DEVELOPMENT AND APPLICATION OF A PHYSICALLY BASED LANDSCAPE WATER BALANCE IN THE SWAT MODEL

Eric D White¹, Zachary M. Easton¹, Daniel R. Fuka¹, Amy S. Collick², Enyew Adgo³,⁴, Matthew McCartney⁵, Seleshi B. Awulachew⁵, Yihenew G. Selassie³,⁴, and Tammo S. Steenhuis¹,²

¹Dept. of Biological and Environmental Engineering, Cornell University, Ithaca, NY, USA
²Faculty of Engineering, Bahir Dar University, Bahir Dar, Ethiopia
³Faculty of Agriculture, Bahir Dar University, Bahir Dar, Ethiopia
⁴Amhara Regional Agricultural Institute, Bahir Dar, Ethiopia
⁵International Water Management Institute, Nile Basin and East Africa Office, Addis Ababa, Ethiopia
ABSTRACT

Watershed scale hydrological and biogeochemical models rely on the correct spatial-temporal prediction of processes governing water and contaminant movement. The Soil and Water Assessment Tool (SWAT) model, one of the most commonly used watershed scale models, uses the popular Curve Number (CN) method to determine the respective amounts of infiltration and surface runoff. While appropriate for flood forecasting in temperate climates, the CN method has been shown to be less than ideal in many situations (e.g., monsoonal climates and areas dominated by variable source area hydrology). The CN model is based on the assumption that there is a unique relationship between the average moisture content and the CN for all hydrologic response units, and that the moisture content distribution is similar for each runoff event, which in many regions is not the case. A physically based water balance was developed and coded in the SWAT model to replace the CN method of runoff generation. To compare this new water balance SWAT (SWAT-WB) to the original CN based SWAT (SWAT-CN), two watersheds were initialized: one in the headwaters of the Blue Nile in Ethiopia and one in the Catskill Mountains of New York State. In the Ethiopian watershed streamflow predictions were significantly better using SWAT-WB than SWAT-CN (Nash-Sutcliffe efficiencies (NSE) of 0.76 and 0.67, respectively). In the temperate Catskills, SWAT-WB and SWAT-CN predictions were approximately equivalent (NSE>0.5). Interestingly, and perhaps more importantly, the spatial distribution of runoff generating areas differed greatly between the two models, with SWAT-WB providing a more realistic distribution of saturated and thus runoff source areas. These results suggest that the addition of a water balance in SWAT significantly improves model predictions in monsoonal climates, and provides equally acceptable levels of
accuracy in stream flow prediction under temperate northeastern USA conditions.

Spatially distributed watershed areas are predicted realistically with SWAT-WB.
Introduction

Non-point source runoff can contribute significant quantities of nutrients, chemicals, and sediments to stream and water bodies. To locate these “non-point” sources of pollution in a landscape, many watershed managers and researchers frequently use watershed scale models. One of the most commonly used watershed scale models is the US Department of Agriculture (USDA) Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998).

SWAT, like any water quality model, must first accurately simulate hydrologic processes before it can realistically predict pollutant transport. Many different approaches to modeling hydrologic processes have been presented in the scientific literature over the past several decades, but SWAT currently uses two methods to model surface runoff: the Curve Cumber (CN) (USDA-SCS, 1972) and the Green-Ampt routine (Green and Ampt, 1911). The Green-Ampt method is a physically-based infiltration excess, rainfall-runoff model. Therefore, Green-Ampt is not suitable for regions where the rainfall rate seldom exceeds the saturated conductivity of the soil, such as in the Northeastern USA (Walter et al., 2000). As a result, the empirically based CN method, due to its ease of use and simplifying assumptions, is the most commonly used runoff routine in the SWAT model (King et al., 1999; Gassman, 2005).

SWAT and other CN-based models are frequently used on watersheds around the world where the climate and landscape vary greatly from that of the United States, where the empirical relationships used in the CN were developed.
SWAT has been applied in areas ranging from China, India, Australia, the UK, France, Belgium, Algeria, Tunisia, Italy, and Greece with little recognition that the underlying runoff calculations were neither developed nor validated for these regions (Gassman et al., 2007).

Another region where CN models have been applied is in the Blue Nile Basin of Africa. Located in the monsoonal climate of the Ethiopian highlands, the temporal runoff dynamics in the Blue Nile Basin are poorly captured by the CN method, which assumes that the moisture content distribution of the watershed is similar for each runoff event (Steenhuis et al., 2009; Collick et al., 2009). Previous work has shown that for a given amount of rain, runoff volumes will vary throughout the rainy season. Liu et al. (2008) demonstrated that in the Ethiopian Highlands less runoff was generated at the beginning of the rainy season as compared to the same rain event at the end of the season. Lui et al. (2008) showed that when watershed discharge is plotted against effective precipitation (i.e., precipitation minus potential evapotranspiration) there is a relatively strong, linear relationship, indicating that the proportion of the rainfall that became runoff was constant during the remainder of the rainy season. These dynamics cannot be predicted by the CN, and are, in fact, in direct contrast to the official method’s literature, which states that there is no correlation between antecedent precipitation and a watershed’s maximum retention beyond five days (NRCS, 2004).

In order to correct the CN to better work in monsoonal climates, various temporally-based values and initial abstractions have been suggested. For instance, Bryant et al. (2006) suggest that a watershed’s initial abstraction
should vary as a function of storm size. While this is a valid argument, the introduction of an additional variable reduces the appeal of the one-parameter CN model. Kim and Lee (2008) found that SWAT was more accurate when CN values were averaged across each day of simulation, rather than using a CN that described moisture conditions only at the start of each day. White et al. (2009) showed that SWAT model results improved when the CN was changed seasonally to account for watershed storage variation due to plant growth and dormancy. Wang et al. (2008) improved SWAT results by using a different relationship between antecedent conditions and watershed storage. While these variable CN methods improve runoff predictions, they are not easily generalized for use outside of the watershed they are tested for due to the fact that the CN method is a statistical relationship and is not physically based.

In many regions, surface runoff is produced by only a small portion of a watershed that expands with an increasing amount of rainfall. This concept is often referred to as a variable source area (VSA); a phenomenon actually envisioned by the original developers of the CN method (Hawkins, 1979), but never implemented in the original CN method as used by the NRCS. Since the method’s inception, numerous attempts have been made to justify its use in modeling VSA-dominated watersheds. These adjustments range from simply assigning different CNs for wet and dry portions to correspond with VSAs (Sheridan and Shirmohammadi, 1986; White et al., 2009), to full reinterpretations of the original CN method (Hawkins, 1979; Steenhuis et al., 1995; Schneiderman et al., 2008; Easton et al., 2008).
To determine what portion of a watershed is producing surface runoff for a given precipitation event, the re-interpretation of CN method presented by Steenhuis et al. (1995) and incorporated into SWAT by Easton et al. (2008) assumes that rainfall infiltrates when the soil is unsaturated or runs off when the soil is saturated. It has been shown that this saturated contributing area of a watershed can be accurately modeled spatially by linking this re-interpretation of the CN method with a topographic index (TI), similar to those used by the topographically driven TOPMODEL (Beven and Kirkby, 1979; Lyon et al., 2004). This linked CN-TI method has since been used in multiple models of watersheds in the northeastern US, including the Generalized Watershed Loading Function (GWLF) (Schneiderman et al., 2007) and SWAT (Easton et al., 2008). While the re-conceptualized CN model is applicable in temperate US climates, it is limited by the fact that it imposes a distribution of storages throughout the watershed that need to fill up before runoff occurs. While this limitation does not seem to affect results in temperate climates, it results in poor model results in monsoonal climates.

SWAT-VSA, the CN-TI adjusted version of SWAT (Easton et al., 2008), returned hydrologic simulations as accurate as the original CN method, however the spatial predictions of runoff producing areas and as a result the predicted phosphorus export were much more accurate. While SWAT-VSA is an improvement upon the original method in watersheds where topography drives flows, ultimately, it still relies upon the CN to model runoff processes and therefore is limited when applied to the monsoonal Ethiopian highlands. Water balance models are relatively simple to implement and have been used frequently in the Blue Nile Basin (Johnson and Curtis, 1994; Conway, 1997;
Ayenew and Gebreegziabher, 2006; Liu et al., 2008; Kim and Kaluarachchi, 2008; Steenhuis et al., 2009). Despite their simplicity and improved watershed outlet predictions they fail to predict the spatial location of the runoff generating areas. Collick et al. (2009), and to some degree Steenhuis et al. (2009), present semi-lumped conceptualizations of runoff producing areas in the water balance model. SWAT, a more distributed (semi-distributed) model can predict these runoff source areas in greater detail, assuming the runoff processes are correctly modeled.

We propose incorporating elements from the spatially adjusted water balance models, as proposed by Lyon et al. (2004) and implemented by Easton et al. (2008), into SWAT-VSA. To accomplish this, we develop and test a CN-free version of SWAT. This new version of SWAT, SWAT-WB, calculates runoff volumes based on the available soil storage capacity of a soil, and then partitions excess moisture to runoff and infiltrating fractions. This can lead to more accurate simulation of where runoff occurs in watersheds dominated by saturation-excess processes. Both the original CN method used by SWAT and the new, water balance (SWAT-WB) method are tested on two watersheds that vary widely in climate, geology, and data availability: one in the monsoonal Blue Nile Basin in Ethiopia, and one in the Catskill Mountains of New York State.
Model Overview

Summarized SWAT Description

SWAT is a basin-scale model designed to simulate hydrologic processes, nutrient cycling, and sediment transport throughout a watershed. SWAT has been applied to catchments ranging from 0.15 km\(^2\) (Chaney et al., 2003) to as large as 491,700 km\(^2\) (Arnold et al., 2000). To initialize the model, SWAT requires soils data, land use/management information, and elevation data to drive flows and direct sub-basin routing. The hydrologic response unit (HRU) is the smallest unit in the SWAT model and is used to simulate processes such as rainfall, runoff, infiltration, plant dynamics (including uptake of water and nutrients, biomass, etc.), erosion, nutrient cycling, and leaching of pesticides and nutrients. Traditionally, HRUs are defined by the coincidence of soil type (Hydrologic Soil Group, USDA 1972) and land use. The predictions from each HRU are aggregated for each subbasin, and routed through the internal channel network. Simulations require meteorological input data including precipitation, temperature, wind, humidity and solar radiation. All of these inputs are initialized using a GIS system (ArcGIS 9.2). More detail on SWAT can be found at http://www.brc.tamus.edu/swat/doc.html.

Model Development

Original Curve Number Approach

Historically, when initializing SWAT a CN is assigned for each specific landuse/soil combination in the watershed, and these values are read into the model. SWAT then calculates upper and lower limits for each CN following a probability function described by the NRCS to account for varying antecedent
moisture conditions (CN-AMC) (USDA-NRCS, 2004). SWAT determines a CN for each simulated day by using this CN-AMC distribution in conjunction with daily soil moisture values determined by the model. This daily CN is then used to determine a theoretical storage capacity, \( S \), of the watershed for each day the model is run. The storage is then indirectly used to calculate runoff volume, \( Q \), via:

\[
Q = \frac{(P - I_a)^2}{(P - I_a) + S}
\]

where \( S \) is watershed storage, \( P \) is precipitation, and \( I_a \) is initial abstraction. All terms are in mm of water, and by convention \( I_a \) is assumed to be equal to 0.2*\( S \).

**Water Balance Approach**

A daily soil water balance was used to determine the saturation deficit of each hydrologic response unit (HRU) in SWAT, which was then used, instead of the CN method, to determine daily runoff volume. To replace the CN, a simple soil profile water balance was calculated for each day of simulation. While SWAT’s soil moisture routine greatly simplifies processes that govern water movement through porous media (in particular, partly-saturated regions), for a daily basin scale model the approach is generally acceptable (Guswa et al., 2002). Thus the model already provides a convenient platform on which to incorporate a water balance. SWAT’s inherent soil moisture routines are then used by SWAT-WB to determine the degree of saturation-deficit for each soil profile for each day of simulation. This saturation-deficit (in mm H\(_2\)O) is termed the available soil storage, \( \tau \):
where \( EDC \) is the effective depth of the soil profile (unitless), \( \varepsilon \) is the total soil porosity (mm), and \( \theta \) is the volumetric soil moisture for each day (mm). The porosity is a constant value for each soil type, whereas \( \theta \) varies by the day and is determined by SWAT’s soil moisture routines. The effective depth coefficient, \( EDC \), a parameter ranging from zero to one, is used to partition soil moisture in excess of \( \varepsilon \) into infiltrating and runoff fractions. By including this adjustment to the available storage, the amount of water able to infiltrate each day is controlled by the \( EDC \). \( EDC \) is spatially varied based on a saturation risk. \( EDC \) values approaching one are assigned to regions expected to produce little saturation excess runoff while values approaching zero indicate an area likely to produce large saturation excess volumes. This spatially adjusted available storage is then used to determine what portion of rainfall events will infiltrate and what portion will runoff:

\[
\tau = EDC(\varepsilon - \theta)
\]

\[ \text{eq. 2} \]

\[
Q = \begin{cases} 
0, & \text{if } P < \tau \\
|P - \tau|, & \text{if } P \geq \tau 
\end{cases}
\]

\[ \text{eq. 3} \]

The available storage, \( \tau \), is calculated each day prior to the start of any rain event. Once precipitation starts, a portion of the rain, equal in volume to \( \tau \), will infiltrate the soil. If the rain event is larger in volume than \( \tau \), the soil profile will saturate and surface runoff will occur. If the rainfall is less than \( \tau \), the soil is unsaturated and there will be no surface runoff and SWAT’s internal soil moisture routing will calculate the flux.
HRU Definition

Traditionally, HRUs are defined in SWAT as being unique occurrences of soil type, land cover, and slope class. Any parcels of land within one subbasin that share the same combination of these three features will be considered one HRU. SWAT models all landscape processes for each unique HRU in the watershed, independent of position within each subbasin. In basins dominated by VSA hydrology this HRU definition has been shown to be less than ideal for describing the spatial and temporal evolution of hydrologic processes (Schneiderman et al., 2007; Easton et al., 2008). In VSA watersheds, runoff-generating areas are likely to occur in portions of the landscape with shallow, low conductive soils, large contributing areas, mild slopes, or any combination of the three. While SWAT’s inclusion of slope classes in HRU delineation begins to address these issues, there is currently no way to include upslope contributing area while defining HRUs. To correct for this, a soil topographic index (STI) was integrated with existing soils data in the SWAT HRU definition process (e.g., Easton et al., 2008).

Topographic indices and their various derivatives have been used to model runoff-contributing areas for quite some time [e.g., TOPMODEL (Beven and Kirkby, 1979)]. Recently, soil topographic indices have been incorporated into CN-based watershed models for use in VSA dominated regions (Lyon et al., 2004; Schneiderman et al., 2007; Easton et al., 2008). SWAT-VSA integrated STIs into SWAT in order to improve determination of runoff-generating areas and the subsequent nutrient loads from these areas in the Catskills Mountains of New York State (Easton et al., 2008). SWAT-VSA provided more accurate
predictions of runoff source areas (as validated by distributed measures in the watershed) than the original SWAT, thus, we included the HRU definition process similar to SWAT-VSA in SWAT-WB.

To initialize SWAT-WB the first step was to create a soil topographic index for the watershed being modeled. The STI is defined as:

\[ STI = \ln\left(\frac{A}{\tan(\beta)DK_s}\right) \]  

eq. 4

The upslope contributing area, A, and the slope, \( \tan(\beta) \), are both obtained from a DEM, while the soil depth, D, and saturated hydraulic conductivity, \( K_s \), are obtained from a soil survey. We assume that STI values relate to a location’s likelihood of saturation, and therefore the likelihood to contribute surface runoff. Higher STI values are the result of either a large contributing area, or small values for slope, soil depth, or saturated conductivity, and therefore are indicative of areas with a higher probability for saturation.

Following the process outlined for SWAT-VSA (Easton et. al., 2008), an areally weighted STI (e.g., wetness classes) is used to represent a location’s likelihood to saturate. The wetness classes determined for the two watersheds used in this study are shown in Fig. 1 for the Ethiopian Watershed and Fig. 2 for the New York watershed. This wetness class map is then substituted for the soils map in the HRU definition process. While the wetness classes can be used in HRU delineation instead of a soil map, SWAT still requires specific soil properties that are commonly associated with the soils map (e.g., SSURGO Database). Thus, in SWAT-WB soil properties required by SWAT were areally weighted and averaged for each wetness class. This practice will not
drastically affect model results for two reasons. First, in Ethiopia, soil survey
information is rare or nonexistent, and, to our knowledge, no defined database
exists that would contain the parameters needed by SWAT. Thus SWAT-WB
utilized the UN-FAO's World's Soil Map (supplemented by literature values) as
the base map (FAO-AGL, 2003), which classifies only five distinct soil types in
all of the 1270 km$^2$ Blue Nile sub-catchment which was modeled. This soils
data, with such a coarse spatial resolution, will not be adversely affected by
the averaging process used in SWAT-WB. Second, in New York State, where
soils information is more readily available, soil formation (in glaciated areas) is
at least partially driven by topography (Page et al., 2005; Sharma et al., 2006;
Thompson et al., 2006; Easton et al., 2008). Therefore, averaging across
topographic features with the wetness index should not pose any problems.

**Watershed Descriptions**

**Gumera Watershed, Blue Nile Basin, Ethiopia**

SWAT-WB was tested on the Gumera River watershed, a heavily cultivated
region in the Ethiopian highlands. Located approximately 30 km northeast of
Bahir Dar (11.83°N, 37.63°E); this 1270 km$^2$ watershed drains into Lake Tana,
the headwaters of the Blue Nile River (Fig. 3). Land use in the Gumera
watershed consists of 96% agriculture and 4% brush (or pasture). Elevation
(determined from a 90 meter DEM ranged from 1797 to 3708 meters above
sea level with slopes ranging from 0% to 79%. Predominant soils were
gathered from the FAO World Soils map and were classified as haplic and
chromic luvisols (58% and 22%, respectively). Other soils present in the basin
were eutric fluvisols (8%), eutric leptosols (8%), eutric vertisols (3%), with
minimal areas classified as urban (>1%) (FAO-AGL, 2003).
Precipitation and temperature data were gathered from the National Meteorological Agency of Ethiopia for the Debre Tabor station, the closest rain gauge to the Gumera basin. Daily precipitation data from 1992 through 2003 was used for model calibration and validation. Other required climatic data included relative humidity, wind speed, and solar radiation. These data were obtained for the nearby city of Bahir Dar through the United States National Climatic Data Center (NCDC, 2007).

**Townbrook Watershed, Catskills, New York**

SWAT-WB was also tested on the Townbrook watershed (Fig. 2) in the United States; a 37 km$^2$ sub-catchment of the Cannonsville Reservoir Basin. The region is typified by steep to moderate hillslopes of glacial origins with shallow permeable soils, underlain by a restrictive layer. The climate is humid with an average annual temperature of 8ºC and average annual precipitation of 1123 mm. Elevation in the watershed ranges from 493 to 989 m above mean sea level. The slopes are quite steep with a maximum of 91%, and a mean of 21%.

Soils are mainly silt loam or silty clay loam with soil hydrologic group C ratings (USDA-NRCS, 2000). Soil depth ranges from less than 50 cm to greater than 1 m and is underlain by a fragipan restricting layer (e.g. coarse-loamy, mixed, active, mesic, to frigid Typic Fragiudepts, Lytic or Typic Dystrudepts common to glacial tills) (Schneiderman et al., 2002). The lowland portion of the watershed is predominantly agricultural, consisting of pasture and row crops (20%) or shrub land (18%) while the upper slopes are forested (60%). Water and wetlands comprise 2%. Impervious surfaces occupy <1% of the watershed and were thus excluded from consideration in the model. Several
studies in this watershed or nearby watersheds have shown that variable source areas control overland flow generation (Frankenberger et al., 1999; Mehta et al., 2004; Lyon et al., 2006a, 2006b; Schneiderman et al., 2007; Easton et al., 2008) and that infiltration-excess runoff is rare (Walter et al., 2003).

**Model Calibration**

To calibrate both the SWAT-WB and SWAT-CN models for Gumera and Townbrook, we utilized the Dynamically Dimensioned Search (DDS) algorithm (Tolson and Shoemaker, 2005). The DDS calibration routine allows for parameters to be calibrated at the watershed, subbasin, HRU, or wetness class level, which in turn allowed for $EDC$ to be calibrated separately for each wetness class. For SWAT-CN the CN was calibrated for each wetness class instead of the $EDC$. In addition 11 other hydrologic parameters were calibrated in both models.

Streamflow at the Gumera watershed outlet was calibrated over a period of eight years, from 1996 to 2003, and streamflow in Townbrook was calibrated from 1998 to 2002.

**Model Validation**

Streamflow data from 1992 through 1995 was used to validate the Gumera model. For Townbrook, streamflow data from 2002 through 2004 was used to validate the model. To test SWAT-CN’s and SWAT-WB’s abilities to correctly predict distributed hydrology, we used measurements by Lyon et al. (2006a) of height of water table above the restricting layer for a section of the Townbrook
watershed. Briefly, 44 pieziometric data loggers, installed to depths of ~50 cm, recorded the water table depth in 15 minute intervals from April 2004 to September 2004. The field site encompassed five wetness index classes (Fig. 3) and three land use types, pasture (PAST), shrub (RNGB), and mixed forest (FRST). To compare the measured and SWAT-VSA water table heights the piezometer data were averaged across index classes; there were between two and 32 piezometers per index (i.e., two piezometers on index class six, 32 on index class 10, etc). To compare measured water table heights with SWAT-CN water table heights, we averaged across land use; there were four to 32 measurements per land use. SWAT (and SWAT-WB) reports soil water in mm of water integrated over the soil profile (i.e. cumulative water depth for all soil layers). Thus, we converted the model predicted mm of soil water to and equivalent depth by dividing by the SSURGO reported porosity and assumed the SSURGO reported soil depth represented the depth to the restricting layer. According to the SSURGO data base, the depth of the local restricting layer is 1.2 – 1.4 m.

**Model Evaluation**

Criteria used to assess the ability of the models to predict discharge in Gumera and Townbrook included: a visual comparison between the modeled and the observed hydrographs, Nash-Sutcliffe Efficiencies (NSE) (Nash and Sutcliffe, 1970), and coefficient of determination, $R^2$. 

14
Results

Model Comparison

To determine if SWAT-WB was indeed a more accurate than the standard CN-based SWAT, the two models were calibrated and validated using the same automatic procedure for Gumera antiTown brook watersheds.

Gumera Basin

Following calibration, SWAT-WB returned more accurate results than both of the SWAT-CN models of the Gumera watershed (based on statistics and visual comparison of hydrographs, Fig. 4) than SWAT-CN. A daily NSE value of 0.70 for the calibration period was achieved, with an $R^2$ of 0.71. SWAT-WB accuracy increased for the validation period, with NSE and $R^2$ values of 0.76 and 0.81, respectively (Table 1 and Fig. 4). SWAT-CN model of Gumera previously published by Stegen et al (2008) can be used as a benchmark. They used a twelve year calibration period that resulted in a model with a daily NSE of 0.61 and a $R^2$ of 0.71. Validation results for their model returned a NSE of 0.61 and an $R^2$ of 0.70 similar to our results for SWAT-CN. Additionally, intra-watershed runoff producing areas were modeled with higher spatial resolution than SWAT-CN due to the inclusion of the STI-based HRU delineation process as introduced in SWAT-VSA (Easton et al., 2008).

Perhaps more interesting is how SWAT-WB and SWAT-CN differ in the predicted distribution of runoff in the watershed. For one large storm in October, 1997 (104 mm of rain), SWAT-CN predicted that all HRUs within the watershed would contribute runoff; with a minimum depth of 17 mm of runoff and a maximum of 71 mm (Fig. 5A). For the same storm, SWAT-WB predicted that some HRUs would produce no runoff, while others produced as much as...
97 mm of runoff (Fig. 5B). Both models predicted some surface runoff for some upland areas, but SWAT-CN predicted much less runoff being generated in the low-lying, flatter areas near the watershed outlet, where SWAT-WB predicted the most runoff.

**Townbrook Watershed**

SWAT-WB results for Townbrook were compared to results from SWAT-CN. Predicted streamflow for the SWAT-CN Townbrook model resulted in a daily $NSE$ and $R^2$ of 0.43 and 0.59, respectively, for calibration, and 0.62 and 0.69, respectively, for validation. Thus in the Townbrook watershed, SWAT-WB outperformed SWAT-CN during the calibration period. However, SWAT-WB’s validation period was not as accurate as its calibration period, while the CN based model performed better during validation. A visual comparison of SWAT-WB’s hydrograph with the measured hydrograph (Figs. 4 and 6) indicates that the model performs fairly well for the Townbrook watershed, a fact supported by the reasonably high daily $NSE$ values of 0.64 and 0.52 for the calibration and validation periods, respectively (Table 2).

Similar to the Gumera results, differences in spatial distribution of runoff is evident when the same event from November 2003 is compared between the Townbrook models (Figs. 7A and B). As expected, SWAT-CN predicts some surface runoff from the majority of the watershed and it is clearly driven by differences in landuse, whereas SWAT-WB predicts substantial portions of the watershed producing no surface runoff, not surprising considering the emphasis the models place on topographic position as it pertains to runoff generation. For this particular storm SWAT-WB predicted that most of the
wetness classes in the low lying areas of the watershed would be saturated at the start of this event, leading to these low-lying wet areas producing nearly identical volumes of runoff (i.e. almost the entire volume of precipitation).

Comparing the SWAT-CN and SWAT-WB predicted water table heights to those measured by Lyon et al. (2006) shows that the SWAT-WB predicted soil water table height agreed with measurements across the monitored hill side in the watershed with $R^2 = 0.68$ (Fig. 7a). There was a slight tendency for SWAT-WB to under predict water table height for large water table heights (Fig. 7a), and slightly under predict small water table heights. SWAT-CN, however, systematically under predicted water table height for all conditions (Fig. 7b).

Discussion

While the CN, a loosely constrained calibration parameter in most applications was removed from SWAT; the WB routine adds another calibration parameter ($EDC$), negating the potential reduction in calibrated parameters. The need to calibrate $EDC$ became evident when the saturation deficit for each soil profile was calculated. For instance, if the entire soil profile was included in the calculation of the available storage, $\tau$, the model would not simulate any surface runoff, all precipitation would infiltrate. If only the uppermost soil horizon were used to determine $\tau$, then essentially all precipitation would runoff, resulting in no infiltration. By examining a range of soil depths used to calculate $\tau$, it became clear that the total depth used to determine surface runoff had to be adjusted; hence the introduction of $EDC$, the effective depth coefficient. This issue has been realized in previous iterations of water balance models, and many previous water balance models of the Blue Nile were
limited to application at a monthly time-step due to inabilities to successfully partition moisture between baseflow, interflow, and surface runoff (Johnson and Curtis, 1994; Conway, 1997). When no EDC was used in SWAT-WB, these same issues were present; high \( \tau \) values resulted in all moisture as baseflow, and when \( \tau \) was too high, all moisture in excess of soil capacity became surface runoff with minimal baseflow contributions.

Interestingly, the EDC solution to these issues is remarkably similar to a recent water balance model developed for the Blue Nile by Kim and Kaluarachchi (2008) that combines a water balance with a traditional tank model. To differentiate between surface and various subsurface flows, Kim and Kaluarachchi (2008) developed a model using two ‘tanks’. The upper tank, described by an upper zone soil moisture term, was used to calculate surface runoff, and a lower zone term was used to model baseflow. The upper layer would produce no surface runoff until a “runoff orifice” depth was filled by rainfall. This upper zone soil layer with its runoff orifice depth is analogous to SWAT-WB’s EDC term; both parameters acknowledge that in saturation excess dominated areas only a certain portion of the soil profile plays a role in runoff generation.

Clear improvements were made to SWAT in the Ethiopian watershed by removal of the CN, however the results are not as definitive for the Townbrook watershed in New York State. While SWAT-WB has substantially higher model accuracy for the calibration period, it does not perform as well during validation as SWAT-CN. By comparing the hydrograph from the Townbrook outlet (Fig. 6) and the model statistics, it is clear that SWAT-WB performs at least as well
as SWAT-CN. Thus in case where rainfall is evenly distributed throughout to
year both models perform equally well in predicting discharge as would have
been expected from earlier studies (Steenhuis et al Easton
In Townbrook, Easton et al. (2008) showed that the inclusion of the STI based
wetness index better captured the spatial distribution of water table depths
and, by extension, runoff producing areas. SWAT-WB provided a similar level
of accuracy in predicting water table heights (Fig. 7). While we have no
distributed runoff data for either of the watersheds, including the wetness
index in SWAT-WB resulted in what appears to be more a realistic distribution
of runoff generating areas than SWAT-CN. Indeed, in VSA dominated
watersheds, runoff generation is closely related to soil moisture levels (as
controlled by perched water table levels), which on turn is governed, to a large
extent, by topographic position. In many SWAT-CN applications the location of
an HRU within each subbasin is not a concern and thus, any locations that
share landuse and soil are treated identically, regardless of its topographic
position and the corresponding likelihood to produce runoff. In SWAT-WB,
STIs were used to link HRUs by similar topographic position, giving model
users the capability to examine intra-watershed runoff dynamics.

This difference in spatial distribution of runoff generating areas predicted by
SWAT-WB and SWAT-CN is clearly demonstrated for both Gumera (Fig. 5)
and Townbrook in Fig. 8. For the same large storm event in the Gumera basin
(Fig. 5), SWAT-WB did not generate surface runoff for all HRUs, whereas
SWAT-CN predicted that the entire watershed would contribute surface runoff
more or less evenly. Due to the imposed topographical controls, SWAT-WB
predicted that the wettest portions of the watershed would contribute more
runoff than drier areas. In addition to the fact that SWAT-CN predicts a nearly uniform runoff volume for the entire watershed, there are two other points of interest that should be discussed. First, is the fact that SWAT-CN predicts that the area nearest Gumera’s outlet produces the least amount of surface runoff, exactly opposite of SWAT-WB’s results which predict that this area produces the highest runoff volumes. These differences between the models can easily be explained by the inclusion of slope in the HRU delineation (and therefore EDC calibration). Again, holding with VSA principles, SWAT-WB assumes that these flat, near-stream regions will wet up and contribute the most runoff, because of reduced lateral flow whereas SWAT-CN treats these HRUs the same as any upland region with the same soil and land cover. The second interesting point is that both models predict that certain upland regions generating a significant portion of surface runoff from the test storm. These soils have a low saturated conductivity. For SWAT-CN this result in a high curve number and for SWAT-WB an increase into STI value. Both increase runoff.

SWAT-CN and most other watershed models have been developed for temperate climates where rainfall is generally well distributed throughout the year. Utilizing models developed in a temperate climate for Ethiopia conditions, with a monsoonal climate, is problematic. Temperate models assume that there is a nearly unique relationship between precipitation amounts or intensity and runoff generated. This is not the case for Ethiopia as demonstration by the results of Liu et al. (2008) where for three watersheds with more than 16 years of record, the rainfall relationship was far from unique. The first rains after the dry season all infiltrate and nearly no runoff is
generated. As the rainfall season progresses more and more rainfall becomes runoff. Since the intensity of the rain did not affect the runoff amounts for a given storm, Liu et al. (2008) concluded that the runoff mechanism was dominated by saturation excess processes.

Water balance models are consistent with saturation excess runoff process because the runoff is related to the available watershed storage capacity and the amount of precipitation. The implementation of water balances into runoff calculations in the Blue Nile Basin is not a novel concept and others have shown that water balance type models often perform better than more complicated models in Ethiopian type landscapes (Johnson and Curtis, 1994; Conway, 1997; Ayenew and Gebreegziabher, 2006; and Liu et al., 2008).

However, these water balance models are typically run on a monthly or yearly time steps because the models are generally not capable of separating base- inter- and surface runoff flow. To truly model erosion and sediment transport (of great interest in the Blue Nile Basin), large events must be captured by the model and daily simulations are required to do so. Thus SWAT-WB not only maintains a water balance but also calculates the interflow and the base flow component, and gives a reasonable prediction of peak flows. SWAT-WB is therefore more likely to be capable of predicting sediment transport than either SWAT-CN or water budget models with monthly time steps.

**Conclusion**

Daily modeling of stream flow and surface runoff in a monsoonal watershed was substantially improved by replacement of the CN method with a simplified water balance routine in the SWAT model. SWAT-WB uses calculated
saturation-deficit values with an effective depth coefficient, $EDC$, to partition rainfall into surface runoff and infiltrating water. This $EDC$-based water balance method is analogous to other tank models that have been successfully applied in monsoonal regions.

SWAT-WB was as accurate in predicting discharge at the outlet as the CN method in a watershed that experiences evenly distributed rainfall throughout the year (New York). SWAT-WB also predicted the distribution of water table heights on a hillslope in the watershed significantly better than SWAT-CN, giving us increased confidence that the spatial distribution of runoff dynamics are more realistically captured.

These results indicate that SWAT performs better in saturation-excess controlled areas when a simple saturation-deficit water balance model is used to calculate runoff volumes. With this physically-based, and easy-to-use model, effective water and land management schemes will be easier to successfully implement in watersheds dominated by saturation excess runoff generation, particularly data-poor regions where use of the CN methodology has not been validated.
REFERENCES


NNDC Climate Data Online. Retrieved November 24, 2007, from
www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD.
Page, T., Haygarth, P. M., Beven K. T., Joynes, A., Butler, T., Keeler, C., Freer,
in relation to the topographic index and critical source areas: Sampling
for assessing risk to water quality. Journal of Environmental Quality. 34,
2263-2277.
the Lake Tana basin, Ethiopia using SWAT model. The Open Hydrology
Journal. 2, 49-62.
Modeling the hydrochemistry of the Cannonsville Watershed with
Generalized Watershed Loading Functions (GWLF). Journal of the
Schneiderman, E. M., Steenhuis, T. S., Thongs, D. J., Easton, Z. M., Zion, M. S.,
source area hydrology into a curve-number-based watershed model.
Hydrological Processes. 21, 3420-3430.
vegetation attributes for developing pedotransfer functions. Soil Science
Society of America Journal. 70, 1430-1440
procedure on coastal plain watersheds. Proceedings from: The 1986
Winter Meeting of the American Society of Agricultural Engineers. Paper
No. 862505. Chicago, IL: ASAE.
Journal of Irrigation and Drainage Engineering. 121(3), 234-238.
Steenhuis, T. S., A. S. Collick, Z. M. Easton, E. S. Leggesse, H. K. Bayabil, E. D.
White, S. B. Awulachew, E. Adgo, and A. A. Ahmed. 2009. Predicting
discharge and erosion for the Abay (Blue Nile) with a simple model.
modeling across a physiographic region: Topographic patterns and
model transportability. Geoderma. 133, 57-70.
algorithm for computationally efficient watershed model calibration. Water
Resources Research. 43, W01413-W01429.
Handbook, Part 630 Hydrology, Section 4, Chapter 10.


Table 1. Model statistics for daily streamflow in Gumera Basin.

<table>
<thead>
<tr>
<th></th>
<th>SWAT-WB</th>
<th></th>
<th>SWAT-CN¹</th>
<th></th>
<th>SWAT-CN²</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration</td>
<td>Validation</td>
<td>Calibration</td>
<td>Validation</td>
<td>Calibration</td>
<td>Validation</td>
</tr>
<tr>
<td>NSE</td>
<td>0.70</td>
<td>0.76</td>
<td>0.64</td>
<td>0.67</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>R²</td>
<td>0.71</td>
<td>0.81</td>
<td>0.65</td>
<td>0.73</td>
<td>0.71</td>
<td>0.70</td>
</tr>
</tbody>
</table>

¹ same input as SWAT-WB
² from Setegn et al., 2008
Table 2. Model statistics for daily streamflow in Townbrook.

<table>
<thead>
<tr>
<th></th>
<th>SWAT-WB</th>
<th></th>
<th>SWAT-VSA</th>
<th></th>
<th>SWAT-CN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration</td>
<td>Validation</td>
<td>Calibration</td>
<td>Validation</td>
<td>Calibration</td>
<td>Validation</td>
</tr>
<tr>
<td>NSE</td>
<td>0.64</td>
<td>0.52</td>
<td>0.56</td>
<td>0.68</td>
<td>0.43</td>
<td>0.62</td>
</tr>
<tr>
<td>R²</td>
<td>0.69</td>
<td>0.65</td>
<td>0.64</td>
<td>0.74</td>
<td>0.59</td>
<td>0.69</td>
</tr>
</tbody>
</table>
Figure 1. Wetness classes for Gumera, Ethiopia.
Figure 2. Wetness classes for Townbrook, located in the Catskill Mountains of New York State.
Figure 3. Elevation of the Gumera Basin, located east of Lake Tana in the Ethiopian Highlands.
Figure 4. Observed and modeled streamflow for Gumera using (a) SWAT-WB, and (b) SWAT-CN.
Figure 5. Spatial distribution of surface runoff in Gumera modeled by: A) SWAT-CN and B) SWAT-WB.
Figure 6. Observed and modeled streamflow in the Townbrook watershed using (a) SWAT-WB and (b) SWAT-CN.
Figure 7. Relationship between SWAT-WB (a) and SWAT-CN (b) predicted water table heights above the restricting layer by index class (SWAT-WB) or landuse (SWAT-CN) and the measured water table heights for March 2004 – September 2004 from Lyon et al. (2006). Individual measured points within an index class or landuse represent the average of the piezometric measurement within the respective classes for a single day.
Figure 8. Spatial distribution of surface runoff in Townbrook modeled by (a) SWAT-WB, and (b) SWAT-CN.