JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

AMERICAN WATER RESOURCES ASSOCIATION

INVESTIGATION OF THE CURVE NUMBER METHOD FOR SURFACE RUNOFF ESTIMATION IN TROPICAL REGIONS[1]

Yihun Taddele Dile, Louise Karlberg, Raghavan Srinivasan, and Johan Rockström[2]

ABSTRACT: This study tests the applicability of the curve number (CN) method within the Soil and Water Assessment Tool (SWAT) to estimate surface runoff at the watershed scale in tropical regions. To do this, surface runoff simulated using the CN method was compared with observed runoff in numerous rainfall-runoff events in three small tropical watersheds located in the Upper Blue Nile basin, Ethiopia. The CN method generally performed well in simulating surface runoff in the studied watersheds (Nash-Sutcliff efficiency [NSE] > 0.7; percent bias [PBIAS] < 32%). Moreover, there was no difference in the performance of the CN method in simulating surface runoff under low and high antecedent rainfall (PBIAS for both antecedent conditions: ~30%; modified NSE: ~0.4). It was also found that the method accurately estimated surface runoff at high rainfall intensity (e.g., PBIAS < 15%); however, at low rainfall intensity, the CN method repeatedly underestimated surface runoff (e.g., PBIAS > 60%). This was possibly due to low infiltrability and valley bottom saturated areas typical of many tropical soils, indicating that there is scope for further improvements in the parameterization/representation of tropical soils in the CN method for runoff estimation, to capture low rainfall-intensity events. In this study the retention parameter was linked to the soil moisture content, which seems to be an appropriate approach to account for *antecedent wetness conditions* in the tropics.

(KEY TERMS: runoff; precipitation; rainfall intensity; watershed; curve number; SWAT; Upper Blue Nile.)

Dile, Yihun Taddele, Louise Karlberg, Raghavan Srinivasan, and Johan Rockström, 2016. Investigation of the Curve Number Method for Surface Runoff Estimation in Tropical Regions. *Journal of the American Water Resources Association* (JAWRA) 1-15. DOI: 10.1111/1752-1688.12446

INTRODUCTION

The curve number (CN) method is a conceptual model supported with empirical data to estimate direct runoff volume from single precipitation events on small agricultural watersheds (Ponce and Hawkins, 1996). The CN method uses several predefined curves to describe the relationship between rainfall and runoff depending on soil type, land-use type, and surface treatment conditions. However, the method does not take into account the spatial and temporal variability in infiltration and other abstractive losses; rather, it aggregates these into the total depth loss for a given rainfall event and drainage area (Ponce and Hawkins, 1996).

The CN method was originally developed in 1954 by the Soil Conservation Service (SCS) of the U.S. Department of Agriculture, since renamed the Natural Resources Conservation Service (NRCS, 2004).

[1]Paper No. JAWRA-15-0102-P of the *Journal of the American Water Resources Association* (JAWRA). Received June 25, 2015; accepted May 27, 2016. **Discussions are open until six months from issue publication**.

[2]Postdoctoral Research Associate (Dile) and Professor (Srinivasan), Department of Ecosystem Science and Management, Texas A&M University, 1500 Research Park, Suite B221, College Station, Texas 77801; Research Fellow (Karlberg), Resources and Development Unit, Stockholm Environment Institute, Stockholm, Sweden; and Professor (Rockström), Stockholm Resilience Centre, Stockholm University, Stockholm Sweden (E-Mail/Dile: yihundile@tamu.edu).

The origins of the methodology can be traced back to thousands of infiltrometer tests carried out by SCS during the 1930s and 1940s at experimental sites (Ponce and Hawkins, 1996; Williams *et al.*, 2012). Although it was originally developed mainly for agricultural watersheds, the CN method has since been adapted for urbanized and forested watersheds (Cronshey *et al.*, 1986).

Among the perceived advantages of the CN method are its simplicity, practicality, predictability, stability, reliance on a single parameter, and responsiveness to watershed properties such as soil type, land use, surface condition, and antecedent condition (Ponce and Hawkins, 1996; Yu, 2012). However, the marked sensitivity of runoff estimation to the choice of CN, the lack of clear guidance on how to vary the antecedent condition, the method's variable accuracy for different biomes, and the fixing of the initial abstraction ratio at 0.2 have all been cited as weaknesses (Ponce and Hawkins, 1996). The most criticized assumption in the CN method is that the ratio of actual retention to potential retention is the same as the ratio of actual runoff to potential runoff.

The CN method is now integrated into more complex simulation models applied throughout the world (e.g., Williams *et al.*, 1985; Arnold *et al.*, 1998; HEC-HMS, 2000). In continuous models — including the Soil and Water Assessment Tool (SWAT) — the CN is updated daily to take account of the *antecedent wetness condition* (also commonly referred to as the "antecedent runoff condition") (Williams *et al.*, 2012). The daily CN in SWAT is determined based on daily soil moisture content or plant evapotranspiration (Neitsch *et al.*, 2012).

The CN method has been applied throughout the world (e.g., Hjelmfelt, 1991; Krysanova *et al.*, 2005; Abbaspour *et al.*, 2007; Krysanova and Arnold, 2008; Rostamian *et al.*, 2008; Schuol *et al.*, 2008; Stehr *et al.*, 2008; Setegn *et al.*, 2010a, b; Yu, 2012). However, several authors (e.g., Liu *et al.*, 2008; Collick *et al.*, 2009; Steenhuis *et al.*, 2009; White *et al.*, 2011) have questioned the representativeness of the approach in different climates and geological settings as it was originally developed for applications in temperate regions, especially in the United States. Tropical regions tend to have highly variable rainfall with distinct dry seasons, rainfall events of high intensity, and soil types that are highly crusted, have low organic matter content, and have dry surface conditions. In temperate regions the rainfall intensity is moderate and the soil has high infiltrability and conductivity.

The underlying mechanism for the CN method is perceived to be infiltration excess mechanism (e.g., Liu *et al.*, 2008; Tilahun *et al.*, 2015). Liu *et al.* (2008) studied the effective rainfall-discharge relationships for three small watersheds in the Ethiopian highlands, namely, Anjeni, Andit Tid, and Maybar. They found that after 500 mm rainfall threshold, approximately 50% of any further rainfall on these watersheds will directly contribute to catchment runoff, instead suggesting that saturation excess process is the dominant mechanism for runoff generation in the Ethiopian highlands. Thereafter, more field and modeling studies have been conducted in watersheds in the Ethiopian highlands (e.g., Collick *et al.*, 2009; Steenhuis *et al.*, 2009; Bayabil *et al.*, 2010; White *et al.*, 2011; Tilahun *et al.*, 2013a, b, 2015) and demonstrated that saturation excess is the main runoff generation mechanism. On the other hand, SWAT and the CN method has been widely applied in the Lake Tana watersheds in particular and the Ethiopian highlands in general (e.g., Setegn *et al.*, 2009, 2010a,b; Betrie *et al.*, 2011; Dile *et al.*, 2013, 2016; Gebremicael *et al.*, 2013; Dile and Srinivasan, 2014) and provided satisfactory results. Most of these studies, however, have not directly evaluated the surface runoff estimation with the CN method. Their evaluation was based on the total streamflow, where interception, infiltration, plant and soil evaporation, channel losses, such as evaporation and seepages, and groundwater recharge and return flow are part of the process in evaluating the hydrological budget (cf. Tessema *et al.*, 2014). Evaluation of SWAT model performance based on total streamflow may not fully evaluate the surface runoff estimation by the CN method.

This study aims to test the applicability of the CN method for estimating surface runoff in tropical regions by comparing runoff estimates through the CN method with measured surface runoff in three experimental small watersheds in the Upper Blue Nile basin of Ethiopia. This is, to our knowledge, the first study to use data on measured surface runoff to test the applicability of the CN method for estimating surface runoff in tropical watersheds. This study evaluates the CN method focusing only on direct runoff on cultivated and grazing landscapes in three small watersheds. At the small watershed, with highly controlled environment of one soil type, one land-use type, and slope, it is more convenient examining the surface runoff estimation through the CN method in tropical regions.

Furthermore, to assess the usability of the CN method in tropical regions, we specifically tested the method under conditions typical of tropical regions that may affect the runoff process, in one of the watersheds. Five-day antecedent rainfall was used as a proxy to capture conditions of very dry and wet soils that commonly occur in monsoonal climates. Moreover, we compared high and low rainfall-intensity events to test the ability of the method to capture tropical storms (e.g., very high intensity events). Surface runoff was estimated using the CN approach within the ArcSWAT modeling environment (Neitsch *et al.*, 2012).

ArcSWAT is an ArcGIS extension for the SWAT graphical user interface (Winchell *et al.*, 2013). The theoretical framework for the CN method within SWAT model is presented as Supporting Information (SI).

MATERIALS AND METHODS

This research was based on data from the southeastern part of the Lake Tana basin, which is located in the upper part of the Upper Blue Nile basin in Ethiopia. The field data were collected from three small watersheds: Gegudeguade, Aletu, and Shimbraye (Figure 1). These watersheds were chosen to capture rainfall-runoff dynamics in different landscapes. Two were located on agricultural land and the third on grazing land. Large parts of the agricultural watersheds were covered by fields of teff (*Eragrostis tef*, a native grain crop). As the watersheds were close to each other — it was possible to place a temporary weather station within 1.2 km from any of them — rainfall variation between them was assumed small. Table 1 gives the basic characteristics of the three watersheds. Secondary data for use in setting up the model were collected from national and global databases.

Field Data

The field data were collected in the rainy seasons (July-September) of 2011 and 2012. Climatic and surface runoff data were collected from all three watersheds. Climatic variables were monitored with a Vantage Davis Pro2 weather station (Davis Instruments Corporation, Hayward, California) at a five-min time step. The parameters monitored included rainfall, maximum and minimum temperatures, solar radiation, wind speed, and relative humidity.

Surface runoff was measured in the watersheds using 2.5 H flumes (Plasti-Fab, Tualatin, Oregon). A combination of manual and automatic stage recorders

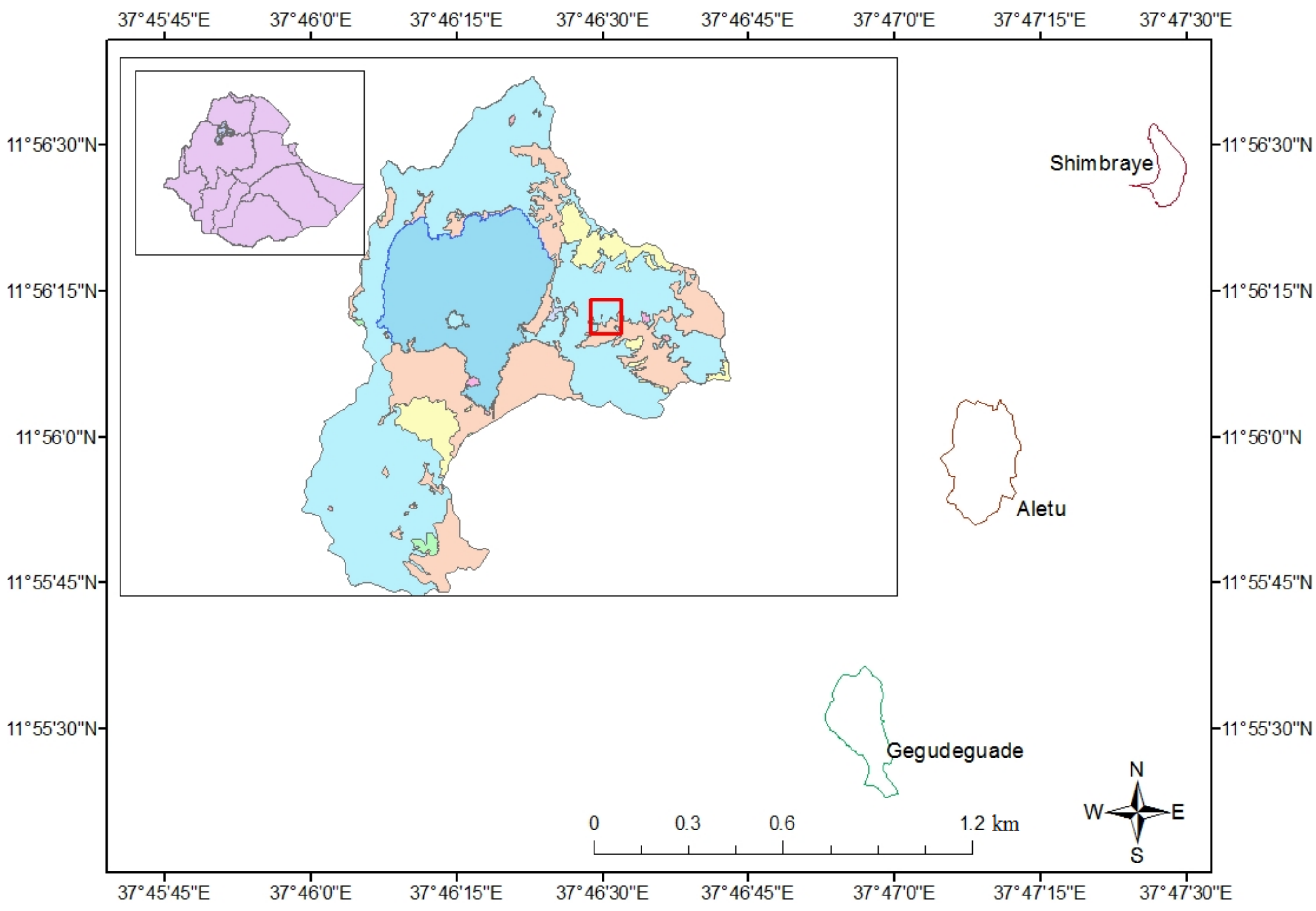

FIGURE 1. The Location of the Watersheds Included in This Study. The red box in the large inset shows their location in the Lake Tana basin. The small inset shows the basin's location in Ethiopia.

TABLE 1. Characteristics of Three Small Watersheds in the Lake Tana Basin, Ethiopia.

Micro-Watershed	Area (ha)	Elevation (m a.s.l.) Max.	Elevation (m a.s.l.) Min.	Land-Use Type	Distance from Temporary Weather Station (m)
Gegudeguade	4.63	1,920	1,879	Crops	726
Aletu	6.71	1,892	1,866	Crops	416
Shimbraye	1.84	1,925	1,881	Grazing	1,140

Note: m a.s.l., meters above sea level.

was employed: the Stage Discharge Recorder (SDR-0001-1) from the Sutron Corporation (Sterling, Virginia) and the Schlumberger Micro-diver (Delft, the Netherlands). The automatic data loggers (SDR or Micro-diver) measured stages at five-min intervals. Manual measurements were taken at small time intervals so as to produce a smooth hydrograph. The stages were converted into discharge (flow rates) using the standard 2.5 H flume rating curve equation (Q = 0.001499−0.01992H$^{0.4}$ +0.727294H$^{1.4}$ + 1.698273H$^{2.5}$). The flow rates for the events were converted into depth of water to estimate the surface runoff that could be generated uniformly over the entire catchment area.

As the watersheds were small there were no large groundwater reservoirs and the base-flow contribution was negligible. Measurements show that the generation of runoff in the watershed channels stops only a few minutes after the secession of rainfall, suggesting that the flow in the channels mainly consists of overland flow. Therefore, this study assumed that the groundwater contribution from interflow and bypass flows was negligible, and base-flow separation was therefore unnecessary.

Runoff data were collected for 70 rainfall events. However, only those generating runoff and for which the hydrograph was complete were used in the analysis. Of the total observed runoff events, 53 events (which accounts 75.7%) were considered in the analysis. The remaining events (24.3% of the total observed runoff events) were excluded as part of the data were lost due to either equipment failure or weir overflow. The analysis used 31 events for Gegudeguade, 9 for Aletu, and 13 for Shimbraye. These were considered sufficient for the purposes of the study (cf. Blume *et al.*, 2010).

To see if the studied rainfall-runoff events cover the entire rainfall distribution in the study area, we compared all observed rainfall events that generated runoff (i.e., events with a total rainfall amount which exceeded 0.8 mm, which amounted to 84 in total) to those events used in the rainfall-runoff analysis. We found that of the rainfall events that generate runoff, nearly 60% of the events had a rainfall amount of 0.8-5 mm, and only three events exceeded 25 mm (Figure 2a). Comparing with the rainfall events used to study runoff generation in this study, we found that our subset of rainfall events is representative of rainfall events in the area in general (Figure 2a). For instance, of the 53 rainfall events that were used in this manuscript to study runoff, 45% of them had a rainfall amount of 0.8-5 mm and 2 events were above 25 mm, thus suggesting that the rainfall events that were used for the rainfall-runoff analysis is a good representation of the entire rainfall distribution.

The temporal distribution of the rainfall events included in the study varied over the rainy season (Figure 2b). Data collection started in the beginning of July and continued until mid-September. Eight of the 53 events occurred by mid-July. The maximum amount of rainfall of a particular event during the early phase of the rainfall season (i.e., until July 15) was 21.2 mm. Most of the studied events (37 of 53) occurred between July 16 and August 15. The highest rainfall amounts were observed in the first two weeks of August. After mid-August, both the frequency of rainfall events and amount of rainfall decreased. Eight of the studied events happened between mid-August to mid-September.

Secondary Data

The secondary data used in setting up the model included a digital elevation model (DEM), soil, land cover, and historical climatic data. The DEM was needed to delineate the watersheds in the ArcSWAT interface, and the land-use and soil data were important to define the hydrologic response units (HRUs). Historical weather data were used to simulate the hydrological processes in SWAT. The weather data consisted of daily rainfall and maximum and minimum temperatures at the nearby Woreta weather station between 1990 and 2012. The built-in weather generator for SWAT (Neitsch *et al.*, 2012) was used to fill gaps in the data.

The DEM data were obtained from the CGIAR Consortium for Spatial Information website (CGIAR-CSI, 2012) and had a resolution of 90 m × 90 m. The stream network, land use, and soil maps of the study

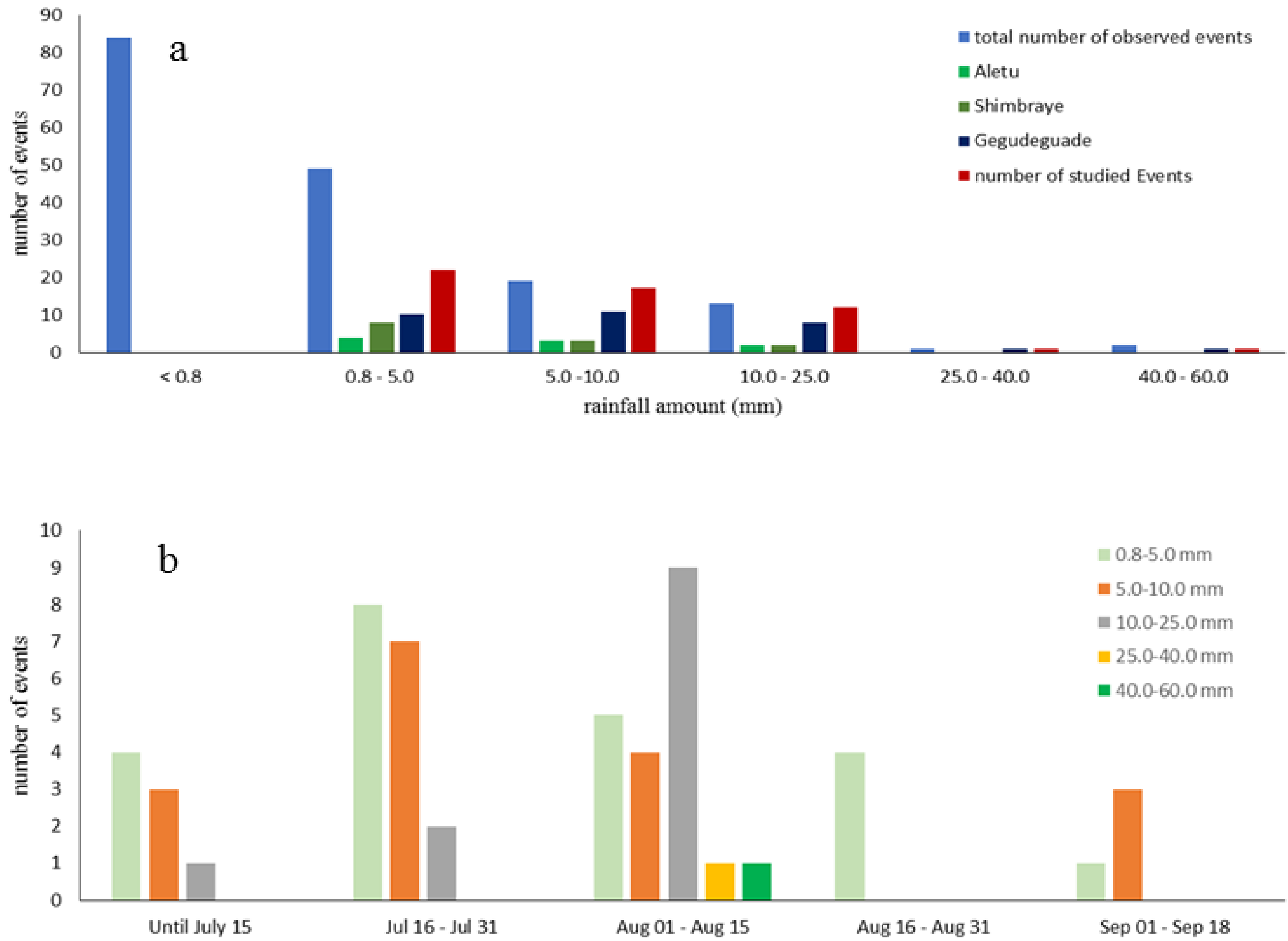

FIGURE 2. (a) Comparison between Observed Rainfall Distribution in the Study Watershed, and Rainfall Events That Were Used for the Rainfall-Runoff Analysis in This Study, (b) Temporal Distribution of the Rainfall Events Included in This Study.

area were from the Ethiopian Ministry of Water Resources (MoWR, 2009). The physical and chemical properties parameters for soil that are required by SWAT were derived from the Africa map sheet of the 1995 CD-ROM edition of the Soil Map of the World (FAO, 1995). The weather data was provided by the Ethiopian National Meteorological Services Agency (ENMSA, 2012).

Model Setup

An ArcSWAT model environment (Winchell *et al.*, 2013) was used to test the CN method in this study. SWAT is a physically based model developed to predict the impact of land management practices on water, sediment, and agricultural chemical yields in watersheds with varying soil, land use, and management conditions (Neitsch *et al.*, 2012). The model can simulate hydrological cycles, vegetation growth, and nutrient cycling with a daily time step by disaggregating a river basin into subbasins and HRUs. HRUs are lumped land areas within the subbasin representing unique land cover, soil, and management combinations.

Five slope classes were implemented in defining the HRUs. These were slope <2, 2-8, 8-12, 12-20, and >20%. A threshold area of 0.5 ha was used for watershed delineation at each outlet, to replicate the characteristics of the watersheds. Multiple HRUs were created within each subbasin, and zero percent threshold area was used to define the HRUs (i.e., all land use, soil, and slope classes in a subbasin were considered in creating the HRUs). The management data used in setting up the model for the two agricultural watersheds are shown in Table 2.

The SWAT model was run from 1990 to 2012, including five years of warm-up period (Daggupati

TABLE 2. Land Management Data Used to Set Up a Soil and Water Assessment Tool Model for the Agricultural Small Watersheds.

Operation	Tillage	Planting	Fertilizer* Urea	Fertilizer* DAP	Harvest and Kill
Timing	April 1-July 22	July 22	July 22 and August 22	July 21	December 5
Amount/frequency	Four times		18.75 kg/ha and 18.75 kg/ha	37.5 kg/ha	

*Urea is nitrogen-based fertilizer, and is applied twice per cropping period (at planting and prior to flowering). DAP is phosphorous-based fertilizer.

et al., 2015), using the CN method to simulate surface runoff. The retention parameter was updated based on soil moisture content. Hargreaves method was used to determine potential evapotranspiration, and the variable storage routing method was used to route the flow of water in the channels. As the CN method assigns CN values for a given land use and soil combinations, it was appropriate to evaluate the method with actual values, i.e., without any parameter manipulations. With good model calibration skills and possibly with model overparameterization, one can easily fit observed and simulated values, especially using automatic model calibration tools. Therefore, we were interested to explore if the CN method works well without any parameter calibrations. Although the model was not calibrated, due emphasis was given to correctly simulating other biophysical processes. For example, as crop growth and biomass production significantly affect the runoff generation and evapotranspiration processes, we ensured that crop growth was well represented by the modeling and that the teff yield (~1.0 ton/ha) was in agreement with regional census data (Dile and Srinivasan, 2014). We also compared the evapotranspiration estimation from the SWAT model to literature. The simulated average actual evaporation in the studied watershed was 610 mm. Using the Hargreaves method in SWAT, Setegn *et al.* (2010a, b) estimated that the average evapotranspiration in the entire Lake Tana basin including the Lake is ~758 mm. The difference between our estimate and that of Setegn *et al.* (2010a, b) was likely due to the inclusion of evaporation from Lake Tana in their calculation. The actual evaporation from Lake Tana ranges 1,478-1,789 mm (Kebede *et al.*, 2006; Wale *et al.*, 2009; Rientjes *et al.*, 2011; Dessie *et al.*, 2015). Given Lake Tana's large area coverage in the basin, which is ~20% of the Lake Tana basin, it can be expected that the estimate from Setegn *et al.* (2010a, b) is a bit higher.

The model setup created 2 HRUs each in the Gegudeguade and Shimbraye watersheds, and 16 HRUs in the Aletu watershed. For the simulation period of 1995-2012 (excluding the model warm-up period), the CN value for the Gegudeguade and Aletu ranges 70.68-92.13 (average at 83.34), and for Shimbraye, it ranges from 68.91 to 92.13 (average at 82.95).

Statistical Evaluation of the SWAT Simulations

Rainfall-runoff data were collected from the watersheds per event, whereas surface runoff simulations from SWAT were often made daily. The collected rainfall-runoff events and the respective event rainfall-runoff simulations (i.e., the event rainfalls were simulated as daily rainfalls in SWAT) were compared.

Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) and percent bias (PBIAS) were used to evaluate the agreement between the overall observed and simulated surface runoff (Table 3). Moriasi *et al.* (2007) provided clear guidance on how to evaluate model performance in terms of these goodness-of-fit criteria. NSE can range from $-\infty$ to 1. An NSE value of 1 indicates a perfect match between observed and simulated data. An NSE value between 0 and 1 is theoretically considered acceptable, whereas a value of 0 or less suggests that the observed mean is a better predictor than the model. PBIAS indicates the average tendency of the simulated values to be higher or lower than the corresponding observed values. The optimal PBIAS value is 0; positive values indicate that the model tends to underestimate, and negative values that it tends to overestimate (Gupta *et al.*, 1999). Moriasi *et al.* (2007) suggest that PBIAS can easily quantify water balance errors and thus indicate model performance. In general, model simulations of streamflow are considered satisfactory if NSE > 0.50, and PBIAS $< \pm 25\%$ (Moriasi *et al.*, 2007).

To study how the CN method performed under different monsoonal climatic conditions, rainfall-runoff events were divided into low and high rainfall-intensity categories, and into low and high antecedent rainfall categories. In the case of continuous numerical data such as rainfall intensity and five-day antecedent rainfall, conditioning can be done by grouping a set of observations into bins. The analysis for the grouping was done using the lattice package in R statistical computing environment (Sarkar, 2008).

TABLE 3. Statistical Goodness-of-Fit Criteria Used to Evaluate the Performance of the Curve Number Method.

No.	Statistical Goodness-of-Fit Criteria	Equation	Reference
1	Nash-Sutcliffe efficiency (NSE)	$\mathrm{NSE} = 1 - \left[\frac{\sum_{i=1}^{n}\left(Q_o^i - Q_s^i\right)^2}{\sum_{i=1}^{n}\left(Q_o^i - \bar{Q}_o\right)^2}\right]$	Legates and McCabe (1999) and Moriasi *et al.* (2007)
2	Percent bias (PBIAS)	$\mathrm{PBIAS} = \left[\frac{\sum_{i=1}^{n}\left(Q_o^i - Q_s^i\right)\times 100}{\sum_{i=1}^{n} Q_o^i}\right]$	Green and Stephenson (1986) and Moriasi *et al.* (2007)
3	Mean absolute error (MAE)	$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left\lvert Q_o^i - Q_s^i\right\rvert$	Blume *et al.* (2010)
4	Sum of absolute errors (G)	$G = \sum_{i=1}^{n}\left\lvert Q_o^i - Q_o^i\right\rvert$	Green and Stephenson (1986)
5	Root mean square error (RMSE)	$\mathrm{RMSE} = \left(\frac{1}{n}\sum_{i=1}^{n}\left(Q_o^i - Q_s^i\right)^2\right)^{1/2}$	Green and Stephenson (1986), Patry and Mariño (1983)
6	Proportional error of estimate (PEE)	$\mathrm{PEE} = \left[\sum_{i=1}^{n}\left(\frac{Q_o^i - Q_s^i}{Q_o^i}\right)^2\right]^{1/2}$	Green and Stephenson (1986) and Manley (1978)
7	Standard error of estimate (SEE)	$\mathrm{SEE} = \left(\sum_{i=1}^{n}\frac{\left(Q_o^i - Q_s^i\right)^2}{(n-2)}\right)^{1/2}$	Green and Stephenson (1986) and Jewell *et al.* (1978)
8	Modified NSE (E_j)	$E_j = 1 - \frac{\sum_{i=1}^{n}\lvert Q_o^i - Q_s^i\rvert^j}{\sum_{i=1}^{n}\lvert Q_o^i - \bar{Q}_o\rvert^j}$	Krause *et al.* (2005) and Legates and McCabe (1999)
9	Relative efficiency criterion (E_{rel})	$E_{\mathrm{rel}} = 1 - \frac{\sum_{i=1}^{n}\left(\frac{Q_o^i - Q_s^i}{Q_o^i}\right)^2}{\sum_{i=1}^{n}\left(\frac{Q_o^i - \bar{Q}_o}{\bar{Q}_o}\right)^2}$	Krause *et al.* (2005)

Note: Q_o^i = observed data at the ith time step; Q_s^i = simulated data at the ith time step; $Q_{\mathrm{obs}}^{\mathrm{mean}}$ = mean of the observed data; and n = total number of observations.

R uses shingles as a way to represent intervals. The intervals are created using overlapping values (like roof shingles) to make smooth transition between the shingles. Gegudeguade watershed, which yielded the most observed surface runoff data, was used for this analysis. Events with rainfall intensity of 1.93-9.62 mm/hr were placed in the low rainfall-intensity category, and those with 7.32-31.82 mm/hr in the high rainfall-intensity category. The low antecedent rainfall category included events with average five-day antecedent rainfall of 4.09-12.24 mm, and the high antecedent rainfall category included events with 8.87-17.89 mm average five-day antecedent rainfall. Each rainfall-runoff category had 20 events with 11 overlapping events. The 11 overlapping events happened to create the smooth transition in the intervals (Sarkar, 2008). Although the overlapping events were necessary to create smooth transitions, they create some redundancy which may introduce bias to the results. The performance of the CN method is, therefore, evaluated in each category.

A variety of goodness-of-fit criteria were used to assess the performance of the CN method under these climatic conditions (Table 3). It was difficult to choose a single goodness-of-fit criterion for this assessment, as different criteria may give more weight to certain aspects of disagreements between simulated and observed data (Green and Stephenson, 1986). For example, NSE involves squaring of residuals. As a result it is highly sensitive to extreme values (Green and Stephenson, 1986; Legates and McCabe, 1999), which may create bias in the analysis toward higher estimates. For this reason, such types of statistical goodness-of-fit criterion (e.g., NSE) were not used in the runoff categorizing analysis. The literature (e.g., Green and Stephenson, 1986; Legates and McCabe, 1999; Krause *et al.*, 2005) suggests instead using the means, residual error estimation, and modified efficiency estimation methods for such analyses. It also suggests modified efficiency estimations, such as the modified NSE and the relative efficiency criterion (E_{rel}). Modified NSE (Table 3, No. 8, with j = 1) quantifies the difference between observation and simulation using the absolute values. In modified NSE, an overestimate or underestimate in higher values has a greater influence than one in lower values. The relative efficiency criterion (Table 3, No. 9) counteracts this deficiency by reducing the absolute differences during high flows. However, the relative efficiency criterion increases the influence of absolute lower differences during low-flow periods (Krause *et al.*, 2005). Because of these inherent biases, multiple statistical goodness-of-fit criteria were used to evaluate the performance of the CN method in calculating runoff at different rainfall-intensity and antecedent rainfall categories.

Further Exploration of Factors Influencing Runoff

Observed rainfall and runoff data were also analyzed without applying the CN method to obtain a better understanding of the rainfall-runoff relationship under different rainfall-intensity and antecedent rainfall categories. This was done to explore factors that play a major role in the hydrological process in the tropics, and get insight on how to adjust the CN parameter.

The 31 observed events in the Gegudeguade watershed were divided based on combinations of rainfall intensity and antecedent rainfall independently. One category consisted of events with high average five-day antecedent rainfall (i.e., higher than the median of 49.40 mm) and low rainfall intensity (i.e., lower than the median of 8.40 mm/hr). The other category consisted of events with low five-day antecedent rainfall and high rainfall intensity.

The rainfall-runoff events were further divided into three antecedent rainfall categories representing dry, average, and wet antecedent wetness conditions. The "dry" category had low five-day antecedent rainfall of 4.10-9.90 mm; the "average" category had five-day antecedent rainfall of 8.40-13.10 mm; and the "wet" category had five-day antecedent rainfall of 9.90-17.90 mm. There are eight overlapping rainfall-runoff events in each category.

RESULTS

Estimating Runoff with the CN Method

The assessment of the performance of the CN method to estimate surface runoff in the Upper Blue Nile basin was based on 53 rainfall-runoff events collected from three small watersheds (Figure 3). The largest amount of rainfall in a single event was ~45 mm. Over the three watersheds, the maximum measured runoff depth was ~20 mm, whereas the corresponding maximum simulated surface runoff depth was ~26 mm. The overall measured and simulated surface runoff showed reasonably good agreement (Table 4). The CN method generally gave lower surface runoff than what was observed. For the observed events in the three watersheds, the total simulated surface runoff was 79% of the measured

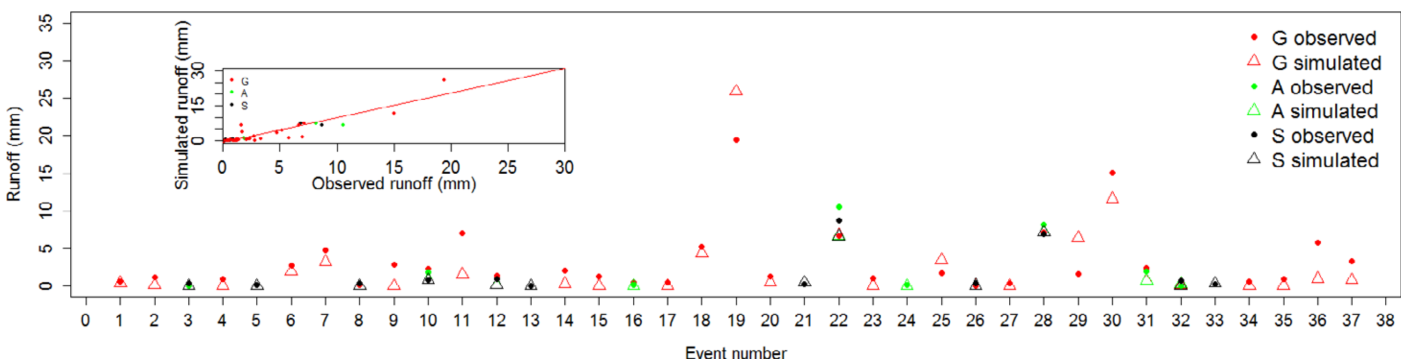

FIGURE 3. Comparison between the Observed Surface Runoff and Runoff Simulated Using the Curve Number Method for Individual Rainfall-Runoff Events in Three Watersheds: Gegudeguade (G), Aletu (A), and Shimbraye (S). The subplot shows the observed *vs.* simulated surface runoff. The regression line represents the Gegudeguade data (31 events).

TABLE 4. Comparison between Observed and Simulated Surface Runoff Using the Curve Number Method.

	Mean (mm)		Standard Deviation (mm)		Maximum (mm)			
Micro-Watershed (no. of events)	**Observed**	**Simulated**	**Observed**	**Simulated**	**Observed**	**Simulated**	**NSE**	**PBIAS**
Gegudeguade (31)	3.15	2.48	4.37	5.17	19.45	25.90	0.72	21
Aletu (9)	2.61	1.79	3.94	2.94	10.55	7.16	0.85	32
Shimbraye (13)	1.79	1.26	2.82	2.51	8.70	7.14	0.87	28

surface runoff in Gegudeguade, 68% in Aletu, and 72% in Shimbraye. No significant difference was observed in the accuracy of runoff simulations between different land-use types (i.e., cropping and grazing).

An evaluation of simulated and observed surface runoff using NSE showed reasonable agreement, with an NSE of more than 0.7. Santhi *et al.* (2002) suggest that a NSE value of more than 0.5 is satisfactory for calibrated SWAT simulations. The PBIAS of 21% for the Gegudeguade watershed demonstrated satisfactory model performance, but those in Aletu (32%) and Shimbraye (28%) were unsatisfactory. Moriasi *et al.* (2007) suggest that a calibrated model can be considered satisfactory if the PBIAS value is <25%. It should be noted that our evaluations were for an uncalibrated model, and calibration of the model is assumed to improve the results.

The difference between the simulated and observed surface runoff (i.e., the distance between the triangles and circles in Figure 3) was small in most events. Overall, the surface runoff estimated with the CN method showed underestimations in all of the watersheds, but overestimations in only a few events (Figure 3, subplot).

Evaluating the CN Method for Different Rainfall-Intensity and Antecedent Rainfall Categories

Based on the mean, PBIAS, and proportional error of estimate, the CN method performed well during high rainfall-intensity conditions (Table 5). However, going by Moriasi *et al.* (2007), the PBIAS value of 61% for observed *vs.* simulated runoff during low rainfall-intensity events indicated poor performance. The CN method also performed better at higher rainfall-intensity category according to efficiency criteria such as modified NSE and the relative efficiency criterion.

Conversely, the mean absolute errors, sums of absolute errors, and standard errors of estimate between observations and measurements indicated that the CN method performed better under low rainfall-intensity conditions. However, as the absolute amount of rainfall was lower in low-intensity events, these results did not strongly affect the overall

TABLE 5. Goodness-of-Fit Calculations Comparing Observed and Simulated Surface Runoff Using the Curve Number Method at Low and High Rainfall Intensity under Low and High Antecedent Rainfall Categories.

	Low Rainfall Intensity		High Rainfall Intensity		Low Five-Day Antecedent Rainfall		High Five-Day Antecedent Rainfall	
Goodness-of-Fit Indicators	**Observed**	**Simulated**	**Observed**	**Simulated**	**Observed**	**Simulated**	**Observed**	**Simulated**
Mean	2.08 mm	0.80 mm	4.24 mm	3.61 mm	3.35 mm	2.37 mm	2.68 mm	1.92 mm
PBIAS	61.41		14.85		29.30		28.42	
MAE	1.45		1.76		1.63		1.42	
G	29.08		35.17		32.58		28.45	
RMSE	2.02		2.65		2.42		2.03	
PEE	3.82		2.71		3.40		4.70	
SEE	2.13		2.79		2.56		2.14	
E_j	0.08		0.53		0.44		0.41	
E_{rel}	0.21		0.37		0.65		0.35	

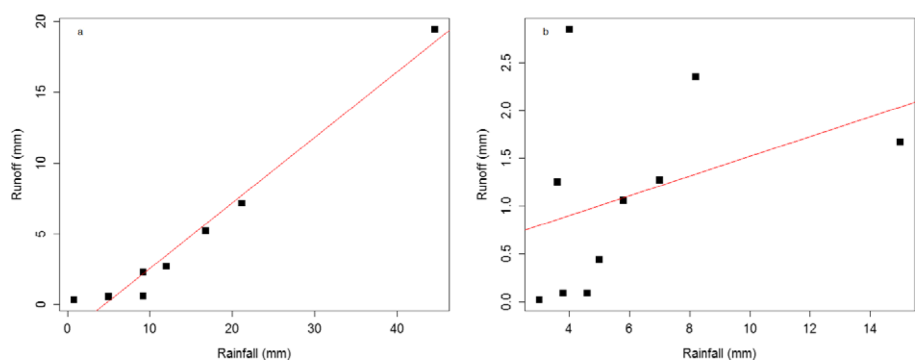

FIGURE 4. The Effect of Five-Day Antecedent Rainfall and Rainfall Intensity on the Rainfall-Runoff Relationship for (a) Events with Low Five-Day Antecedent Rainfall (below the median of 49.4 mm) and High Rainfall Intensity (above the median of 8.4 mm/hr) in Gegudeguade Watershed, and (b) Events with High Five-Day Antecedent Rainfall (above the median) and Low Rainfall Intensity (below the median).

performance of the model. Overall, it appears that the CN method accurately estimated runoff following high rainfall-intensity events, but underestimated runoff at low rainfall-intensity events.

The CN method estimated runoff under both low and high antecedent rainfall categories equally well (Table 5). For example, the PBIAS for both antecedent conditions was ~30% and the modified NSE ~0.4. Some goodness-of-fit indicators were better under low antecedent rainfall category and others under high antecedent rainfall.

Understanding the Rainfall-Runoff Relationship from the Field Observations

Analysis of the field data from Gegudeguade shows that rainfall intensity played a major role in runoff generation during low antecedent rainfall (Figure 4a). Events with higher rainfall intensity and low five-day antecedent rainfall generated more runoff than events with high five-day antecedent rainfall but low rainfall intensity. The slope for the best-fit line of high rainfall intensity and low five-day antecedent rainfall is 0.46 (R^2 = 0.97), which was statistically significant ($p < 0.01$). On the other hand the slope for the best-fit line of high five-day antecedent rainfall and low rainfall intensity is 0.1 (R^2 = 0.14), which was not statistically significant ($p > 0.01$). At low rainfall intensity but with high five-day antecedent rainfall, no correlation was found between rainfall and runoff (Figure 4b).

Further analysis of the observed rainfall and runoff data for Gegudeguade confirms that an increase in antecedent rainfall did not result in an increase in runoff (Figure 5). The slopes of the best-fit line with the “dry” and “average” five-day antecedent rainfall categories are both 0.44 (R^2 of 0.91 and 0.71 for “dry” and “average” categories, respectively), whereas the “wet” five-day antecedent rainfall category has a best-fit line slope of 0.41 (R^2 = 0.65). The slope for the best-fit lines for “dry,” “average,” and “wet” categories was statistically significant ($p < 0.01$).

The Impact of Seasonality on Determining CN Curves

Monsoonal climates commonly have distinct wet and dry seasons. The CN method is partly able to account for this by using three curves to represent the rainfall-runoff relationship for dry, average, and wet antecedent wetness conditions. In our study, the average annual CN for the simulation from 1 January 1990 to 31 December 2012 was 81. Taking the average annual CN to be the same as CN_2 (i.e., the CN value representing the average antecedent wetness condition), the corresponding CN values for dry conditions (CN_1) and wet conditions (CN_3) were 64 and 92, respectively. When the rainfall-runoff curves for these CNs were plotted over the measured and

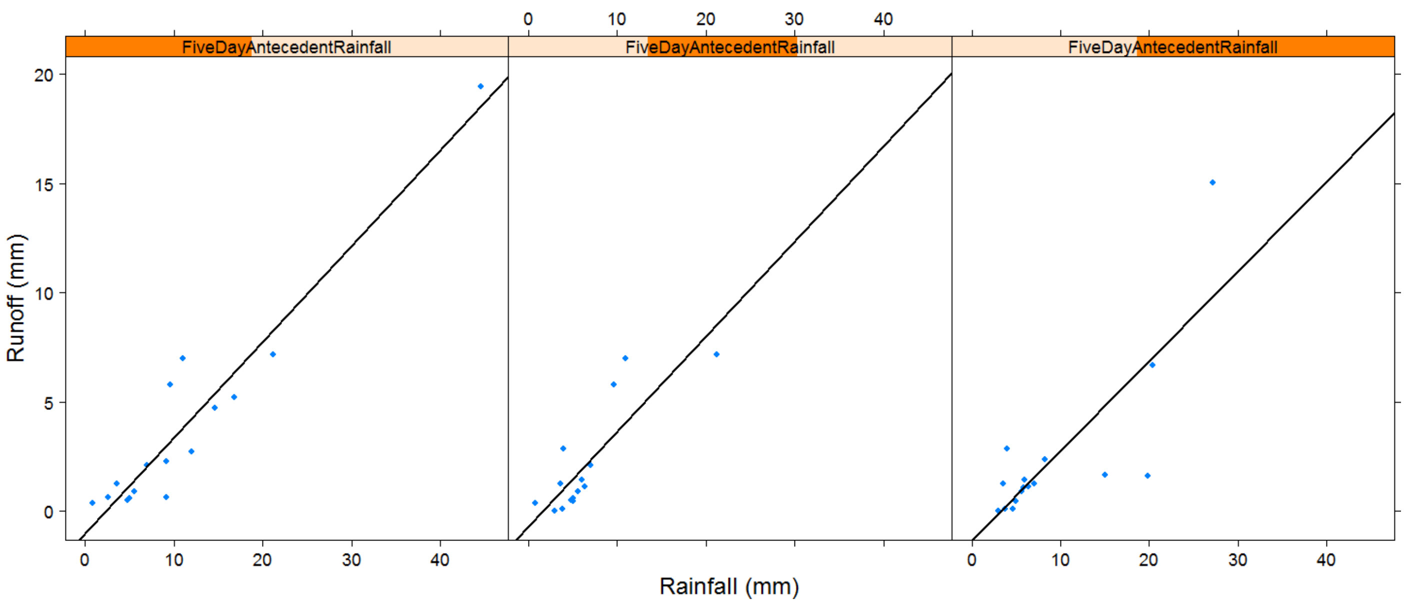

FIGURE 5. Rainfall-Runoff Relationships under Three Average Five-Day Antecedent Rainfall Categories. The left-hand panel, showing 4.1-9.9 mm of rainfall (“dry” antecedent wetness conditions), has a slope of 0.44 (R^2 = 0.91). The middle panel, showing 8.4-13.1 mm of rainfall (“average” antecedent wetness conditions), has a slope of 0.44 (R^2 = 0.91). The right-hand panel, showing 9.9-17.9 mm of rainfall (“wet” antecedent wetness conditions), has a slope of 0.41 (R^2 = 0.65).

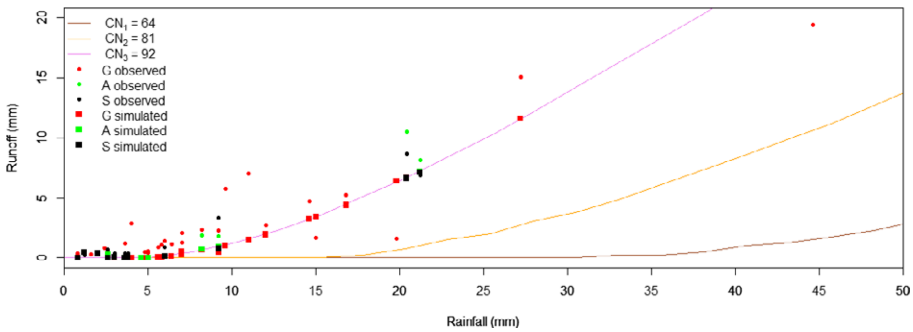

FIGURE 6. Observed *vs*. Simulated Rainfall-Runoff Plots for Three Watersheds in the Lake Tana Basin: Gegudeguade (G), Aletu (A), and Shimbraye (S). The rainfall-runoff curves are with Soil and Water Assessment Tool curve number CN_2, CN_1, and CN_3.

simulated surface runoff, most simulated runoff events were actually situated over the CN_3 curve (Figure 6). This is perhaps not surprising, considering the fact that the simulations were all for events during the wet season. However, the observed surface runoff was higher than can be computed by any of these three CNs, which explains why runoff was on average underestimated by the CN method in the analysis above (without calibration of the model).

The retention parameter values (S) for the observed event rainfall-runoff data were calculated using Equation (3) in SI. These S values were then used to calculate the corresponding CN values for each rainfall-runoff event, using Equation (4) in SI. The calculated CN values for these events ranged from 83 to 99, with a median of 95. The median CN that divides the data into two equal numbers of points located above and below the CN curves is associated with CN_2 (Hjelmfelt, 1991; Woodward *et al.*, 2002). Therefore, the CN_2 for the observed rainfall-runoff data is 95, quite different from the CN_2 of 81 used in the analysis.

To investigate how measurements would be distributed if a CN_2 of 95 was used, the corresponding CN values for dry (CN_1) and wet (CN_3) antecedent wetness conditions were first determined (88 and 98, respectively) and the observed events were then superimposed onto these calculated curves (Figure 7). With these new curves, most of the observations fall between the CN_1 and CN_3 curves, which is in line with the CN method's conceptualization that the CN_1 and CN_3 curves represent extreme runoff situations (Hjelmfelt, 1991).

DISCUSSION

Previous attempts to estimate surface runoff in the Upper Blue Nile basin using the CN method have yielded successful results (e.g., Setegn *et al.*, 2010a, b; Betrie *et al.*, 2011). However, most of these studies used total measured streamflow (i.e., base flow + surface runoff) to evaluate the model simulations, and operated on a larger scale. Moreover, the results of these studies were after model calibration. These approaches tell little about surface runoff estimation using the CN method. Our study tested the performance of the CN method in estimating surface runoff

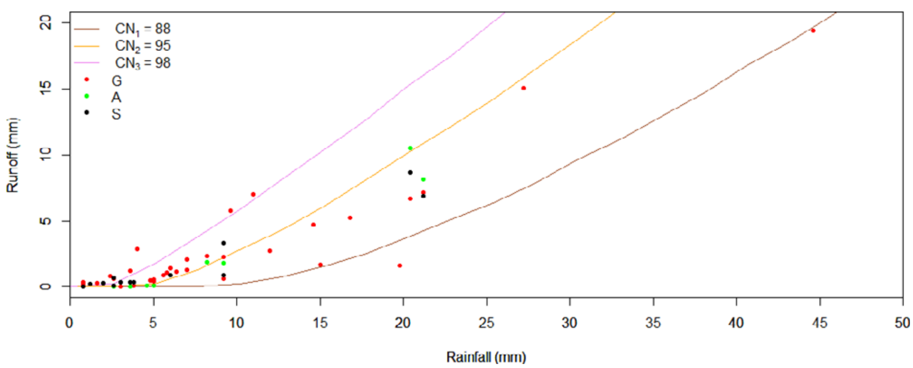

FIGURE 7. Observed Rainfall-Runoff Data and Rainfall-Runoff Curves for Three Watersheds in the Lake Tana Basin: Gegudeguade (G), Aletu (A), and Shimbraye (S). The dots represent rainfall *vs*. observed runoff combinations. The curves are according to calculated CN_1, CN_2, and CN_3 based on the observed rainfall and runoff data.

in three small watersheds without manipulating CN values. Studies in large watersheds involve different processes including interflow and bypass flows. Focusing on small watersheds, we were able to measure surface runoff and compare it with simulated surface runoff with the CN method in the SWAT modeling environment.

In our study, the CN method appears to have performed reasonably well in estimating surface runoff. No difference in the accuracy of runoff estimations was found between different land-use types. The NSE values were more than 0.7 and the PBIAS values were <32%. Moriasi *et al.* (2007) suggest that a model is considered as satisfactory if the NSE value is more than 0.5, and PBIAS value is <25% for a monthly time-step simulation. Although the PBIAS values for our results were higher than this, Moriasi *et al.*'s (2007) suggestion is for calibrated models, whereas the objective of our study was to evaluate the CN method without any parameter manipulation. Calibrating the CN to reflect uncertainties in the soil data, soil residue cover conditions, and tillage practices could improve the flow simulations further. The CN computations from our measured rainfall-runoff data (Figure 7) suggest that choosing a higher CN through calibration would have improved the estimates.

Our results suggest that the CN method performs better in high rainfall-intensity situations, but underestimates runoff in low rainfall-intensity situations when used in tropical regions. One probable explanation is that the SWAT model updates the CN daily based on estimated soil moisture conditions, and surface runoff is estimated at a daily time step. In practice, events can happen within hours and the effect of antecedent soil moisture on surface runoff will be immediate and high. In these conditions, the CN method in SWAT cannot keep track of the soil moisture content sufficiently well because of difference in time step, and could underestimate surface runoff in low rainfall-intensity situations. Perhaps more importantly, water may runoff the rocky and surface-crusted soil found in some parts of the watersheds even in low rainfall-intensity situations (Figure 8), before the model would recognize sufficient soil moisture to produce runoff (cf. Steenhuis *et al.*, 2009). Future research could explore how to better represent tropical soils and/or modify the CN method to account for such low infiltrability, which is common in tropical soils.

FIGURE 8. Photo Showing Rocky Surfaces in the Shimbraye Watershed. The picture also shows the location of the weir and the boundary of the watershed. Photo was taken on July 22, 2012, and photo credit goes to Yihun Dile.

There was no observable difference in the performance of the CN method under low and high antecedent rainfall categories. An increase in five-day antecedent rainfall did not result in an increase in runoff generation. This contradicts the CN method's original conceptualization that runoff increases as five-day antecedent rainfall increases, and suggests that five-day antecedent rainfall alone is not a good indicator to use to adjust the CN for different antecedent wetness conditions. In the Gegudeguade watershed, rainfall intensity proved to be a better indicator of the antecedent wetness condition than the average five-day antecedent rainfall. Liu *et al.* (2008) has also shown that when the rainy monsoon progresses the same storm in the beginning of the rainy phase produces less runoff than at the end of rainy phase. Other research in the Ethiopian highlands (e.g., Bayabil *et al.*, 2010; Tilahun *et al.*, 2015) also showed that formations of saturated areas at the valley bottom of the watersheds play a dominant role in generating surface runoff. They suggested that saturated excess principles could better help to understand the hydrological processes in the Ethiopian highlands. Other factors such as rainfall duration, rainfall amount, and evaporation may also have an effect on the antecedent wetness conditions. However, data on these factors may not be readily available for hydrological modeling purposes. The approach implemented in SWAT — determining the retention parameter directly from the soil moisture content or plant evapotranspiration — might address these issues, which could partly explain the relatively good agreement between measured surface runoff and surface runoff estimated with the CN method within SWAT environment.

Our findings demonstrate that process understanding is needed to truly trust the application of the CN method. However, the fact that the CN method worked better in high rainfall-intensity conditions suggests that it can be useful for water management in tropical regions. The average runoff amount in low rainfall-intensity conditions is small, resulting in only small errors in the estimates of absolute runoff, despite the current problems with underestimation. For water management applications such as irrigation, it is the highest rainfall-intensity events, generating a large amount of runoff, that count as they help to fill reservoirs and ponds. We therefore conclude that the CN method is useful to estimate surface runoff for water management purposes in the tropics at the watershed scale, although there is scope for improvement in runoff estimates at low rainfall intensities. Such a modification of the CN method should be based on an understanding of the hydrological processes determining the rainfall-runoff relationships in different landscape settings (Steenhuis *et al.*, 2009).

CONCLUSIONS

Based on our analysis, the CN method can simulate surface runoff at the watershed scale in tropical regions with a satisfactory level of accuracy. We found that the CN method performed equally well in simulating surface runoff during both low and high antecedent rainfall conditions characteristic of monsoonal climates. Another common feature of monsoonal climates is high rainfall intensity. The CN method could satisfactorily simulate runoff for high rainfall-intensity events; however, at low rainfall intensity, runoff was underestimated. A possible explanation could be low infiltrability due to crust formation or rocky surfaces, common in tropical soils, and formation of saturated areas at the valley bottom of the watershed (Bayabil *et al.*, 2010; Tilahun *et al.*, 2015). Therefore, a modification of the CN method to account for low soil infiltrability and saturated areas in the valley bottom of the watershed could be developed in the future to further improve the accuracy of the CN method at low rainfall intensities in tropical watersheds.

SUPPORTING INFORMATION

Additional supporting information may be found online under the Supporting Information tab for this article: The curve number (CN) method within SWAT: A theoretical background.

ACKNOWLEDGMENTS

Field equipment was bought with support from a Swedish Research Council for the Environment, Agricultural Sciences and Spatial Planning (FORMAS) center of excellence grant under the Resilience and Sustainability: Integrated Research on Social-Ecological Systems project. The authors would like to thank Dr. Jennie Barron for her help in raising this grant and Dr. Melesse Temesegne for help in designing the field research. We appreciate the useful comments from Dr. Prasad Daggupati on the earlier version of the manuscript. The first author was sponsored by FORMAS, and was at Texas A&M University Spatial Sciences Laboratory while doing part of the analysis and drafting the manuscript. We appreciate the two reviewers for their constructive comments.

LITERATURE CITED

Abbaspour, K.C., J. Yang, I. Maximov, R. Siber, K. Bogner, J. Mieleitner, J. Zobrist, and R. Srinivasan, 2007. Modelling Hydrology and Water Quality in the Pre-Alpine/Alpine Thur Watershed Using SWAT. Journal of Hydrology 333:413-430.

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams, 1998. Large Area Hydrologic Modeling and Assessment Part I: Model Development. Journal of the American Water Resources Association 34:73-89.

Bayabil, H.K., S.A. Tilahun, A.S. Collick, B. Yitaferu, and T.S. Steenhuis, 2010. Are Runoff Processes Ecologically or Topographically Driven in the (Sub) Humid Ethiopian Highlands? The Case of the Maybar Watershed. Ecohydrology 130:126-130, DOI: 10.1002/eco.

Betrie, G.D., Y.a. Mohamed, A. van Griensven, and R. Srinivasan, 2011. Sediment Management Modelling in the Blue Nile Basin Using SWAT Model. Hydrology and Earth System Sciences 15:807-818.

Blume, T., E. Zehe, and A. Bronstert, 2010. Rainfall-Runoff Response, Event-Based Runoff Coefficients and Hydrograph Separation. Hydrological Sciences Journal 52:843-862.

CGIAR-CSI (CGIAR Consortium for Spatial Information), 2012. SRTM 90 m Digital Elevation Data [WWW Document]. http://srtm.csi.cgiar.org/, *accessed* November 2012.

Collick, A.S., Z.M. Easton, T. Ashagrie, B. Biruk, S. Tilahun, E. Adgo, S.B. Awulachew, G. Zeleke, and T.S. Steenhuis, 2009. A Simple Semi-Distributed Water Balance Model for the Ethiopian Highlands. Hydrological Processes 23:3718-3727.

Cronshey, R., R.H. McCuen, N. Miller, W. Rawls, S. Robbins, and D. Woodward, 1986. Urban Hydrology for Small Watersheds—TR-55. U.S. Department of Agriculture (USDA), Washington, D.C.

Daggupati, P., N. Pai, S. Ale, K.R. Doulgas-Mankin, R. Zeckoski, J. Jeong, P. Parajuli, D. Saraswat, M. Youssef, 2015. A Recommended Calibration and Validation Strategies for Hydrological and Water Quality Models. Transactions of ASABE 58(6):1705-1719.

Dessie, M., N.E.C. Verhoest, V.R.N. Pauwels, E. Adgo, J. Deckers, J. Poesen, and J. Nyssen, 2015. Water Balance of a Lake with Floodplain Buffering: Lake Tana, Blue Nile Basin, Ethiopia. Journal of Hydrology 522:174-186, DOI: 10.1016/j.jhydrol.2014.12.049.

Dile, Y.T., R. Berndtsson, and S.G. Setegn, 2013. Hydrological Response to Climate Change for Gilgel Abay River, in the Lake Tana Basin—Upper Blue Nile Basin of Ethiopia. PLoS ONE 8: e79296, DOI: 10.1371/journal.pone.0079296.

Dile, Y.T., L. Karlberg, P. Daggupati, R. Srinivasan, D. Wiberg, and J. Rockström, 2016. Assessing the Implications of Water Harvesting Intensification on Upstream-Downstream Ecosystem Services: A Case Study in the Lake Tana Basin. Science of the Total Environment 542:22-35, DOI: 10.1016/j.scitotenv.2015.10.065.

Dile, Y.T. and R. Srinivasan, 2014. Evaluation of CFSR Climate Data for Hydrologic Prediction in Data-Scarce Watersheds: An Application in the Blue Nile River Basin. Journal of the American Water Resources Association 50(5):1226-1241, DOI: 10.1111/jawr.12182.

ENMSA (Ethiopian National Meteorological Services Agency), 2012. Meteorological Data: The Ethiopian National Meteorological Services Agency, Addis Ababa, Ethiopia.

FAO (Food and Agriculture Organization), 1995. World Soil Resources: An Explanatory Note on the FAO World Soil Resources Map at 1:25 000 00 Scale. FAO, Rome, Italy.

Gebremicael, T.G., Y.a. Mohamed, G.D. Betrie, P. van der Zaag, and E. Teferi, 2013. Trend Analysis of Runoff and Sediment Fluxes in the Upper Blue Nile Basin: A Combined Analysis of Statistical Tests, Physically-Based Models and Landuse Maps. Journal of Hydrology 482:57-68, DOI: 10.1016/j.jhydrol.2012.12.023.

Green, I.R.A. and D. Stephenson, 1986. Criteria for Comparison of Single Event Models. Hydrological Sciences Journal 31:395-411.

Gupta, H., S. Sorooshian, and P. Yapo, 1999. Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. Journal of Hydrologic Engineering 4:135-143.

HEC-HMS, 2000. Hydrologic Modeling System HEC-HMS. Technical Reference Manual. U.S. Army Corps of Engineers, Hydrologic Engineering Center.

Hjelmfelt, A.T., 1991. Investigation of Curve Number Procedure. Journal of Hydraulic Engineering Division of the American Society of Civil Engineers 117:725-737.

Jewell, T.K., D.D. Adrian, and T.J. Nunno, 1978. Methodology for Calibrating Stormwater Models. Journal of Environmental Engineering, ASCE 104:485-501.

Kebede, S., Y. Travi, T. Alemayehu, and V. Marc, 2006. Water Balance of Lake Tana and Its Sensitivity to Fluctuations in Rainfall, Blue Nile Basin, Ethiopia. Journal of Hydrology 316:233-247, DOI: 10.1016/j.jhydrol.2005.05.011.

Krause, P., D.P. Boyle and F. Bäse, 2005. Comparison of Different Efficiency Criteria for Hydrological Model Assessment. Advances in Geosciences 5:89-97.

Krysanova, V. and J.G. Arnold, 2008. Advances in Ecohydrological Modelling with SWAT—A Review. Hydrological Sciences Journal 53:939-947.

Krysanova, V., F. Hattermann, and F. Wechsung, 2005. Development of the Ecohydrological Model SWIM for Regional Impact Studies and Vulnerability Assessment. Hydrological Processes 19:763-783.

Legates, D.R. and G.J. McCabe, 1999. Evaluating the Use of "Goodness-of-fit" Measures in Hydrologic and Hydroclimatic Model Validation. Water Resources Research 35:233-241.

Liu, B.M., A.S. Collick, G. Zeleke, E. Adgo, Z.M. Easton, and T.S. Steenhuis, 2008. Rainfall-Discharge Relationships for a Monsoonal Climate in the Ethiopian Highlands. Hydrological Processes 22:1059-1067.

Manley, E.R., 1978. Calibration of Hydrological Model Using Optimization Technique. Journal of the Hydraulics Division, ASCE 104:189-202.

Moriasi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, and T.L. Veith, 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE 50:885-900.

MoWR (Ministry of Water Resources), 2009. Spatial Data: The Ethiopian Ministry of Water Resources. MoWR, Addis Ababa, Ethiopia.

Nash, J.E. and J.V. Sutcliffe, 1970. River Flow Forecasting Through Conceptual Models: Part 1. — A Discussion of Principles. Journal of Hydrology 10:282-290.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams, 2012. Soil and Water Assessment Tool Theoretical Documentation: Version 2009. Grassland Soil and Water Research Laboratory, Agricultural Research Service, Blackland Research Center, Texas Agricultural Experiment Station, Temple, Texas.

NRCS (Natural Resources Conservation Service), 2004. Estimation of Direct Runoff from Storm Rainfall, Part 630. *In*: Hydrology: National Engineering Handbook, The U.S. Department of Agriculture, Washington, D.C., p. 79.

Patry, G. and M. Mariño, 1983. Nonlinear Runoff Modeling: Parameter Identification. Journal of the Hydraulics Division, ASCE 109:865-880.

Ponce, B.V.M. and R.H. Hawkins, 1996. Runoff Curve Number—Has It Reached Maturity? Journal of Hydrologic Engineering 1:11-19.

Rientjes, T.H.M., B.U.J. Perera, A.T. Haile, P. Reggiani, and L.P. Muthuwatta, 2011. Regionalisation for Lake Level Simulation—The Case of Lake Tana in the Upper Blue Nile, Ethiopia. Hydrology and Earth System Sciences 15:1167-1183, DOI: 10.5194/hess-15-1167-2011.

Rostamian, R., A. Jaleh, M. Afyuni, S.F. Mousavi, M. Heidarpour, A. Jalalian, and K.C. Abbaspour, 2008. Application of a SWAT Model for Estimating Runoff and Sediment in Two Mountainous Basins in Central Iran. Hydrological Sciences Journal 53:977-988.

Santhi, C., J.G. Arnold, J.R. Williams, W.A. Dugas, R. Srinivasan, and L.M. Hauck, 2002. Validation of the SWAT Model on a Large River Basin with Point and Nonpoint Sources. Journal of the American Water Resources Association 37:1169-1188.

Sarkar, D., 2008. Lattice: Multivariate Data Visualization with R. Springer, Seattle, Washington, DOI: 10.1007/978-0-387-75969-2.

Schuol, J., K. Abbaspour, R. Srinivasan, and H. Yang, 2008. Estimation of Freshwater Availability in the West African Sub-Continent Using the SWAT Hydrologic Model. Journal of Hydrology 352:30-49.

Setegn, S.G., B. Dargahi, R. Srinivasan, and A.M. Melesse, 2010a. Modeling of Sediment Yield from Anjeni-Gauged Watershed, Ethiopia Using Swat Model. Journal of the American Water Resources Association 46:514-526.

Setegn, S.G., R. Srinivasan, B. Dargahi, and A.M. Melesse, 2009. Spatial Delineation of Soil Erosion Vulnerability in the Lake Tana Basin, Ethiopia. Hydrological Processes 23:3738-3750.

Setegn, S.G., R. Srinivasan, A.M. Melesse, and B. Dargahi, 2010b. SWAT Model Application and Prediction Uncertainty Analysis in the Lake Tana Basin, Ethiopia. Hydrological Processes 24 (3):357-367.

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 Sediment for the Abay (Blue Nile) with a Simple Model. Hydrological Processes 3737:3728-3737.

Stehr, A., P. Debels, F. Romero, and H. Alcayaga, 2008. Hydrological Modelling with SWAT under Conditions of Limited Data Availability: Evaluation of Results from a Chilean Case Study. Hydrological Sciences Journal 53:588-601.

Tessema, S.M., S.W. Lyon, S.G. Setegn, and U. Mörtberg, 2014. Effects of Different Retention Parameter Estimation Methods on the Prediction of Surface Runoff Using the SCS Curve Number Method. Water Resources Management 28:3241-3254, DOI: 10.1007/s11269-014-0674-3.

Tilahun, S.A., C.D. Guzman, A.D. Zegeye, D.C. Dagnew, A.S. Collick, B. Yitaferu, and T.S. Steenhuis, 2015. Distributed Discharge and Sediment Concentration Predictions in the Sub-Humid Ethiopian Highlands: The Debre Mawi Watershed. Hydrological Processes 29:1817-1828, DOI: 10.1002/hyp.10298.

Tilahun, S.A., C.D. Guzman, A.D. Zegeye, T.A. Engda, A.S. Collick, A. Rimmer, and T.S. Steenhuis, 2013a. An Efficient Semi-

Distributed Hillslope Erosion Model for the Subhumid Ethiopian Highlands. Hydrology and Earth System Sciences 17:1051-1063, DOI: 10.5194/hess-17-1051-2013.

Tilahun, S.A., R. Mukundan, B.A. Demisse, T.A. Engda, C.D. Guzman, B.C. Tarakegn, Z.M. Easton, A.S. Collick, A.D. Zegeye, E.M. Schneiderman, J.-Y. Parlange, and T.S. Steenhuis, 2013b. A Saturation Excess Erosion Model. Transactions of the ASABE 56:681-695, DOI: 10.13031/2013.42675.

Wale, A., T.H.M. Rientjes, A.S.M. Gieske, and H.A. Getachew, 2009. Ungauged Catchment Contributions to Lake Tana's Water. Hydrological Processes 3693:3682-3693.

White, E.D., Z.M. Easton, D.R. Fuka, A.S. Collick, E. Adgo, M. McCartney, S.B. Awulachew, Y.G. Selassie, and T.S. Steenhuis, 2011. Development and Application of a Physically Based Landscape Water Balance in the SWAT Model. Hydrological Processes 25:915-925.

Williams, J.R., N. Kannan, X. Wang, C. Santhi, and J.G. Arnold, 2012. Evolution of the SCS Runoff Curve Number Method and Its Application to Continuous Runoff Simulation. Journal of Hydraulic Engineering Division of the American Society of Civil Engineers 17:1221-1229.

Williams, J.R., A.D. Nicks, and J.G. Arnold, 1985. Simulation for Water Resources in Rural Basins. Journal of Hydraulic Engineering Division of the American Society of Civil Engineers 111:970-986.

Winchell, M., R. Srinivasan, and M.J. Di Luzio, 2013. ArcSWAT 2.3.4 Interface for SWAT2012. Blackland Research and Extension Center at Texas A&M AgriLife Research and Grassland, Soil and Water Research Laboratory at USDA Agricultural Research Service, Temple, Texas.

Woodward, D.E., R.H. Hawkins, A.T. Hjelmfelt, J.A. Van Mullem, and Q.D. Quan, 2002. Curve Number Method: Origins, Applications and Limitations. *In*: Hydrologic Modeling for the 21st Century: 2nd Federal Interagency Hydrologic Modelling Conference, Las Vagas, Nevada, pp. 1-10.

Yu, B., 2012. Validation of SCS Method for Runoff Estimation. Journal of Hydrologic Engineering 17(11):1158-1163.