Modeling at catchment scale and associated uncertainties

Adel Shirmohammadi\textsuperscript{1)*}, Tzyy-Woei Chu\textsuperscript{2)} & Hubert J. Montas\textsuperscript{1)}

\textsuperscript{1)} Fischell Department of Bioengineering, University of Maryland, College Park, MD 20742, USA (*e-mail: ashirmo@umd.edu)
\textsuperscript{2)} Department Civil Engineering, National Taipei University of Technology, Taipei, Taiwan

Received 7 Feb. 2007, accepted 18 Oct 2007 (Editor in charge of this article: Raija Laiho; guest editor: Ülo Mander)

This study describes application of a catchment scale model, SWAT (Soil Water Assessment Tool), to a small scale agricultural watershed in northern Maryland. It covers the steps involved in model application and associated model uncertainty as affected by variability in input parameters using Latin Hyper Cube Sampling (LHS) with Constrained Monte Carlo Simulation (MCS). SWAT model predictions of the impact of environmentally friendly practices are discussed within the context of input variability. Results indicate that SWAT is a reasonable monthly predictor of hydrology, but does not provide strong association between measured and simulated nitrate loss at that time scale. SWAT was found to perform very well when used for annual nitrate loss predictions. Results also show that using average input parameter values without considering their variability due to media heterogeneity produces simulation outputs that can be misleading and should not be given 100\% confidence. It was concluded that in developing TMDL (Total Maximum Daily Load) plans for a given watershed one has to assert associated uncertainty levels in model’s inputs and simulation results for proper resource management.

Introduction

To address the interaction between human life and the surrounding environment in the landscape, the “peep-hole” principle has mostly been used (Hagerstrand 1992a, 1992b). The result is that the landscape mantle is understood to a limited degree only, mainly as related to biological systems and to components of economic importance related to the use of natural resources. Recent heightened concern for sustainability has encouraged scientists to evaluate the multi-cause problems of the environment in relation to human and animal life under diverse conditions (Shirmohammadi \textit{et al}. 2005, Falkenmark and Mikulski 1994). Efforts to respond to the issue of sustainability have produced multicomponent water quality models describing hydrologic and water quality responses of the landscape under diverse climatic and managerial conditions. In most cases, these models have used the systems approach in describing a natural event rather than looking at each event as isolated phenomena (e.g., continuous simulation models such as CREAMS (Knisel 1980), GLEAMS (Leonard 1987), SWAT (Arnold \textit{et al}. 1998)).

A model may have different interpretations based on its discipline of use. In hydrology, water quality, and engineering, models are used to explain natural phenomena and under some
conditions make deterministic and/or probabilistic predictions. In other words, a modeler tries to use established laws or circumstantial evidences in order to represent the real life scenario mathematically, producing and end product called “model”. Although each modeler tries to represent the real system, the strengths and weaknesses of their models depend upon the modeler’s background, theoretical and empirical algorithms used in the model, the application conditions, and the scale of application. One should note that Aristotle and his ideas that “inaccessible is more challenging to explore than the accessible in everyday’s world” seem to have had a guiding influence on the development of water quality models. Additionally, the “particle theory” of Einstein that “universe has a grain structure and each grain is in a relative state with respect to the others”, has formed the basis for describing inter-relationships between different components of water quality models. For instance, a natural scientist is concerned about the inter-relationships governing the state of a given environment and tries to understand such relationship using experimental procedures and biological principles. The products of such studies are generally a set of factual data and possibly some empirical models describing such relationships. Such empirical models are developed under specific conditions and their use for conditions other than the one under which they have been developed may generate significant errors in model predictions. A physicist and an engineer on the other hand, try to use physical laws and mechanistic approaches to describe inter-relationships governing the state of an event and produce deterministic and mechanistic models. Such models are not complete until they have been calibrated, validated and tested against experimental data (Shirmohammadi et al. 2001). In addition, measuring or determining proper parameter values for such models is a challenge at best.

Watershed scale hydrologic and non-point source pollution models are useful tools in assessing the environmental condition of a watershed and evaluating the potential effects of implementing Best Management Practices (BMPs) to help reduce the damaging effects of storm runoff, baseflow, and groundwater on water bodies and the landscape. Although their complexities may differ, such models may also be useful tools in the development of TMDL (Total Maximum Daily Load) plans to meet various water quality standards, as required by the Clean Water Act. Numerous models have been developed and are in use either as research, management, or regulatory tools (Shirmohammadi et al. 2001).

Watershed scale nonpoint source pollution models use the principles used in the field scale models and extend them to mixed land use scenarios. For example, AGNPS by Young et al. (1989), SWRRB by Arnold et al. (1990), and SWAT by Arnold et al. (1998) are all built upon the strength of the USDA’s CREAMS model (Knisel 1980). They are all continuous simulation models with daily time steps. Some watershed models such as ANSWERS-2000 (Bouraoui and Dillaha 1996) are event based, thus requiring more detailed climatic data. Watershed scale models such as SWAT and ANSWERS-2000 are distributed parameter models, thus enabling the user to consider the diversities in land use, soils, topography, and management alternatives within the watershed. These models generally contain routing algorithms that consider the attenuation of sediment and chemicals through upland areas as well as within the stream system. Such distributed parameter models are also adaptable to the GIS environment. However, variability in input parameter values due to heterogeneity in natural system (e.g., climate, soils, land use, etc.) results in output uncertainty in these models (Haan et al. 1995, Sohrabi et al. 2003, Shirmohammadi et al. 2006). This study intends to illustrate advantages and disadvantages of a watershed scale model such as SWAT in its application to a small agricultural watershed. The goal is to provide a guideline for using such models and calling on the users to be cautious about their output uncertainty.

Materials and methods

Study site and available data

The study site is a small agricultural watershed (346 ha), Warner Creek, in the piedmont physiographic region of Maryland (Fig. 1). The entire Warner Creek watershed was divided into 40
subwatersheds based on topography and similar land use and soils for the SWAT simulation using Digital Elevation Models (DEMs) on the GIS platform. The SWAT model is linked with ArcView GIS, and performs subwatershed divisions for the user based on DEMs. Then, the user can identify smaller virtual hydrologic response units (HRUs) based on soils and land use in each subwatershed. The model routes output from each HRU to the outlet of the corresponding subwatershed. Finally, all discharges from the outlet of subwatersheds are routed to the outlet of the whole watershed, which is station 2A in our site. An automated flowmeter and sampler were used to acquire flow integrated samples at station 2A. Sediment and nutrient concentration were determined by analyzing the water samples in the laboratory using automated ion analyzer based on the colorimetric method for nutrients. Finally, sediment and nutrient loads were computed using flow volume and concentration of constituents. The measured streamflow at station 2A was separated into storm flow (surface runoff) and base flow using the streamflow partitioning method proposed by Linsley et al. (1982). The simplest method of Linsley, which connects the beginning of the rising limb of a hydrograph to the inflection point on the recession limb of the hydrograph, was used for flow separation. The data collection period lasted from 1994 to 2002.

Modeling

This study uses SWAT model (Arnold et al. 1998) for a case study to show advantages and disadvantages of using watershed scale models. SWAT is a complex, physically based model with spatially explicit parameterization capability. A complete description of SWAT’s components is found in Arnold et al. (1998). In brief, SWAT is a continuous simulation model and operates on a daily time step to perform simulations up to one hundred years using measured and/or stochastically generated weather data. A GIS-based user interface, AVSWAT (ArcView SWAT), developed by Diluzio et al. (2004), was used in this study to help reduce spatial data collection and manipulation time, and also allows the user to modify and analyze various alternative management practices efficiently.

The major components of SWAT model include hydrology, weather, sedimentation, soil temperature, crop growth, nutrients, pesticides, agricultural management, channel routing, and reservoir routing (Arnold et al. 1998). The hydrology component consists of weather data extracted from input files of measured data or from the weather component of model (precipitation, maximum/minimum air temperature, solar radiation, wind speed, and relative humidity), surface runoff, percolation, lateral subsur-
face flow, groundwater flow, evapotranspiration, snow melt, transmission losses, and channel flood routing. The SWAT model estimates soil erosion and sediment yield from the landscape and in-stream depositional and degrading processes. Sediment yield from the landscape is calculated by the Modified Universal Soil Loss Equation (MUSLE). MUSLE follows the structure of the Universal Soil Loss Equation (USLE) with the exception that the rainfall factor is replaced by the runoff factor (Blaszczynski 2003). Nitrogen and phosphorus processes in SWAT are handled in a similar manner as in the Erosion Productivity Impact Calculator (EPIC) model (Williams 1995). SWAT also adopted a modified version of QUAL2E model (Brown and Barnwell 1987, Ramanarayanan et al. 1996) to simulate in-stream nutrient transformations. QUAL2E is intended for use as a water quality planning tool, which can be operated as a steady state or as a dynamic model. The sub-components of QUAL2E include models of the biochemical dynamics of algae as chlorophyll a, dissolved oxygen, carbonaceous oxygen demand, organic nitrogen, ammonium nitrogen, nitrite nitrogen, nitrate nitrogen, organic phosphorus and soluble phosphorus.

It is convenient to think of the SWAT model as having \( k \) input variables: \( X_1, X_2, \ldots, X_k \), and producing from them a set of \( j \) output variables: \( Y_1, Y_2, \ldots, Y_j \). The input and output variables can be parameters, time series or spatial series. Model operation is denoted by \( f \) so that the relationship between input and output becomes:

\[
(Y_1, Y_2, \ldots, Y_j) = f(X_1, X_2, \ldots, X_k) \quad (1)
\]

In deterministic applications of the model, the uncertainty in \( X_1, X_2, \ldots, X_k \) is not considered leading to deterministic outputs \( Y_1, Y_2, \ldots, Y_j \). In stochastic applications, the input variables, \( X_1, X_2, \ldots, X_k \), are considered to be random with specific statistics and distributions representing their uncertainty and the model produces output variables that are also random variables. This study applied SWAT deterministically for calibration and validation and then stochastically to illustrate the effects of uncertainty as described in the next section.

**Data analysis and model uncertainty**

The calibration and validation of SWAT were performed by comparing predicted surface runoff, the sum of lateral subsurface flow and groundwater flow, and nitrate concentrations to measured surface runoff, measured base flow and measured nitrate concentrations, respectively. Monthly measured data, from April 1994 through December 1995, were used for model calibration. The data for the remaining period (1996 through 2002) were used for model validation.

Graphical methods (time series plot and scattergram), and statistical measures were used to evaluate the model performance based on the measured data. Four statistical criteria were used to evaluate the hydrologic goodness-of-fit: correlation coefficient \( r \), a measure of the intensity of association between observations and model predictions), coefficient of determination \( r^2 \), the square of \( r \) as used in regression, it is the proportion of the variance of observed values that is explained by the model after the predicted mean has been adjusted to equal the observed mean), Nash-Sutcliffe coefficient \( R^2 \), the same as \( r^2 \) but without adjusting the predicted mean to equal the observed mean, i.e. \( R^2 = 1 – \text{RMS}^2/\text{Variance} \) of observations (Nash and Sutcliffe 1970), and root mean square error (RMS). Detailed procedure, results, and discussions of both calibration and validation periods are presented in Chu and Shirmohammadi (2004) and Chu et al. (2004).

To illustrate the impact of uncertainty in model output, this study used LHS (Latin Hypercube Sampling) with constrained Monte Carlo Simulation (MCS). The LHS procedure described by McKay et al. (1979) and Iman et al. (1981) was used. A detailed uncertainty analysis on SWAT may be found in Sohrabi et al. (2003), however, sample data are presented to compare model output obtained based on the average input parameter values within the output probability distribution range for the Warner Creek Watershed. The technique is essentially an optimized Monte Carlo approach with solid track record in decision theory.

The application of LHS started by identifying the most sensitive hydrologic, soil, chemis-
try, and management parameters and identified appropriate probability distribution functions (pdf, e.g. normal, logarithmic, beta, gamma, uniform) for each one. Groups of parameter values were generated for each sensitive parameter by a targeted randomization strategy that preserves distribution moments. SWAT was then used to predict flow and transport for each group of randomized input variables, producing an ensemble of model outputs. For example, following the formalism of Eq. 1, the model predictions resulting from the $i$th group of input variables would be written as:

$$(Y_1, Y_2, \ldots, Y_j) = f(X_1, X_2, \ldots, X_k)$$  \hspace{1cm} (2)$$

and the targeted LHS input sampling strategy would produce a number $n$ of such predictions. In this study, $n$ was of the order of 500. The output vectors were then processed for specified variables, such as streamflow, sediment, and nutrients, to produce cumulative probability distribution curves.

This stochastic model application procedure was applied to the analysis of the effects of several best management practices for the Warner Creek watershed. For the purpose of this study only one of the practices, BMP4 (contour strip cropping with no-till), was selected for illustration. Deterministic simulation results using average input values were overlaid on the cumulative distribution curves obtained by LHS-MCS to illustrate the differences between the two approaches. It is also notable that several summary statistics such as: (1) the range of $Y$, (2) the mean and variance of $Y$, and (3) the lower and upper quartiles for $Y$ can be readily evaluated from the output of LHS although this will not be performed here (Iman and Helton 1985).

**Results and discussion**

The statistical results of the model performance for the hydrologic parameters during both calibration and validation periods are summarized in Table 1. Previous study (Chu and Shirmohammadi 2004) pointed out the presence of subsurface flow contributions from outside the watershed boundary. Measured base flow was therefore corrected for the extra subsurface flow contribution from outside the watershed using the water balance approach. Increased values of $r$, $r^2$, and $R^2$ all indicate reasonable performance of the SWAT model during the validation period. As compared with results reported earlier (Chu and Shirmohammadi 2004), flow adjustments for contributions from outside the watershed improved model performance significantly. All nutrient loadings leaving the watershed were also adjusted to subtract the chemical transport via subsurface flow contribution from outside.

**Table 1.** Statistical results comparing monthly measured and simulated flow data at station 2A after adjustment to the subsurface flow contribution from outside the watershed.

<table>
<thead>
<tr>
<th>Hydrological components</th>
<th>No. of samples</th>
<th>$r$</th>
<th>$r^2$</th>
<th>$R^2$ (Nash-Sutcliffe)</th>
<th>RMS (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calibration period (April 1994–1995)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stream flow</td>
<td>21</td>
<td>0.83</td>
<td>0.69</td>
<td>0.68</td>
<td>16.2</td>
</tr>
<tr>
<td>Surface runoff</td>
<td>21</td>
<td>0.66</td>
<td>0.43</td>
<td>0.35</td>
<td>11.2</td>
</tr>
<tr>
<td>Subsurface flow</td>
<td>21</td>
<td>0.75</td>
<td>0.57</td>
<td>0.53</td>
<td>12.2</td>
</tr>
<tr>
<td><strong>Validation period (1997–1999)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stream flow</td>
<td>36</td>
<td>0.89</td>
<td>0.78</td>
<td>0.78</td>
<td>19.9</td>
</tr>
<tr>
<td>Surface runoff</td>
<td>36</td>
<td>0.91</td>
<td>0.83</td>
<td>0.74</td>
<td>10.7</td>
</tr>
<tr>
<td>Subsurface flow</td>
<td>36</td>
<td>0.81</td>
<td>0.66</td>
<td>0.62</td>
<td>16.8</td>
</tr>
<tr>
<td><strong>Validation period (1997–2002)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stream flow</td>
<td>72</td>
<td>0.85</td>
<td>0.72</td>
<td>0.71</td>
<td>18.1</td>
</tr>
<tr>
<td>Surface runoff</td>
<td>72</td>
<td>0.84</td>
<td>0.71</td>
<td>0.68</td>
<td>9.4</td>
</tr>
<tr>
<td>Subsurface flow</td>
<td>72</td>
<td>0.79</td>
<td>0.62</td>
<td>0.53</td>
<td>14.7</td>
</tr>
</tbody>
</table>
the watershed. This process permits a fair evaluation of the SWAT model, especially for small watersheds like the one used in this study.

The statistical results of the model performance in monthly and yearly nitrate prediction during calibration and validation periods are summarized in Table 2 (monthly calibration and validation in rows 1 and 2 and combined yearly results in row 3). Low values of $r$, $r^2$, and $R^2$ indicate that despite improvements in model performance during validation, results of monthly simulations of NO$_3$-N are poor. However, despite poor performance in predicting monthly nitrate loadings, the yearly simulations showed a strong agreement (Fig. 2 and 3). Statistical results in Table 2 with high values of $r$, $r^2$, and $R^2$ (0.82, 0.68, and 0.63, respectively) are for the period of 9 years (1994–2002).

Evaluation of Uncertainty in the simulation results of TMDL models such as SWAT is essential (Shirmohammadi et al. 2006). Figure 4 shows the cumulative probability distribution of stream flow obtained by the SWAT model using LHS-MCS strategy. The 9-year average stream

### Table 2. Statistical results comparing measured and simulated NO$_3$-N at station 2A.

<table>
<thead>
<tr>
<th>Nutrient parameters</th>
<th>No. of samples</th>
<th>$r$</th>
<th>$r^2$</th>
<th>$R^2$ (Nash-Sutcliffe)</th>
<th>RMS (kg ha$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monthly</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration period (April 1994–1995)</td>
<td>NO$_3$-N (adjusted)</td>
<td>21</td>
<td>0.52</td>
<td>0.27</td>
<td>0.16</td>
</tr>
<tr>
<td>Validation period (1996–2002)</td>
<td>NO$_3$-N (adjusted)</td>
<td>83</td>
<td>0.53</td>
<td>0.28</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Yearly</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1994–2002)</td>
<td>NO$_3$-N (adjusted)</td>
<td>9</td>
<td>0.82</td>
<td>0.68</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Fig. 2. Time-series plot of measured and simulated yearly nitrate loading at Station 2A.

Fig. 3. Scattergram of measured and simulated yearly nitrate loading at Station 2A during the validation period (1994–2002).
flow (1994–2002) under BMP4 simulated in this study (411 mm) corresponds to cumulative probability of 0.67 as depicted by an arrow in Fig. 4, which means that there is 67% confidence that stream flow is equal or smaller than 411 mm. In other words, probability of stream flow being greater than 411 mm is 33%. It is also apparent that annual streamflow of 1004 mm for a wet hydrologic year (1996) was way outside the probability distribution developed for Warner Creek watershed. This indicates the unusual effects of extreme climatic condition on our models that are generally calibrated and validated for average conditions. Similar probability distributions like in Fig. 4 can provide the reliability of certain quantitative value and associated risk for each model output, which could be used in TMDL assessment and economic analysis of BMPs.

The annual sediment loading in 1996 under BMP4 was 13 000 kg ha$^{-1}$ with about 84% certainty (Fig. 5). The year 1996 was an extremely wet hydrologic year with precipitation exceeding 1800 mm (almost two times that of the normal amount in Maryland). One should be cautious that such an extreme condition often yields more pollutant loadings and creates additional environmental concerns. The cumulative probabilities of annual NO$_3$-N loadings in 1996 under BMP4 were about 39% and 43% with winter crop and without winter crop, respectively (Fig. 6). Such results indicate that using average input parameter values without considering their variability produces simulation outputs that are less than 100% certain to be averages themselves but can be below the true mean (50% point on the cumulative probability function). Therefore, in simulating BMP effects and development of TMDL plans for a watershed, one has to consider output uncertainty due to input variability. If the effectiveness of each BMP is associated with certain reliability, such information can be great aid to prescribe an economically and environmentally feasible BMP or a set of BMPs to resolve the pollution problem.
Conclusions

Overall, this study concluded that SWAT is a reasonable annual predictor of the watershed responses for assessing the impacts of different management systems on water supplies and nonpoint source pollution. However, it fails to do reasonable predictions on short time steps such as daily or monthly basis. Our previous studies and BMP evaluations conducted in this study indicate that simulations results of SWAT model are highly affected by the variability in input parameter values. Inherent spatially heterogeneity in soils, land use, and management scenarios produce variability in physical, hydraulic, and chemical parameter used in the SWAT model, thus they should be represented by appropriate probability distribution functions (pdfs). Our previous study defined appropriate pdfs for these parameters, then, used LHS with constrained MCS techniques to generate a select number of model simulations. The results were then plotted as cumulative probability distribution for each constituent of interest. The SWAT model was also used with average input parameter values without considering their variability. Results indicated that deterministic model output for BMP4 (contour strip cropping with no-till) falls within the range of the cumulative probability distribution for the watershed for each constituent of interest but is not necessarily the mean output. This study concluded that one should consider the stochastic nature of the input parameter values in simulating hydrologic and water quality response of any watershed. Results such as those reported here may provide guidelines and warnings regarding the use of these models such as SWAT model in TDML assessment and planning.

References


Arnold J.G., Williams J.R., Nicks A.D. & Sammons N.B. 1990. SWRRB, a basin scale simulation model for soil and water resources management. Texas A&M University Press, College Station, Texas.

Blaszczynski J. 2003. Estimating watershed runoff and sediment yield using a GIS interface to curve number and MUSLE models. Soils/Geology, Resources Notes 66, National Science and Technology Center, Denver CO.


Hägerstrand T. 1992a. Landskapsmanteln. Input to a dia-

Fig. 6. Model output distribution of NO$_3$-N loading at the watershed outlet (Sohrabi et al. 2003) and NO$_3$-N under BMP4 based on average input parameter values for 1996 (wet year).


