van Liew et al. (2005) suggested that autocalibration be attempted first, followed by manual calibration, to ensure that average annual means and the general balances are correct. Another approach is to perform manual calibration first on the average annual hydrologic balance and average annual loads (minimizing percent bias). This approach was used by Jeong et al. (2010) for the calibration of flow using the subdaily rainfall and runoff version of SWAT. Autocalibration with a narrow window of parameter ranges can then be performed to fine-tune daily and subdaily statistics.

SWAT-CUP provides a decision-making framework that incorporates a semi-automated approach (SUFI2) using both manual and automated calibration incorporating sensitivity and uncertainty analysis. Users can manually adjust parameters and ranges iteratively between autocalibration runs. Users can also use output from sensitivity and uncertainty analysis as they iteratively move between manual and autocalibration. Parameter sensitivity analysis helps focus the calibration, and uncertainty analysis is used to provide statistics for goodness-of-fit. The user interaction or manual component of the calibration forces the user to obtain a better understanding of the overall hydrologic processes (baseflow ratios, ET, sediment sources and sinks, crop yields, and nutrient balances) and of parameter sensitivity. By integrating these tools in the calibration processes, SWAT-CUP provides a powerful approach to watershed calibration.

Watershed Calibration

The ultimate goal of calibrating a watershed model should be to incorporate spatial processes into the calibration techniques. However, the models are only tools to aid in the decision process and are never a substitute for user understanding of the processes and management practices occurring in the watershed to guide calibration. We recommend that users study the watershed thoroughly, understand the processes involved, identify the specific project needs and scenarios to be analyzed, parameterize SWAT for the watershed, compare the model prediction with observed data, and then develop and implement a calibration plan. The SWAT-CUP case study shows the potential to combine spatial and process data along with user understanding of the watershed into the autocalibration process. We recommend continued addition of spatial process information to SWAT-CUP, and incorporating all the checks on processes and errors from White et al. (2012).

In many SWAT applications, additional fine-tuning of the calibration may be needed after scenario analysis is completed. It is extremely important to remember that although some SWAT input parameters are empirical, they are all physically based and must be kept within realistic ranges as influenced by the uncertainty in defining the parameter.

References

Abbaspour K. C. 2005. Calibration of hydrologic models: When is a model calibrated? In *Proc. Intl. Congress on Modelling and Simulation (MODSIM'05)*, 2449-2455. A. Zerger and R. M. Argent, eds. Melbourne, Australia: Modelling and Simulation Society of Australia and New Zealand.

Abbaspour, K. C., A. Johnson, and M. Th. van Genuchten. 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J.* 3(4): 1340-1352.

Abbaspour K. C., M. Vejdani, and S. Haghighat. 2007. SWAT-CUP calibration and uncertainty programs for SWAT. In *Proc. Intl. Congress on Modelling and Simulation (MODSIM'07)*, 1603-1609. L. Oxley and D. Kulasiri, eds. Melbourne, Australia: Modelling and Simulation Society of Australia and New Zealand.

Abbaspour, K. C., M. Faramarzi, S. S. Ghasemi, and H. Yang. 2009. Assessing the impact of climate change on water resources in Iran. *Water Resour. Res.* 45: W10434, doi: 10.1029/2008WR007615.

Ahl, R. S., S. W. Woods, and H. R. Zuuring. 2008. Hydrologic calibration and validation of SWAT in a snow-dominated Rocky Mountain watershed, Montana, U.S.A. *J. American Water Resour. Assoc.* 44(6): 1411-1430.

Akhavan, S., J. Abedi-Koupai, S. F. Mousavi, M. Afyuni, S. S. Eslamian, and K. C. Abbaspour. 2010. Application of SWAT model to investigate nitrate leaching in Hamadan-Bahar watershed, Iran. *J. Agric. Ecosystem and Environ.* 139(4): 675-688.

Alibuyog, N. R., V. B. Ella, M. R. Reyes, R. Srinivasan , C. Heatwole, and T. Dillaha. 2009. Predicting the effects of land use change on runoff and sediment yield in Manupali River subwatersheds using the SWAT model. *Intl. Agric. Eng. J.* 18(1-2): 15-25.

Andersson, J. C. M., A. J. B. Zehnder, G. P. W. Jewitt, and H. Yang. 2009. Water availability, demand, and reliability of *in situ* water harvesting in smallholder rainfed agriculture in the Thukela River basin, South Africa. *Hydrol. Earth System Sci.* 13(12): 2329-2347.

Andersson, J. C. M., A. J. B. Zehnder, J. Rockstrom, and H. Yang. 2011. Potential impacts of water harvesting and ecological sanitation on crop yield, evaporation, and river flow regimes in the Thukela River basin, South Africa. *Agric. Water Mgmt.* 98(7): 1113-1124.

Arabi, M., R. S. Govindaraju, M. M. Hantush, and B. A. Engel. 2006. Role of watershed subdivision on modeling the effectiveness of best management practices with SWAT. *J. American Water Resour. Assoc*. 42(2): 513-528.

Arnold, J. G., and P. M. Allen. 1999. Automated methods for estimating baseflow and groundwater recharge from streamflow records. *J. American Water Resour. Assoc*. 35(2): 411-424.

Arnold, J. G., and N. Fohrer. 2005. SWAT2000: Current capabilities and research opportunities in applied watershed modeling. *Hydrol. Proc.* 19(3): 563-572.

Arnold, J. G., and J. R. Williams. 1987. Validation of SWRRB: Simulator for water resources in rural basins. *J. Water Resour. Plan. Manage. ASCE* 113(2): 243-256.

Arnold, J. G., P. M. Allen, R. Muttiah, and G. Bernhardt. 1995a. Automated baseflow separation and recession analysis techniques. *Ground Water* 33(6): 1010-1018.

Arnold, J. G., J. R. Williams, and D. R. Maidment. 1995b. Continuous-time water and sediment-routing model for large basins. *J. Hydraul. Eng. ASCE* 121(2): 171-183.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. 1998. Large-area hydrologic modeling and assessment: Part I. Model development. *J. American Water Resour. Assoc.* 34(1): 73-89.

Arnold, J. G., R. S. Muttiah, R. Srinivasan, and P. M. Allen. 2000. Regional estimation of baseflow and groundwater recharge in the Upper Mississippi River basin. *J. Hydrol.* 227(1-4): 21-40.

Arnold, J. G., P. M. Allen, and D. S. Morgan. 2001. Hydrologic model for design and constructed wetlands. *Wetlands* 21(2): 167-178.

Arnold, J. G., P. M. Allen, M. Volk, J. R. Williams, and D. D. Bosch. 2010a. Assessment of different representations of spatial variability on SWAT model performance. *Trans. ASABE* 53(5): 1433-1443.

Arnold, J. G., P. W. Gassman, and M. J. White. 2010b. New developments in the SWAT ecohydrology model. In *Proc. 21st Watershed Technology Conf.: Improving Water Quality and Environment*. ASABE Publication No. 701P0210cd. St. Joseph, Mich.: ASABE.

ASCE. 1993. Criteria for evaluation of watershed models. *J. Irrig. Drainage Eng*. 119(3): 429-442.

Balascio, C. C., D. J. Palmeri, and H. Gao. 1998. Use of a genetic algorithm and multi-objective programming for calibration of a hydrologic model. *Trans. ASAE* 41(3): 615-619.

Behera, S., and R. K. Panda. 2006. Evaluation of management alternatives for an agricultural watershed in a subhumid subtropical region using a physical process-based model. *Agric. Ecosystems Environ.* 113(1-4): 62-72.

Bekele, E. G., and H. V. Knapp. 2010. Watershed modeling to assessing impacts of potential climate change on water supply availability. *Water Resour. Mgmt.* 24(13): 3299-3320.

Bekele, E. G., and J. W. Nicklow. 2007. Multi-objective automatic calibration of SWAT using NSGA-II. *J. Hydrol.* 341(3-4): 165-176.

Benaman, J., and C. A. Shoemaker. 2004. Methodology for analyzing ranges of uncertain model parameters and their impact on total maximum daily load process. Reston, Va.: ASCE.

Benaman, J., C. A. Shoemaker, and. D. A. Haith. 2005. Calibration and validation of Soil and Water Assessment Tool on an agricultural watershed in upstate New York. *J. Hydrol.* 10(5): 363-374.

Bende-Michl, U., M. Volk, R. D. Harmel, L. Newham, and T. Dalgaard. 2011. Monitoring strategies and scale-appropriate hydrologic and biogeochemical modelling for natural resource management: Conclusions and recommendations from a session held at the iEMSs 2008. *Environ. Modelling Software* 26(4): 538-542.

Beven, K. 2006. On undermining the science? *Hydrol. Proc.* 20(14): 3141-3146.

Bosch, D. D., J. M. Sheridan, H. L. Batten, and J. G. Arnold. 2004. Evaluation of the SWAT model on a coastal plain agricultural watershed. *Trans. ASAE* 47(5): 1493-1506.

Bouraoui, F., L. Galbiati, and G. Bidoglio. 2002. Climate change impacts on nutrient loads in the Yorkshire Ouse catchment (U.K.). *Hydrol. Earth Syst. Sci.* 6(2): 197-209.

Bouraoui, F., S. Benabdallah, A. Jrad, and G. Bidoglio. 2005. Application of the SWAT model on the Medjerda River basin (Tunisia). *Physics Chem. Earth.* 30(8-10): 497-507.

Boyle, D. P., H. V. Gupta, and S. Sorooshian. 2000. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resour. Res.* 36(12): 3663-3674.

Brown, L. C., and T. O. Barnwell Jr. 1987. The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and user manual. EPA/600/3-87/007. Athens, Ga.: U.S. EPA, Environmental Research Laboratory.

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

Cheng, H., W. Ouyang, F. Hao, X. Ren, and S. Yang. 2007. The nonpoint-source pollution in livestock-breeding areas of the Heihe River basin in Yellow River. *Stochastic Environ. Res. Risk Assess.* 21(3): 213-221.

Chiang, L., I. Chaubey, M. W. Gitau, and J. G. Arnold. 2010. Differentiating impacts of land use changes from pasture management in a CEAP watershed using the SWAT model. *Trans. ASABE* 53(5): 1569-1584.

Chin, D. A., D. Sakura-Lemmessy, D. D. Bosch, and P. A. Gay. 2009. Watershed-scale fate and transport of bacteria. *Trans. ASABE* 52(1): 145-154.

Chu, T. W., A. Shirmohammadi, H. Montas, and A. Sadeghi. 2004. Evaluation of the SWAT model's sediment and nutrient components in the Piedmont physiographic region of Maryland. *Trans. ASAE* 47(5): 1523-1538.

Coffey, M. E., S. R. Workman, J. L. Taraba, and A. W. Fogle. 2004. Statistical procedures for evaluating daily and monthly hydrologic model predictions. *Trans. ASAE* 47(1): 59-68.

Coffey, R., E. Cummins, N. Bhreathnach, V. O. Flaherty, and M. Cormican. 2010. Development of a pathogen transport model for Irish catchments using SWAT. *Agric. Water Mgmt.* 97(1): 101-111.

Conan, C., G. de Marsily, F. Bouraoui, and G. Bidoglio. 2003. A long-term hydrological modelling of the Upper Guadiana River basin (Spain). *Physics Chem. Earth* 28(4-5): 193-200.

Debele, B., R. Srinivasan, and J. Y. Parlange. 2008. Coupling upland watershed and downstream waterbody hydrodynamic and water quality models (SWAT and CE-QUAL-W2) for better water resources management in complex river basins. *Environ. Modeling Assess*. 13(1): 135-153.

Di Luzio, M., and J. G. Arnold. 2004. Formulation of a hybrid calibration approach for a physically based distributed model with NEXRAD data input. *J. Hydrol.* 298(1-4): 136-154.

Doherty, J. 2004. *PEST: Model-Independent Parameter Estimation User Manual.* 5th ed. Brisbane, Australia: Watermark Numerical Computing. Available at: www.pest

homepage.org/Model-independence.php. Accessed 25 June 2012.
Douglas-Mankin, K. R., R. Srinivasan, and J. G. Arnold. 2010. Soil and Water Assessment Tool (SWAT) model: Current development and applications. *Trans. ASABE* 53(5): 1423-1431.
Du, B., A. Saleh, D. B. Jaynes, and J. G. Arnold. 2006. Evaluation of SWAT in simulating nitrate nitrogen and atrazine fates in a watershed with tiles and potholes. *Trans. ASABE* 49(4): 949-959.
Duan, Q. D. 2003. Global optimization for watershed model calibration. In *Calibration of Watershed Models*, 89-104. Water Science and Application Series, Vol. 6. Washington, D.C.: American Geophysical Union.
Duan, Q. D., V. K. Gupta, and S. Sorooshian. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res*. 28(4): 1015-1031.
Easton, Z. M., D. R. Fuka, M. T. Walter, D. M. Cowan, E. M. Schneiderman, and T. S. Steenhuis. 2008. Re-conceptualizing the Soil and Water Assessment Tool (SWAT) model to predict runoff from variable source areas. *J. Hydrol.* 348(3-4): 279-291.
Eawag. 2009. SWAT-CUP. Dübendorf, Switzerland: Swiss Federal Institute of Aquatic Science and Technology. Available at: www.eawag.ch/organisation/abteilungen/siam/software/swat/index_EN.
Eckhardt, K., and J. G. Arnold. 2001. Automatic calibration of a distributed catchment model. *J. Hydrol.* 251(1-2): 103-109.
Eckhardt, K., S. Haverkamp, N. Fohrer, and H. G. Frede. 2002. SWAT-G, a version of SWAT99.2 modified for application to low mountain range catchments. *Physics Chem. Earth* 27(9-10): 641-644.
Eckhardt, K., N. Fohrer, and H. G. Frede. 2005. Automatic model calibration. *Hydrol. Proc.* 19(3): 651-658.
Engel, B. A., R. Srinivasan, J. G. Arnold, C. Rewerts, and S. J. Brown. 1993. Nonpoint-source (NPS) pollution modeling using models integrated with geographic information systems (GIS). *Water Sci. Tech.* 28(3-5): 685-690.
Engel, B., D. Storm, M. White, J. Arnold, and M. Arabi. 2007. A hydrologic/water quality model application protocol. *J. American Water Resour. Assoc.* 43(5): 1223-1236.
Faramarzi, M., K. C. Abbaspour, R. Schulin, and H. Yang. 2009. Modeling blue and green water availability in Iran. *Hydrol. Proc*. 23(3): 486-501.
Gan, T. Y., E. M. Dlamini, and G. F. Biftu. 1997. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. *J. Hydrol.* 192(1): 81-103.
Gassman, P. W., M. Reyes, C. H. Green, and J. G. Arnold. 2007. The Soil and Water Assessment Tool: Historical development, applications, and future directions. *Trans. ASABE* 50(4): 1211-1250.
Gassman, P. W., J. G. Arnold, R. Srinivasan, and M. Reyes. 2010. The worldwide use of the SWAT model: Technological drivers, networking impacts, and simulation trends. In *Proc. 21st Century Watershed Technology: Improving Water Quality and Environment*. ASABE Publication No. 701P0210cd. St. Joseph, Mich.: ASABE.
Ghaffari, G., S. Keesstra, J. Ghodousi, and H. Ahmadi. 2010. SWAT-simulated hydrological impact of land-use change in the Zanjanrood basin, northwest Iran. *Hydrol. Proc*. 24(7): 892-903.
Gikas, G., T. Yiannakopoulou, and V. Tsihrintzis. 2006. Modeling of nonpoint-source pollution in a Mediterranean drainage basin. *Environ. Modeling Assess.* 11(3): 219-233.
Gitau, M. W., W. J. Gburek, and P. L. Bishop. 2008. Use of the SWAT model to quantify water quality effects of agricultural BMPs at the farm-scale level. *Trans. ASABE* 51(6): 1925-1936.
Govender, M., and C. S. Everson. 2005. Modelling streamflow from two small South African experimental catchments using the SWAT model. *Hydrol. Proc.* 19(3): 683-692.
Green, C. H., and A. van Griensven. 2008. Autocalibration in hydrologic modeling: Using SWAT2005 in small-scale watersheds. *Environ. Modelling Software* 23(4): 422-434.
Green, C. H., M. D. Tomer, M. Di Luzio, and J. G. Arnold. 2006. Hydrologic evaluation of the Soil and Water Assessment Tool for a large tile-drained watershed in Iowa. *Trans. ASABE* 49(2): 413-422.
Green, C. H., J. G, Arnold, J. R. Williams, R. Haney, and R. D. Harmel. 2007. Soil and Water Assessment Tool hydrologic and water quality evaluation of poultry litter application to small-scale subwatersheds in Texas. *Trans. ASABE* 50(4): 1199-1209.
Green, W. H., and G. A. Ampt. 1911. Studies on soil physics: 1. The flow of air and water through soils. *J. Agric. Sci.* 4(1): 11-24.
Grizzetti, B., F. Bouraoui, and G. De Marsily. 2005. Modelling nitrogen pressure in river basins: A comparison between a statistical approach and the physically based SWAT model. *Physics Chem. Earth* 30(8-10): 508-517.
Haan, C. T., B. Allred, D. E. Storm, G. J. Sabbagh, and S. Prahhu. 1995. Statistical procedure for evaluating hydrologic/water quality models. *Trans. ASAE* 38(3): 725-733.
Harmel, R. D., R. J. Cooper, R. M. Slade, R. L. Haney, and J. G. Arnold. 2006a. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASABE* 49(3): 681-701.
Harmel, R. D., S. Potter, P. Casebolt, K. Reckhow, C. Green, and R. Haney. 2006b. Compilation of measured nutrient load data for agricultural land uses in the United States. *J. American Water Resour. Assoc.* 42(5): 1163-1178.
Harmel, R. D., S. Qian, K. Reckhow, and P. Casebolt. 2008. The MANAGE database: Nutrient load and site characteristics updates and runoff concentration data. *J. Environ. Qual.* 37(6): 2403-2406.
Harmel, R. D., D. R. Smith, K. W. King, and R. M. Slade. 2009. Estimating storm discharge and water quality data uncertainty: A software tool for monitoring and modeling applications. *Environ. Modelling Software* 24(7): 832-842.
Harmel, R. D., P. K. Smith, and K. L. Migliaccio. 2010. Modifying goodness-of-fit indicators to incorporate both measurement and model uncertainty in model calibration and validation. *Trans. ASABE* 53(1): 55-63.
Heuvelmans, G., B. Muys, and J. Feyen. 2004. Evaluation of hydrological model parameter transferability for simulating the impact of land use on catchment hydrology. *Physics Chem. Earth* 29(11-12): 739-747.
Heuvelmans, G., B. Muys, and J. Feyen. 2006. Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets. *J. Hydrol.* 319(1-4): 245-265.
Hu, X., G. F. McIsaac, M. B. David, and C. A. L. Louwers. 2007. Modeling riverine nitrate export from an east-central Illinois watershed using SWAT. *J. Environ. Qual.* 36(4): 996-1005.
Inamdar, S. and A. Naumov. 2006. Assessment of sediment yields for a mixed land use Great Lakes watershed: Lessons from field measurements and modeling. *J. Great Lakes Res.* 32(3): 471-488.
Jeong, J., N. Kannan, J. G. Arnold, R. Glick, L. Gosselink, and R. Srninvasan. 2010. Development and integration of subhourly rainfall-runoff modeling capability within a watershed model. *Water Resour. Mgmt.* 24(15): 4505-4527.

Jeong, J., C. Santhi, J. G. Arnold, R. Srinivasan, S. Pradhan, and K. Flynn. 2011. Development of algorithms for modeling onsite wastewater systems within SWAT. *Trans. ASABE* 54(5): 1693-1704.

Jha, M. K., K. E. Schilling, P. W. Gassman, and C. F. Wolter. 2010. Targeting land-use change for nitrate-nitrogen load reductions in an agricultural watershed. *J. Soil Water Cons.* 65(6): 342-352.

Kannan, N., S. M. White, F. Worrall, and M. J. Whelan. 2007. Sensitivity analysis and identification of the best evapotranspiration and runoff options for hydrological modeling in SWAT-2000. *J. Hydrol.* 332(3-4): 456-466.

Kemanian, A. R., S. Julich, V. S. Manoranjan, and J. G. Arnold. 2011. Integrating soil carbon cycling with that of nitrogen and phosphorus in the watershed model SWAT: Theory and model testing. *Ecol. Modelling* 222(12): 1913-1921.

Knisel, W. G. 1980. CREAMS: A field-scale model for chemicals, runoff, and erosion from agricultural management systems. Conservation Research Report No. 26. Washington, D.C.: USDA National Resources Conservation Service.

Krause, P., D. P. Boyle, and F. Base. 2005. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* 5: 89-97.

Krysanova, V., and J. G. Arnold. 2008. Advances in ecohydrological modeling with SWAT: A review. *Hydrol. Sci. J.* 53(5): 939-947.

Legates, D. R., and G. J. McCabe. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* 35(1): 233-241.

Lemonds, P. J., and J. E. McCray. 2007. Modeling hydrology in a small rocky mountain watershed serving large urban populations. *J. American Water Resour. Assoc.* 43(4): 875-887.

Leonard, R. A., W. G. Knisel, and D. A. Still. 1987. GLEAMS: Groundwater loading effects on agricultural management systems. *Trans. ASAE* 30(5): 1403-1418.

Ma, L., J. C. Ascough II, L. R. Ahuja, M. J. Shaffer, J. D. Hanson, and K. W. Rojas. 2000. Root Zone Water Quality Model sensitivity analysis using Monte Carlo simulation. *Trans. ASAE* 43(4): 883-895.

Maski, D., K. R. Mankin, K. A. Janssen, P. Tuppad, and G. M. Pierzynski. 2008. Modeling runoff and sediment yields from combined in-field crop practices using the Soil and Water Assessment Tool. *J. Soil Water Cons.* 63(4): 193-203.

Meng, H., A. M. Sexton, M. C. Maddox, A. Sood, C. W. Brown, R. R. Ferraro, and R. Murtugudde. 2010. Modeling Rappahannock River basin using SWAT-Pilot for Chesapeake Bay watershed. *Trans. ASAE* 26(5): 795-805.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Binger, R. D. Harmel, and T. Veith. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 50(3): 885-900.

Moriasi, D. N., J. G. Arnold, G. G. Vasquez-Amabile, B. A. Engel, and C. G. Rossi. 2011. Shallow water table depth algorithm in SWAT: Recent developments. *Trans. ASABE* 54(5): 1705-1711.

Moriasi, D. N., C. G. Rossi, J. G. Arnold, and M. D. Tomer. 2012. Evaluating hydrology of SWAT with new tile drain equations. *J. Soil Water Cons.* (accepted).

Mukundan, R., D. E. Radcliffe, and L. M. Risse. 2010. Spatial resolution of soil data and channel erosion effects on SWAT model predictions of flow and sediment. *J. Soil Water Cons.* 65(2): 92-104.

Muleta, M. K., and J. W. Nicklow. 2005. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *J. Hydrol.* 306(1-4): 127-145.

Nair, S. S., K. W. King, J. D. Witter, B. L. Sohngen, and N. R. Fausey. 2011. Importance of crop yield in calibrating watershed water quality simulation models. *J. American Water Res. Assoc.* 47(6): 1285-1297.

Narasimhan, B., R. Srinivasan, S. T. Bednarz, M. R. Ernst, and P. M. Allen. 2010. A comprehensive modeling approach for reservoir water quality assessment and management due to point and nonpoint source pollution. *Trans. ASABE* 53(5): 1605-1617.

Nash, J. E., and J. E. Sutcliffe. 1970. River flow forecasting through conceptual models: Part I. A discussion of principles. *J. Hydrol.* 10(3): 282-290.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, R. Srinivasan, and J. R. Williams. 2002. Soil and Water Assessment Tool, User Manual, Version 2000. Temple, Tex.: Grassland, Soil and Water Research Laboratory.

Ng, T. L., J. W. Eheart, X. Cai, and F. Miguez. 2010a. Modeling *Miscanthus* in the Soil and Water Assessment Tool (SWAT) to simulate its water quality effects as a bioenergy crop. *Environ. Sci. Tech.* 44(18): 7138-7144.

Ng, T. L., J. W. Eheart, and X. M. Cai. 2010b. Comparative calibration of a complex hydrologic model by stochastic methods GLUE and PEST. *Trans. ASABE* 53(6): 1773-1786.

Olivera, F., M. Valenzuela, R. Srinivasn, J. Choi, H. Cho, S. Koka, and A. Agrawal. 2006. ArcGIS-SWAT: A geodata model and GIS interface for SWAT. *J. American Water Resources Assoc*. 42(2): 295-309.

Parajuli, P. B., N. O. Nelson, L. D. Frees, and K. R. Mankin. 2009. Comparison of AnnAGNPS and SWAT model simulation results in USDA-CEAP agricultural watersheds in south-central Kansas. *Hydrol. Proc.* 23(5): 748-763.

Qi, C., and S. Grunwald. 2005. GIS-based hydrologic modeling in the Sandusky watershed using SWAT. *Trans. ASAE* 48(1): 169-180.

Razavi, S., B. A. Tolson, L. S. Matott, N. R. Thomson, A. MacLean, and F. R. Seglenieks. 2010. Reducing the computational cost of automatic calibration through model preemption. *Water Resour. Res.* 46: W11523, doi: 10.1029/2009WR008957.

Reckhow, K. H. 1994. Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. *Ecol. Model.* 72(1-2): 1-20.

Refsgaard, J. C. 1997. Parameterisation, calibration, and validation of distributed hydrological models. *J. Hydrol.* 198(1): 69-97.

Rouholahnejad, E., K. C. Abbaspour, M. Vejdani, R. Srinivasan, R. Schulin, and A. Lehmann. 2012a. A parallelization framework for calibration of hydrological models. *Environ. Modelling Software* 31: 28-36.

Rouholahnejad, E., K. C. Abbaspour, R. Srinivasan, R. Schulin, and A. Lehmann. 2012b. Using a hydrological model of the Danube River basin to calculate nitrate and phosphate loads. *J. Hydrol.* (in preparation).

Rykiel, E. J. 1996. Testing ecological models: The meaning of validation. *Ecol. Model.* 90(3): 229-244.

Santhi, C., J. G. Arnold, J. R. Williams, W. A. Dugas, R. Srinivasan, and L. M. Hauck. 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. *J. American Water Resour. Assoc.* 37(5): 1169-1188.

Santhi, C., N. Kannan, J. G. Arnold, and M. Di Luzio. 2008. Spatial calibration and temporal validation of flow for regional-scale hydrologic modeling. *J. American Water Resour. Assoc.* 44(4): 829-846.

Schuol, J., K. C. Abbaspour, R. Srinivasan, and H. Yang. 2008a. Estimation of freshwater availability in the west African subcontinent using the SWAT hydrologic model. *J. Hydrol.* 352(1-2): 30-49.

Schuol, J., K. C. Abbaspour, H. Yang, R. Srinivasan, and A. J. B. Zhender. 2008b. Modeling blue and green water availability in Africa. *Water Resour. Res.* 44: W07406, doi: 10.1029/2007WR006609.

Setegn, S. G., R. Srinivasan, and B. Dargahi. 2009. Hydrological modelling in the Lake Tana basin, Ethiopia, using SWAT model. *Open Hydrol. J.* 2: 49-62.

Shoemaker, C. A., R. G. Regis, and R. C. Fleming. 2007. Watershed calibration using multistart local optimization and evolutionary optimization with radial basis function approximation. *Hydrol. Sci.* 52(3): 450-465.

Srinivasan, R., and J. G. Arnold. 1994. Integration of a basin-scale water quality model with GIS. *Water Resour. Bull.* 30(3): 453-462.

Srinivasan, R., X. Zhang, and J. G. Arnold. 2010. SWAT ungauged: Hydrological budget and crop yield predictions in the Upper Mississippi River basin. *Trans. ASABE* 53(5): 1533-1546.

Starks, P. J., and D. N. Moriasi. 2009. Spatial resolution effect of precipitation data on SWAT calibration and performance: Implications for CEAP. *Trans. ASABE* 52(4): 171-1180.

Sudheer, K. P., I. Chaubey, V. Garg, and K. W. Migliaccio. 2007. Impact of time scale of the calibration objective function on the performance of watershed models. *Hydrol. Proc.* 21(25): 3409-3419.

Sui, Y., and J. R. Frankenberger. 2008. Nitrate loss from subsurface drains in an agricultural watershed using SWAT2005. *Trans. ASABE* 51(4): 1263-1272.

Tuppad, P., K. R. Douglas-Mankin, T. Lee, R. Srinivasan, and J. G. Arnold. 2011. Soil and Water Assessment Tool (SWAT) hydrologic/water quality model: Extended capability and wider adoption. *Trans. ASABE* 54(5): 1677-1684.

van Griensven, A. 2002. Developments towards integrated water quality modeling for river basins. Publication No. 40. Brussels, Belgium: Vrije Universiteit, Department of Hydrology and Hydraulic Engineering.

van Griensven, A., and W. Bauwens. 2003. Multiobjective autocalibration for semidistributed water quality models. *Water Resour. Res.* 39(12): 1348-1356.

van Griensven, A., and W. Bauwens. 2005. Application and evaluation of ESWAT on the Dender basin and Wister Lake basin. *Hydrol. Proc.* 19(3): 827-838.

van Griensven, A., T. Meixner, R. Srinivasan, and S. Grunwald. 2008. Fit-for-purpose analysis of uncertainty using split-sampling evaluations. *Hydrol. Sci.* 53(5): 1090-1103.

Van Liew, M. W., and J. Garbrecht. 2003. Hydrologic simulation of the Little Washita River experimental watershed using SWAT. *J. American Water Resour. Assoc.* 39(2): 413-426.

Van Liew, M. W., J. G. Arnold, and D. D. Bosch. 2005. Problems and potential of autocalibrating a hydrologic model. *Trans. ASAE* 48(3): 1025-1040.

Van Liew, M. W., C. H. Green, and P. J. Starks. 2007. Unit source area data: Can it make a difference in calibrating the hydrologic response for watershed-scale modeling? *J. Soil Water Cons.* 62(3): 162-170.

Wang, X., and A. M. Melesse. 2005. Evaluation of the SWAT model's snowmelt hydrology in a northwestern Minnesota watershed. *Trans. ASAE* 48(4): 1359-1376.

Wang, X., N. Kannan, C. Santhi, S. R. Potter, J. R. Williams, and J. G. Arnold. 2011. Integrating APEX output for cultivated cropland with SWAT simulation for regional modeling. *Trans. ASABE* 54(4): 1281-1298.

White, K. L., and I. Chaubey. 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. *J. American Water Resour. Assoc.* 41(5): 1077-1089.

White, M. J., and J. G. Arnold. 2009. Development of a simplistic vegetative filter strip model for sediment and nutrient retention at the field scale. *Hydrol. Proc.* 23(11): 1602-1616.

White, M. J., D. E. Storm, P. R. Busteed, M. D. Matlock, and R. R. West. 2011. Evaluating potential phosphorus management impacts in the Lake Eucha basin using SWAT. *Trans. ASABE* 54(3): 827-835.

White, M. J., R. D. Harmel, J. G. Arnold, and J. R. Williams. 2012. SWAT Check: A screening tool to assist users in the identification of potential model application problems. *J. Environ. Quality* (in press). Available at: http://swatmodel.tamu.edu/software/swat-check.

Whittaker, G. R. 2004. Use of a Beowulf cluster for estimation of risk using SWAT. *Agron. J.* 95(5): 1495-1497.

Williams, J. R., and H. D. Berndt. 1977. Sediment yield prediction based on watershed hydrology. *Trans. ASAE* 20(6): 1100-1104.

Williams, J. R., J. G. Arnold, J. R. Kiniry, P. W. Gassman, and C. H. Green. 2008. History of model development at Temple, Texas. *Hydrol. Sci.* 53(5): 948-960.

Yalew, S. G., A. van Griensven, and L. Kokoszkiewcz. 2010. Parallel computing of a large-scale spatially distributed model using the Soil and Water Assessment Tool (SWAT). In *Proc. 2010 Intl. Congress on Environ. Modeling and Software: Modeling for Environment's Sake*. Manno, Switzerland: International Environmental Modeling and Software Society. Available at: www.iemss.org/iemss2010/index.php?n=Main.Proceedings. Accessed 1 October 2011.

Yang, J., K. C. Abbaspour, P. Reichert, and H. Yang. 2008. Comparing uncertainty analysis techniques for a SWAT application to Chaohe basin in China. *J. Hydrol.* 358(1-2): 1-23.

Zhang, X., R. Srinivasan, B. Debele, and F. Hao. 2008a. Runoff simulation of the headwaters of the Yellow River using the SWAT model with three snowmelt algorithms. *J. American Water Resour. Assoc.* 44(1): 48-61.

Zhang, X., R. Srinivasan, and M. Van Liew. 2008b. Multisite calibration of the SWAT model for hydrologic modeling. *Trans. ASABE* 51(6): 2039-2049.