Post-validation of SWAT model in a coastal watershed for predicting land use/cover change impacts
Harsh Vardhan Singh, Latif Kalin, Andrew Morrison, Puneet Srivastava, Graeme Lockaby and Shufen Pan

ABSTRACT
Watershed models are typically calibrated and validated with the same land use and land cover (LULC) dataset and later used in assessing impacts of changing LULC, such as urbanization, on hydrology and/or water quality. However, their performance in predicting water quality/quantity in response to changing LULC is rarely assessed. The main objective of this paper was to explore the performance of the soil and water assessment tool (SWAT) in predicting water quality and quantity in response to changing LULC in a coastal watershed in Alabama, USA. Using the 1992 LULC as the input, the model was calibrated and validated for flow for the period 1990–1998, and for total suspended solids (TSS), nitrate (NO3⁻/C0), and organic phosphorus for the period 1994–1998 at several sites within the watershed. The model was then driven with the 2008 LULC data and its performance in predicting flow and TSS, NO3⁻, and total-P loads during the period 2008–2010 was evaluated (post-validation). SWAT showed good performance in predicting changes in flow and water quality during the post-validation period. The study also highlighted the importance of using the most up-to-date LULC data for effectively predicting the impacts of LULC changes on water quality.

Key words | land use, modelling, SWAT, urban development, water quality, watershed

INTRODUCTION
Land use and land cover (LULC) have direct effects on various hydrologic processes such as infiltration, interception loss, and evapotranspiration, and therefore changes in LULC can alter watershed hydrology. Studies have shown that urbanization decreases the amount of water that infiltrates into the soil (Dunne & Leopold 1978; Klein 1979), causes quicker and larger pulses in the flow hydrograph (Dunne & Leopold 1978; Neller 1988), increases surface runoff and decreases base flow (Zhou et al. 2013) and increases the frequency of extremes, both at the high and low-flow end (Lazar 1990; Rose & Peters 2001). LULC have also been shown to affect water quality (Nelson & Booth 2002; Schilling & Spooner 2006). Various studies have shown that an increase in impervious surface area as low as 10% can result in stream degradation (Schueler 1995; Bledsoe & Watson 2001). High proportions of impervious surfaces can also lead to increased loadings of nutrients and sediments (Harden 1992; Nelson & Booth 2002) and heavy metals (Callender & Rice 2000) to streams.

When developing future watershed management plans for the protection of water resources (both for quality and quantity), it is imperative to consider the implications of potential LULC changes. This requires understanding the nexus between LULC and water quality/quantity. One way of establishing this relationship is through long-term water quality monitoring, which is an expensive and time-consuming process. As a result, watershed models are commonly utilized to assess or forecast the impacts of LULC changes on water quality/quantity (LeBlanc et al. 1997; Bhaduri et al. 2000; King & Balogh 2001; Tang et al. 2005; Tu 2009). The conventional practice in such studies is to calibrate and validate a model based on current (or past) LULC conditions and observations, and then use this calibrated model to predict future water quality/quantity due
to projected changes in LULC (Kalin & Hantush 2006b; Kalin & Hantush 2009; Tong et al. 2009). While Klemes (1986), in his paper on operational testing of hydrologic models, proposed model ‘post-validation’ (he actually uses the term ‘model transportability’) as a required step; whenever, a model is to be used to simulate flows under different conditions, very few studies essentially followed this blueprint in the field of watershed hydrology. Studies assessing the impacts of LULC on water quality fall under this category. Reliability of a model without post-validation, therefore, cannot be fully established.

The process-based soil and water assessment tool (SWAT) is widely used for evaluating changes in water quality in response to LULC alterations and management practices (Fohrer et al. 2001; King & Balogh 2001; Vache et al. 2007; Li et al. 2009; Ullrich & Volk 2009; Ghaffari et al. 2010). Except for Li et al. (2009), the focus of which was annual runoff, studies that relied on SWAT to explore LULC impacts on flow or water quality did not perform post-validation. In other words, the impacts of LULC changes on hydrology or water quality, which are either projections or purely scenario analysis, were not assessed in terms of model performance. Hence, there is a need for performing post-validation in order to assess the ability of a calibrated and validated SWAT model in predicting the future changes in water quality in response to LULC changes.

The overarching goal of this paper is to evaluate the performance of the SWAT model in predicting impacts of LULC changes on water quality in a coastal watershed in Alabama, USA. The study watershed drains to a sub-estuary of Mobile Bay in northern Gulf of Mexico. The specific objectives are: (1) calibrate and validate the SWAT model for flow and water quality (nitrate [NO₃⁻], organic phosphorus (Org-P) and total suspended solids (TSS)) in the Fish River watershed for the period 1990–1998, where the model is driven by the 1992 LULC dataset; and (2) post-validate the calibrated model by driving it with the 2008 LULC dataset and comparing the model-generated flow and water quality estimates with their observed counterparts. Below we first provide a short description of the SWAT model and then describe the study area. This is followed by the description of the calibration, validation, and post-validation procedures. After presenting and discussing the results, the paper closes with a summary and conclusions.

**METHODOLOGY**

**SWAT model**

SWAT is a watershed scale continuous time simulation model developed by the US Department of Agriculture – Agriculture Research Service to predict flow and water quality. It models physical processes related to water, sediment, and nutrients. In this study, SWAT version 2005 (Neitsch et al. 2005) was used. Model parameterization and processing of input and output files were carried out using ArcSWAT version 2.1.6 (Winchell et al. 2008).

SWAT employs a modified version of the Soil Conservation Service curve number method (Soil Conservation Service (SCS) 1972) for calculating surface runoff. The Penman–Monteith method was used for estimating potential evapotranspiration in this study. Other than the soil water that is taken up by plants or evaporated, the rest of the soil water either percolates into the aquifer or moves laterally in the soil profile and finally contributes to the streamflow. SWAT uses the modified universal soil loss equation for calculating sediment loading. Transformation of nitrogen (N) and phosphorus (P) is controlled during their respective cycles. Nitrogen and phosphorus cycles are connected to water, atmosphere, and land. Detailed description of flow and water quality processes used in SWAT is provided in Neitsch et al. (2005), and thus is not repeated here.

**Study area**

The 398 km² Fish River watershed, located in coastal Alabama, USA, drains into Weeks Bay, which is a small sub-estuary of the Mobile Bay (Figure 1). Approximately 75% of Weeks Bay’s fresh water input comes from the Fish River with the rest mainly coming from the Magnolia River (Figure 1). Weeks Bay is one of only three designated Outstanding National Resource Waters in the state of Alabama, where no direct point sources of discharges are permitted. However, this pristine bay is under the threat of
water quality impairment (Alabama Department of Environmental Management 2010). In the Weeks Bay watershed, there has been a significant conversion of natural land cover to urban development over the past two decades. Baldwin County has the second fastest population growth rate in the state of Alabama. The assessment of the impact of these LULC changes on the quantity and quality of the major fresh water supplier of the bay, the Fish River, is therefore, of paramount importance.

Water quality data have been collected in the past at different times in the tributaries of the Fish River from 1994 to 2001 by Basnyat (1998) and Geological Survey of Alabama, GSA (Neil & Chandler 2003). Hence, we found a unique opportunity in this watershed for assessing the SWAT model’s predictive capability of LULC change impacts on water quality. The watershed lies in two physiographic districts, the Southern Pine Hills and the Coastal Lowlands. Table 1 depicts major LULC changes from 1992 to 2008 in the Fish River watershed. A large fraction of the watershed was covered with pasture in the early 1990s, which has since been gradually converted into agricultural and medium-density residential housing due to

![Figure 1](image-url)
urbanization and the expansion of sod farming. A major portion of the watershed is covered with sandy loam soil. The remainder is covered mainly with either sandy or loamy soil.

Model inputs

Input datasets required by ArcSWAT include a digital elevation model, LULC, soil, and climate data (Figure 2). National Elevation Dataset with a one-third arc second resolution (10 m pixels) was used for delineating the Fish River watershed and the drainage area of the sampling sites. The county level Soil Survey Geographic dataset was used in deriving soil parameters. Rainfall data were available from three climate stations providing good coverage (Figure 1). Temperature data were available only from the Robertsdale station (Figure 1). Wind speed, relative humidity, and solar radiation data were obtained from the Mobile Regional Airport (~50 km west of the study area).

Two different LULC datasets were used as input for creating two different models. The 1992 National Land Cover Data (NLCD) were the most representative LULC dataset for the water quality data and were used to calibrate and validate the SWAT model. LULC data for 2008, which were specifically developed for this study (details given below), were used as input for the second model to perform post-validation during 2008–2010. In 1992, the watershed was dominantly pasture, forest, row crops, and urban. Owing to its proximity to Gulf of Mexico beaches, Weeks Bay watershed attracts a large number of people every year, which has led to increases in medium-density residential, large density residential and commercial areas. The urban area has increased from 2.1% in 1992 to 22.8% in 2008. Table 1 shows the changes in LULC from 1992 to 2008 at the sub-watersheds draining to various sampling sites within the Fish River watershed. The general trend is increase in urban and reduction in pasture land.

LULC data for 2008 were developed from the Landsat TM image acquired on 25 March 2008. Unsupervised classification was performed, producing 100 spectral clusters. Each spectral cluster was visually checked against the Landsat imagery as well as the ancillary data such as aerial photographs, existing LULC data, and the national wetland inventory, and was labeled with the land cover type it represented. Post refinements were performed, especially for developed areas, including commercial/transportation/industrial, high residential, medium residential, and low residential by comparing the original TM images with the aerial photographs of 2005 and LULC data of 2005 developed by Baldwin County as well as ground-truthed data. The misclassified areas were manually corrected. An overall accuracy of 85.4% was achieved.

Table 1 | LULC change from 1992 to 2008 at the studied sites

<table>
<thead>
<tr>
<th>Sites:</th>
<th>G4</th>
<th>G5</th>
<th>G6</th>
<th>G7</th>
<th>G9</th>
<th>G10</th>
<th>USGS Outlet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drainage area (km$^2$)</td>
<td>30.8</td>
<td>10.5</td>
<td>12.7</td>
<td>24.8</td>
<td>43.9</td>
<td>5.5</td>
<td>143.8</td>
</tr>
<tr>
<td>Forest (%)</td>
<td>1992</td>
<td>20.1</td>
<td>5.8</td>
<td>19.9</td>
<td>21.6</td>
<td>61.5</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>18.3</td>
<td>6.2</td>
<td>19.6</td>
<td>21.9</td>
<td>38.9</td>
<td>11.0</td>
</tr>
<tr>
<td>2008–1992</td>
<td>−1.8</td>
<td>+0.4</td>
<td>−0.3</td>
<td>+0.3</td>
<td>−22.6</td>
<td>−3.4</td>
<td>−7.1</td>
</tr>
<tr>
<td>Cropland (%)</td>
<td>1992</td>
<td>24.9</td>
<td>32.3</td>
<td>22.7</td>
<td>28.0</td>
<td>9.5</td>
<td>25.9</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>28.9</td>
<td>73.3</td>
<td>33.4</td>
<td>38.9</td>
<td>10.6</td>
<td>44.9</td>
</tr>
<tr>
<td></td>
<td>2008–1992</td>
<td>+4.0</td>
<td>+41.0</td>
<td>+10.7</td>
<td>+10.9</td>
<td>+1.1</td>
<td>+19.0</td>
</tr>
<tr>
<td>Past (%)</td>
<td>1992</td>
<td>49.8</td>
<td>56.1</td>
<td>49.3</td>
<td>37.8</td>
<td>16.0</td>
<td>53.0</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>7.2</td>
<td>8.6</td>
<td>20.2</td>
<td>7.4</td>
<td>5.6</td>
<td>22.2</td>
</tr>
<tr>
<td></td>
<td>2008–1992</td>
<td>−42.6</td>
<td>−47.5</td>
<td>−29.1</td>
<td>−30.4</td>
<td>−10.4</td>
<td>−30.8</td>
</tr>
<tr>
<td>Urban (%)</td>
<td>1992</td>
<td>0.05</td>
<td>0.05</td>
<td>0.02</td>
<td>0.1</td>
<td>0.3</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>36.2</td>
<td>6.0</td>
<td>15.1</td>
<td>22.4</td>
<td>18.0</td>
<td>21.1</td>
</tr>
<tr>
<td></td>
<td>2008–1992</td>
<td>+36.1</td>
<td>+5.9</td>
<td>+15.1</td>
<td>+22.3</td>
<td>+17.7</td>
<td>+19.4</td>
</tr>
</tbody>
</table>

Other minor LULC types not included in the table are water, bare land, scrublands, grass, forested wetland, shrub wetland, emergent wetland.
Model calibration and validation

Before model calibration, the literature was first consulted to identify a set of sensitive parameters (Santhi et al. 2001; Srivastava et al. 2006). Then, an informal sensitivity analysis was performed to identify the most sensitive parameters (see Table 2). Only those parameters listed in Table 2 were calibrated. In this study, manual calibration was preferred over automated calibration mainly because of our previous experience with the SWAT model in the studied region (i.e. expert knowledge) (see, for example, Singh et al. 2011; Wang & Kalin 2011; Niraula et al. 2012; Wang et al. 2014). While autocalibration can produce better model performances, it can also yield parameter estimates that do not represent the actual conditions of the watershed (Van Liew et al. 2003). Both quantitative and qualitative measures were used during model calibration. We tried to maximize the model performance statistics (more on these below) at different time scales. For flow we also checked the flow duration curve to make sure that the model is performing well in predicting both low and high flows. The flow component of the SWAT model was calibrated and validated at the US Geological Survey (USGS) site (Figure 1) for the time periods January 1990–December 1994 and January 1995–December 1998, respectively. A 10-year warm up period was used to minimize the uncertainties associated with the initial unknown conditions, such as antecedent soil moisture and groundwater level.

Spatial calibration/validation was performed for water quality (TSS, NO\textsubscript{3}/C\textsubscript{0}\textsubscript{3}, and Org-P) due to the nature of the data. Water quality data were only available from several tributaries of the Fish River (Figure 1) for the period 1994–1998 (Basnyat 2012; Neil & Chandler 2003). Unfortunately, no past water quality data were available at the USGS site. Therefore, split sampling was done in the spatial domain. Sites G5, G7, and G9 were used for model calibration, whereas sites G4, G6, and G10 were used for validation. These sites were selected in such a way that each represented a different LULC type. While watershed G5 had more than 56% pasture land, watershed G7 had a higher percentage of agricultural land, and G9 had the highest forest coverage comprising more than 62% of its watershed area (Table 1). Since no hydrologic or water quality data were available at the watershed outlet (due to tidal influence; which is why USGS has its site further upstream at its current location) no calibration, validation, or post-validation was carried out at the main outlet. The outlet is shown in Figure 1 because it is the pouring point to the Weeks Bay.

Although SWAT can be run at an hourly time scale (requires accurate hourly climate data), a review of the literature reveals that results at hourly time resolution are not embraced yet by the scientific community. Applications of SWAT are typically at daily, and more commonly at monthly time scale (Saleh et al. 2000; Spruill et al. 2000; Santhi et al. 2001; Van Liew et al. 2007). On the other hand, due to the constraint of available resources, water quality is often only measured instantaneously. While we made an attempt to compare the instantaneous observed results with the model
simulated average daily concentrations, this resulted in large discrepancies. For this reason, observed instantaneous water quality data were converted into continuous monthly loadings using the LOADEST software (Runkel et al. 2004), which has widely been used in water quality studies (Dornblaser & Striegl 2007; Maret et al. 2008). The adjusted maximum likelihood estimation method within LOADEST was used for estimating the continuous loads. LOADEST requires flow data for estimating the water quality loadings. Since flow data were not available at the tributaries where water quality data were available, model simulated flows were used to obtain monthly water quality loadings. It was assumed that once the model is calibrated for flow at the USGS gauge, which drains more than half of the Fish River watershed and is strategically located in the middle of the watershed, it will provide a reasonable flow estimate at sub-watershed level. In a study conducted in Iran, Noor et al. (2014) found that the SWAT model calibrated using data collected at the outlet of the watershed was able to simulate streamflow at multiple gauging stations within the watershed.

During the 1994–1998 period, 59 water quality samples had been collected from the sampling sites and all of the data were used in developing the continuous loading data with LOADEST. Twenty more samples were collected and analyzed for water quality during the post-validation period through automated and grab sampling. Flow was also continuously monitored during this period. Water quality data were collected and analyzed for TSS, NO$_3$ and total phosphorus (TP). Details of data collection and analysis during the post-validation period can be found in Morrison (2011). Samples were collected using automated and grab sampling. Samples were taken during both baseflow and stormflow periods in order to form a flow gradient, which is necessary for reliable LOADEST estimates. Model performance was evaluated using mass balance error (MBE), coefficient of determination ($R^2$), and Nash-Sutcliffe efficiency ($E_N$) as given in Kalin & Hantush (2006a). Table 2 summarizes calibrated model parameters.

### Post-validation

In this step, the calibrated and validated SWAT model was used to predict flow and water quality during the period 08 October–10 March. For this, LULC data representing the year 2008 were used as an input to ArcSWAT in place

---

**Table 2 | Calibrated SWAT parameters along with their default values**

<table>
<thead>
<tr>
<th>Process</th>
<th>Parameters</th>
<th>Description</th>
<th>Calibrated values</th>
<th>Default values</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow</td>
<td>$cn2$</td>
<td>CN</td>
<td>$-10%^a$</td>
<td>0</td>
<td>0–100</td>
</tr>
<tr>
<td></td>
<td><em>Esco</em></td>
<td>Soil evaporation compensation factor</td>
<td>0.9</td>
<td>0.95</td>
<td>0.01–1.0</td>
</tr>
<tr>
<td></td>
<td><em>gw_delay (days)</em></td>
<td>Groundwater delay</td>
<td>150</td>
<td>31</td>
<td>0–100</td>
</tr>
<tr>
<td></td>
<td><em>Gaqmnn (mm)</em></td>
<td>Threshold depth of water in the shallow aquifer required for return flow to occur</td>
<td>4</td>
<td>0</td>
<td>0–500</td>
</tr>
<tr>
<td></td>
<td><em>Gaarevap</em></td>
<td>Groundwater revap coefficient</td>
<td>0.2</td>
<td>0.02</td>
<td>0.02–0.2</td>
</tr>
<tr>
<td></td>
<td><em>Surlag</em></td>
<td>Surface runoff lag coefficient</td>
<td>1.5</td>
<td>4</td>
<td>0–10</td>
</tr>
<tr>
<td>TSS</td>
<td><em>Spcon</em></td>
<td>Linear parameter for estimating sediment routing in channel</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0001–0.1</td>
</tr>
<tr>
<td></td>
<td><em>Spexp</em></td>
<td>Exponential parameter for estimating sediment routing in channel</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0–2.0</td>
</tr>
<tr>
<td>Nitrate</td>
<td><em>Cmn</em></td>
<td>Rate factor for humus mineralization of organic matter</td>
<td>0.003</td>
<td>0.0003</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td><em>Nperco</em></td>
<td>Nitrogen percolation coefficient</td>
<td>0.02</td>
<td>0.2</td>
<td>0.01–1.0</td>
</tr>
<tr>
<td></td>
<td><em>Rsdeo</em></td>
<td>Residue decomposition coefficient</td>
<td>0.1</td>
<td>0.05</td>
<td>0.01–0.05</td>
</tr>
<tr>
<td></td>
<td><em>Sdnco</em></td>
<td>Denitrification threshold water content</td>
<td>1</td>
<td>1.10</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td><em>sol_no3 (mgN/Kg soil)</em></td>
<td>Initial NO$_3$ concentration in the soil layer</td>
<td>0$^b$</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Org-p</td>
<td><em>Phoskd</em></td>
<td>Phosphorus soil partitioning coefficient</td>
<td>80</td>
<td>175</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td><em>Psp</em></td>
<td>Phosphorus availability index</td>
<td>0.5</td>
<td>0.40</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td><em>sol_orgp (mgP/Kg soil)</em></td>
<td>Initial organic P concentration in soil layer</td>
<td>60</td>
<td>0</td>
<td>–</td>
</tr>
</tbody>
</table>

*aPercentage change.

$b$For G9 it is 1.0.
of the previously used 1992 NLCD data. While the calibration process is used to determine and adjust the sensitive parameters in order to bring model simulated flow and water quality values close to their observed counterparts, it is equally important to transfer these calibrated model parameters in a systematic way when the model is set up with a different LULC. This was the case during post-validation of the SWAT model, where LULC from 2008 replaced the 1992 LULC. For this purpose, calibrated model parameters were transferred according to the soil and LULC combinations (see Figure 2). In other words, there was no new model calibration. The same parameter values obtained during the calibration period were used during the post-validation period, too. For instance curve number (CN) was reduced by 10%, and ground water delay was kept as 150 days (Table 2). The model was also run using 1992 LULC data to visualize the differences in model simulations while using two different LULC datasets from 1992 and 2008. This was done to show the importance of using the most up to date and accurate dataset in modelling. Model simulations were compared with observed flow at the USGS site and water quality data at sites G4, G6, G7, and G9 to test whether SWAT is a reliable tool in predicting the effect of LULC changes. These four sites were selected based on the availability and quality of the data.

RESULTS AND DISCUSSION

Calibration and validation

Flow

Model simulated and observed streamflow at the USGS gauge are shown in Figure 3 at a monthly time scale for the calibration (1990–1994) and validation (1995–1998) periods. Model simulations closely followed observed values in both periods (Figure 3) with $R^2 > 0.80$ and $\text{MBE} < 10\%$ in both calibration and validation periods.

![Figure 3](https://iwaponline.com/hr/article-pdf/46/6/837/370754/nh0460837.pdf)
SWAT predicted flow successfully at the daily time step as well with $R^2$ and $E_N$ values of 0.69 and 0.68, respectively. Flow was overestimated by 10% during the validation period. Overall, the model was able to simulate streamflow during both calibration and validation periods with high accuracy. The flow duration curves for daily flows in Figure 3 also show that SWAT performed well in simulating high and low flows.

Note again that since observed flow was not available at the water quality sampling sites, model simulated flows were used at those sites to obtain the continuous monthly loadings for TSS, NO$_3^-$, and Org-P for calibration and validation. The $R^2$ values of the regression equations predicted by the LOADEST model between SWAT generated discharge and the observed instantaneous water quality concentrations varied from 0.82 to 0.86. This provides some evidence that SWAT performed well in predicting flows at sub-watershed level. The justification behind this is the fact that observed flow and constituent concentrations are expected to be highly correlated.

**TSS**

Instantaneous TSS concentrations were available for all the sampling sites (Figure 1) for the period 1994–1998, which were converted into continuous monthly TSS loadings using LOADEST. Estimated monthly TSS loadings (henceforth called observed data) from three sites (G5, G7, and G9) were used for calibrating sediment-related parameters of SWAT. The model was validated using TSS data from sites G4, G6, and G10. TSS is usually composed of clay and silt. Therefore, clay and silt loadings from the model output were added together to get the TSS loadings from the model.

Figure 4 shows that, except for the months of July 1997 and September 1998, the model was successful in capturing TSS loadings at most of the watersheds during the calibration period (1994–1998). The underestimation in those 2 months are due to some very large storms, during which observed TSS may contain particles larger and heavier than silt and clay and even organic detritus. National Climatic Data Center recorded two floods in Baldwin County during July 1997 and September 1998. On 19 July 1997, Hurricane Danny started falling at 21:00 on Baldwin County and the Fish River flooded for several days. The highest estimated rainfall was 760–890 mm with even higher estimations near Weeks Bay. Another hurricane, Hurricane Georges, hit Baldwin County on 28 September 1998 dropping 380–655 mm of rainfall, resulting in a 25-year return interval flood for Fish River. To obtain a better understanding of model performance during the calibration period, we removed the loadings for these 2 months.

Performance statistics confirm that simulated TSS loadings adequately matched observed TSS data at sites G5 and G9 (Figure 4). Removal of July 1997 and September 1998 made a significant change in MBE. At site G5, the change in MBE was 62% (from 42% underestimation to 20% overestimation). At site G7, 62% underestimation was reduced to 50%. Change in MBE at G9 was about 14%. In general, except for some extreme events, the model was able to simulate TSS loadings accurately at the calibration sites. Similar results were obtained for the validation watersheds. TSS loadings during July 1997 and September 1998 were again underestimated substantially. High $R^2$ and $E_N$ values combined with low MBE for watersheds G4 and G6, after removing the loadings of July 1997 and September 1998, indicated that SWAT successfully predicted the TSS loadings for these watersheds (Figure 4). Although the model predicted TSS loadings adequately at site G10, the performance at this site was weaker compared to sites G4 and G6. The main channel where water quality data were collected at site G10 is narrow. During site visits we frequently observed streamflow overflowing the bank. This suggests that sediment data collected at this site are less reliable.

**Nitrate (NO$_3^-$)**

Watershed G9 was calibrated separately from G5 and G7 by adjusting the NO$_3^-$ concentration of the soil (sol_no3) to bring the final NO$_3^-$ loadings for this particular watershed within comparable range to the observed NO$_3^-$ loadings. The main soil type in watershed G9 is loamy soil which is different from the rest of the Fish River watershed, where the soil is either sandy or sandy loam.

Figure 5 graphically compares observed and simulated NO$_3^-$ loadings. $R^2$ and $E_N$ values for NO$_3^-$ loadings are lower in comparison with those for flow and TSS simulations. This is not totally unexpected as errors incurred
during calibration of flow are carried over. Watershed G7 shows a relatively good match of simulated NO$_3$ loadings with the observed values. Better model performance for this particular watershed may be attributed to the fact that this is the closest watershed to the USGS gauging station, which was used for flow calibration. A summary of the literature review presented in Moriasi et al. (2007) shows that the range of reported $E_N$ values for monthly calibration of NO$_3$ lie between $0.08$ and $0.59$, with a median of $0.26$. $E_N$ values for watersheds G5, G7, and G9 are comparable to values reported in the literature.

Figure 5 shows a small lag in the simulated NO$_3$ loadings compared with the observed loadings for watersheds G5 and G7. Low average NO$_3$ loadings for watershed G9 can be explained by the fact that about 62% of the watershed is covered with forests that can act like a sink or buffer,
therefore preventing \( \text{NO}_3^- \) loadings into the tributary of Fish River. The SWAT model was able to predict \( \text{NO}_3^- \) loadings at the validation sites G6 and G10 at an adequate level (Figure 5). At G4, simulated loadings were higher compared with observed data. Since we do not have detailed knowledge on fertilizer application rates and practices or any management practices that may have existed during this period, we have no clear explanation for these anomalies.

**Org-P**

The only available P data were Org-P concentrations. Watersheds G5, G7, and G9 were used for calibration, while the watersheds G4, G6, and G10 were used for validation. Table 2 presents all the parameters that were adjusted for calibrating Org-P loadings. Similar to \( \text{NO}_3^- \), special calibration was needed for watershed G9 where the model
was significantly underestimating Org-P content. To bring the simulated value into a comparable range with the observed data for watershed G9, soil Org-P content was raised to 60 mg/kg. Again, this is most likely due to the differences in soils. Earlier we excluded July 1997 and September 1998 data for sediment simulation because during large storms observed TSS may contain particles larger and heavier than silt and clay. The reason for not excluding July 1997 and September 1998 data from phosphorus simulation is that phosphorus is mainly carried by fine-grained sediment particles (Stone & Mulamoottil 1995) and larger particles are less likely to carry phosphorus.

Calibration sites G7 and G9 show a good match between observed and simulated loadings of Org-P (Figure 6). Model performance measures at the calibration sites show that SWAT slightly over-predicted Org-P loadings except at G9, where the model underestimated Org-P loading by approximately 38%. It is evident from Figure 6 that

![Figure 6](https://iwaponline.com/hr/article-pdf/46/6/837/370754/nh0460837.pdf)

**Figure 6** | Comparison of simulated and observed monthly Org-P loadings for sites G5, G7, and G9 during calibration and for sites G4, G6, and G10 for validation.
the model is overestimating Org-P during high flows. Average loadings for Org-P were between 0.11 and 0.22 kg/km²/day suggesting low phosphorus content at all the sampling watersheds. Model performances at validation sites were also good, especially at sites G4 and G10.

Post-validation (2008–2010)

In this section, the calibrated/validated SWAT model was further tested for its capabilities in predicting effects of changes in LULC on water quality. While the sub-watersheds and the stream network for the Fish River watershed remained the same for both LULC datasets of 1992 and 2008, hydrological response units (HRUs: the smallest computational element in SWAT having uniform soil, LULC, and slope) differed. Comparison of the model simulations for flow, TSS, NO₃⁻, and TP with their observed counterparts for the period October 2008–January 2010, at different sampling sites are presented below.

From 1992 to 2008, all sub-watersheds experienced a reduction in pasture land and an increase in urban and agricultural land (Table 1). Percentage forest cover changed significantly at only one sub-watershed. Based on observed water quality data, the resultant effects of these LULC changes, in general, were increased phosphorus and reduced nitrate loadings. TSS loadings increased in G4 (pasture → urban), G6 (pasture → urban + agg), and G9 (forest → urban).

Flow

Although flow was measured at the sampling sites, due to lack of confidence in flow rating curves (insufficient flow measurements to develop a reliable Q-h relationship) they were not used in this study. Therefore, post-validation of SWAT for flow was limited to the USGS site. Simulated flow values obtained from the model using 1992 and 2008 LULC datasets at the USGS gauge were compared with the observed flow data. R², Eₙ, and MBE values obtained using the 2008 LULC data suggested a good match between the observed and simulated flow values (Figure 7). Although the use of 1992 LULC data also resulted in satisfactory model performance (values in parenthesis in the figure), the 2008 LULC clearly provided better flow estimates.

There were significant changes in LULC from 1992 to 2008 due to conversion of pasture and forested area into mainly urban areas. As a result, the use of 2008 LULC produced better model performance.

TSS

Figure 8 shows the comparison between observed and simulated TSS loadings using 1992 and 2008 LULC datasets. Except for March 2009, the model simulated TSS loadings with good accuracy at all sites using the 2008 LULC data. There was a severe storm on 29 March 2009 which caused the USGS gauge on Fish River to record a stage 1.2 m above the flood level. Large storms such as this could also erode channels and bring in additional sediment from the floodplains. Typically, erosion and sediment deposition occur simultaneously in small events, which can be easily flushed out by a large storm. The inability of SWAT to accurately predict sediment in such events was also evidenced during the calibration stage. Figure 8 shows that the model did not predict TSS loadings accurately using the 1992 LULC data compared with simulations using 2008 LULC data. With the 1992 LULC data, the model underestimated monthly average TSS loadings by more than 70% for watersheds G4, G6, and G9. Overall, the model clearly performed better with the 2008 LULC dataset.
Nitrate

The model performed well for NO$_3$ loadings at all sites from May 2009 onward. Before May 2009, there appears to be a time lag between model simulations and observed data (Figure 9). When the whole period is considered, predicted average NO$_3$ loadings did not deviate more than 40% from the observed NO$_3$ loadings. The model did a relatively good job with the 1992 LULC data, too, in simulating NO$_3$ loadings at site G9. $R^2$, $E_N$, and MBE shown in Figure 9 indicate that SWAT was not able to simulate NO$_3$ as accurately as flow and TSS. This was expected, as model performances in simulating NO$_3$ loadings during the calibration and validation periods were also lower than the model performances in simulating flow, TSS, and Org-P. Overall, the model performance in simulating NO$_3$ loadings decreased compared with the calibration and validation periods. White & Chaubey (2005) also found that SWAT performance was less satisfactory for nitrates and nitrites during their study for a multisite watershed in Arkansas. However, when considering the effect of change in LULC data as input and other unknowns (e.g. actual fertilizer application rates) model simulations can be deemed acceptable.

TP

During this period, TP is used instead of Org-P, which was used during calibration and validation. This was because the laboratory facilities we utilized were only equipped to analyze TP. $R^2$ and $E_N$ values show that the model did a better job in predicting TP loadings than simulating NO$_3$
loadings using 2008 LULC data (Figure 10). The MBE using 1992 LULC data was higher than 80% for three out of four sites that were used for testing the accuracy of the model simulations. Figure 10 shows that although simulated TP loadings using 2008 LULC data closely follow the pattern of loadings in different months, the model underestimated TP loadings at all sites. This could be the result of the fact that the model underestimated TSS during the post-validation period (Figure 8), resulting in underestimation of particulate phosphorus and subsequently the underestimation of TP.

### SUMMARY AND CONCLUSIONS

Watershed models are commonly used in assessing the impacts of LULC changes on hydrology and water quality. They are typically calibrated using existing or past conditions and are then used in forecast mode to analyze various LULC change scenarios. In the absence of data, they are not calibrated at all. In this study, the reliability of the SWAT model in predicting water quality response to LULC changes was explored in a coastal watershed in Alabama, USA. The model was first calibrated and validated for flow, TSS, nitrate, and org-P with 1992 LULC. The calibrated model then was fed with 2008 LULC and the model-produced outputs of flow, TSS, nitrate, and TP were compared with monthly estimated loadings (post-validation). Results showed that the model was able to simulate flow and water quality satisfactorily during calibration, validation, and post-validation periods. The following observations and conclusions are drawn:

1. Model simulations systematically under-predicted TSS loadings during extreme events. Therefore, particles
larger than silt and clay should be accounted for in the simulated results, in order to get better estimates for simulated TSS loadings during extreme events.

2. The model over-predicted TP during the post-validation period, highlighting the importance of having detailed measurements on different P components (organic and mineral forms).

3. SWAT performed less well with the 1992 LULC data compared with the 2008 LULC data during the post-validation period. This establishes the importance of using the most accurate and up-to-date LULC data that better represent the changes in LULC for efficiently using SWAT for predicting the effect of LULC changes.

4. Having good estimates of fertilizer use would improve model estimates. This is particularly the case in agricultural areas, where lumping all the agricultural areas into a single category is an oversimplification.

In summary, this study showed that SWAT could be a reliable tool in predicting potential changes in water quality/quantity in watersheds undergoing LULC changes or to project future water quality/quantity conditions under different land development or watershed management plans. However, it should be noted that this is just one study conducted in coastal Alabama. Similar studies under different physiographic and climatic regions can further strengthen these conclusions.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the funding provided by the Mississippi-Alabama Sea Grant Consortium (MASGC) and the Center for Environmental Studies at the Urban-Rural Interface (formerly Center for Forest Sustainability).
through the Peaks of Excellence program at Auburn University.

REFERENCES


Basnyat, P. 1998 Valuation of Forested Buffers. PhD Dissertation, Auburn University, Auburn, AL, USA.


Morrison, A. 2001 Spatial and Temporal Trends and the Role of Land Use/Cover on Water Quality and Hydrology in the Fish River Watershed. MS thesis, Auburn University, Auburn, AL, USA.


Schueller, T. 1995 The importance of imperviousness. Watershed Protection Techniques 1, 100–111.


Ullrich, A. & Volk, M. 2009 Application of the soil and water assessment tool (SWAT) to predict the impact of alternative management practices on water quality and quantity. Agricultural Water Management 96, 1207–1217.


First received 12 December 2014; accepted in revised form 3 March 2015. Available online 26 March 2015.