Hydrologic Modelling of Data Scarce Basin with SWAT Model: Capabilities and Limitations

Martine Nyeko

Abstract Hydrologic modelling is pre-requisite to water resources management. Unfortunately, hydrologic modelling in data scare basin has always been difficult. The current study, explored the use of "data limited" model Soil Water Assessment Tool (SWAT) in modelling lower Aswa basin located in northern Uganda. The study adopted different techniques in generating and estimating various missing model parameters and input especially solar radiation, saturated soil hydraulic conductivity, available soil water content, Universal Soil Lost Equation erodibility factor and moist soil albedo. Soil Water Assessment Tool model was then manually calibrated using monthly historical streamflow records. The calibration was successful with coefficient of determination (R^2) value of 0.618 and the Nash and Sutcliffe efficiency value of 0.47. Validation of the calibrated model using independent dataset shows even better model performance with Nash and Sutcliffe efficiency value of 0.64 and coefficient of determination (R^2) value of 0.56. Successful calibration of hydrologic model Soil Water Assessment Tool under the data scarcity still proves the potential of the application of the model even in data limited basin, but more especially by water resources managers who needs understanding of existing condition and modelling possible future.

1 Introduction

Water resources management requires modeling of various hydrological processes like infiltration, runoff generation, groundwater recharge, evapotranspiration, and non-hydrological processes like vegetation and crop growth, nitrate and phosphorous dynamics, erosion, sewage system dynamics, water regulations, among others. All these processes combined constitute a complex systems, which require complex models that demand significant amount of input field data and hydrologic observation. There are currently several different hydrologic models with different complexities available to simulate hydrologic processes at watersheds, including physical models, conceptual models and empirical model.

The Système Hydrologique Européen (SHE) (Abbott et al. 1986a and b) is one of a widely used physically based fully distributed watershed models that has several advantages such as

M. Nyeko (✉)
Department of Biosystems Engineering, Faculty of Agriculture and Environment, Gulu University, P. O. Box 166, Gulu, Uganda
e-mail: m.nyeko@gu.ac.ug

providing location specific outputs. The SHE model divide the watershed into cells, this capability allows the model to accommodate significant spatial detail. However, the distributed and physical based nature of SHE model requires that in each application study, vast amount of data and parameters describing the physical characteristics of the watershed are available. The data availability will in any case determines the degree of reliability of the model results. Hamilton (2007), discusses the lack of data to operate hydrologic models and the problems this has created with respect to making decisions from model outputs. Semi-distributed hydrologic model such as the Soil and Water Assessment Tool (SWAT) (Arnold et al. 1998) and Hydrologic Engineering Center, Hydrologic Modeling System (HEC-HMS) (US-Army Corps of Engineers 2000) have also been widely used at different level of complexity in modelling both hydrologic and non-hydrologic processes regarding water management. The Hydrologic Engineering Center's Hydrologic Modelling System (HEC-HMS) for example has been used to simulate precipitation-runoff processes and reservoir operations (Fleming and Neary 2004). The model is capable of integrating features and environment that includes a database, data-entry utilities, computation engine, and results reporting tools (US-Army Corps of Engineers, 2000).

SWAT model is a spatially semi-distributed conceptual hydrological model, which can operate on both daily time-step, monthly or even annually for long term simulation. The model basically require three sets of data; Digital Elevation Model (DEM) necessary for elevation and definition of watershed geomorphology; soil; and land use data. Weather data can be generated by the model during simulation using inbuilt weather generator or provided as input. For validation and calibration purposes, the watershed must be gauged. SWAT model has also been widely used to predict the impact of management on water, sediment, and agricultural chemical yields (Gassman et al. 2007).

In practice however, it is often difficult to determine the capabilities, operational characteristics, and limitations of any hydrologic model just from the documentation, without actual application. In other words, there is no "best" model or no "easy-to-use" model which require low data input and which provide accurate results under all scenarios. Site-specific, tailor-made approaches are therefore needed to supplement model inabilities.

The current study assess the capabilities and limitations of SWAT model in modelling watershed that has limited field and hydrologic data for possible use in water resources management. Different techniques were adopted in generating and estimating various missing model parameters and hydrologic observation including solar radiation, saturated soil hydraulic conductivity, available soil water content, Universal Soil Lost Equation erodibility factor and moist soil albedo.

2 Study Area and Data Management

2.1 Study Area

The study was conducted in Aswa basin located in northern Uganda (Fig. 1). Altitude in the basin ranges between 870 and 1908 m above sea level and slope is gentle with most part (>97 %) having slope less than 20 %. Water resources development and management program in the basin are centered on groundwater, which provides potable water to domestics and livestock. Increased demand for water is expected to be high especially in the agricultural sector, to boost food production and in the rapidly growing urban centers. However, uncertainty remain high in the basin concerning the water resources availability and reliability. Limited attempt has been made to study the hydrologic processes in this basin due to inadequate field data and hydrologic observations required for hydrologic model setup and simulations.
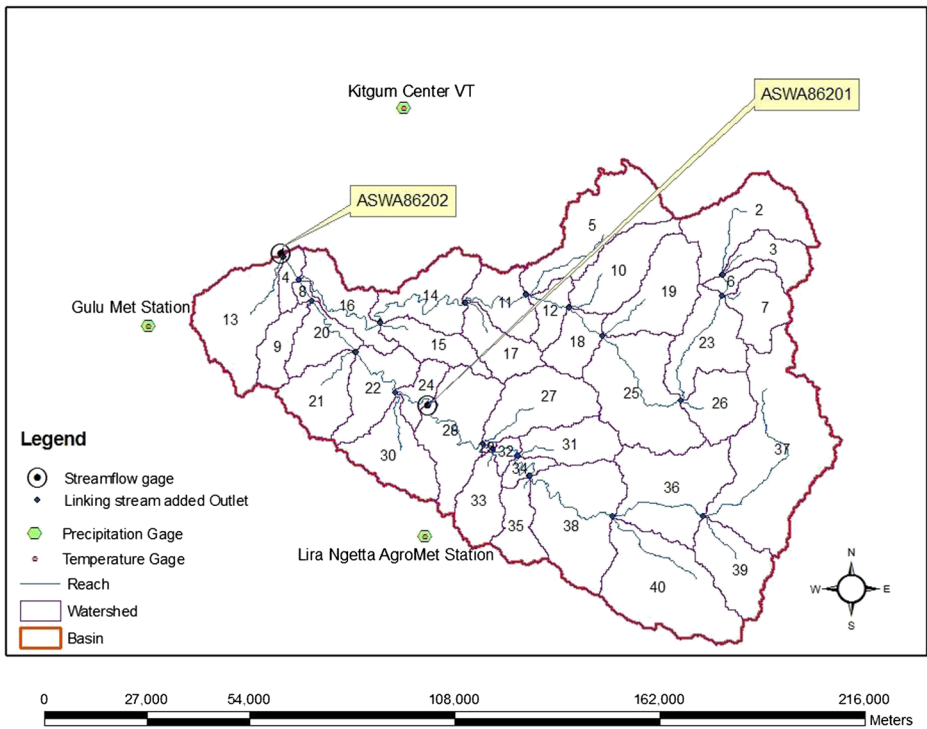
Fig. 1 The sub-basin delineation of Aswa basin showing weather station & streamflow stations

2.2 Data Management

SWAT model requires three key sets of data: land use, soil and elevation data in addition to climatic and stream flow data, which is optional. Land use map was derived from 1986 LANDSAT scene using the spectrally based supervised image classification. six land cover classes were identified and reclassified to match SWAT land cover and crop growth database that is; agricultural (generic), forest (mixed), range land (brush, grass, and semi-arid), wetland (mixed), urban (low density) and water.

The Soil and Terrain Database for north-eastern Africa (SEA), in a CD-ROM at a scale of 1:1,000,000 according to FAO, was used to derive the soil units and some soil properties.

HydroSHED DEM which is derived from Shuttle Radar Topography Mission SRTM at 3 arc-second (approximately 90 m) resolution was downloaded from the SRTM website (http://srtm.csi.cgiar.org/). The DEM was used to delineate the watershed and to derive spatial sub-basin data such as slope gradient, slope length and stream network characteristics.

Daily river flow data were available for two gauges (ASWA86201 and ASWA86202) for the observation period between 1960 and 1978. The streamflow observation were portioned into calibration data (1970 to 1974) and validation data (1975 to 1978).

2.2.1 Estimation of Missing Solar Radiation

Solar radiation data covering the simulation periods (1970 to 1978) were missing and yet it is an important parameter in estimation of evapotranspiration according to Penman-Monteith method.

The missing records of solar radiation were estimated using Angstrom empirical equation (Allen et al. 1998). Historical monthly solar radiation data for 7 years (1965 to 1975) were used to calibrate the Angstrom equation. Due to limited records of sunshine hours, solar radiation estimated were used in the derivation of solar radiation parameter for the SWAT weather generator, which was later used to generate missing solar radiation during SWAT model simulation.

Angstrom empirical equation relates solar radiation to extraterrestrial radiation as;

$$R_s = \left(a + \frac{b \times n}{N}\right)Ra \tag{1}$$

where a is regression constant, expressing the fraction of extraterrestrial R_a radiation reaching the earth on overcast days ($n=0$), and $a+b$ fraction of extraterrestrial radiation reaching the earth on clear days ($n=$N). According to Allen et al. (1998), Angstrom empirical formula for radiation yields fairly good results, as there is a strong link between sunshine hours and net radiation received compared to other methods such as the one proposed by Hargreaves and Samani (1982).

To be able to obtain parameter values for a and b in Eq. 1 using simple linear regression, a linear transformation of the equation was performed, with parameter a representing the y-intercept and parameter b presenting the gradient (Eq. 2). The ration of the solar radiation (R_s) to the extraterrestrial radiation (R_a) was correlated with the relative sunshine duration ($\frac{n}{N}$).

$$\frac{R_s}{R_a} = a + b\left(\frac{n}{N}\right) \tag{2}$$

The extraterrestrial radiation R_a was estimated using the relationships;

$$R_a = \frac{24(60)}{\pi}G_{sc}d_r[\omega_s\sin(\phi)\sin(\delta) + \cos(\phi)\sin(\omega_s)] \tag{3}$$

where

R_a	is extraterrestrial radiation (MJm-2 day-1)
G_{sc}	is the solar constant =0.0820 MJm-2 min-1
d_r	is the inverse relative distance Earth-Sun (Eq. 4)
ω_s	is the sunset hour angle (Eq. 6) (rad)
ϕ	is the latitude (rad)
δ	is the solar decimation (Eq. 5)

$$d_r = 1 + 0.033\cos\left(\frac{2\pi}{365}J\right) \tag{4}$$

$$\delta = 0.409\sin\left[\frac{2\pi}{365}J - 1.39\right] \tag{5}$$

J is the number of the day in the year between 1 (1 January) and 365 or 366 (31 December)

$$\omega_s = ar\cos[-\tan(\phi)\tan(\delta)] \tag{6}$$

The parameters (a and b) in Eq. 2 were estimated by plotting$\frac{R_s}{R_a}$ verses the relative sun shine duration (n/N). From the plot, the values of the parameters were obtained as a=0.219 and b= 0.4297 (Fig. 2). The regression coefficient or coefficient of determination was 75 %. Allen

et al. (1998) proposed the values of $a=0.25$ and $b=0.5$ for use where no actual solar radiation are available and no calibration has been carried to improve a & b parameters. However, in the current study, the coefficient of determination (75 %), is considered adequate and therefore the estimated values of parameters a and b are considered appropriate and used in estimation of missing solar radiation values.

2.2.2 Estimation of Saturated Soil Hydraulic Conductivity and Available Soil Water Content

The SEA soil database however does not contain all the soil parameter necessary for SWAT model setup and simulation. Missing soil parameters were derived from the harmonised world soil database (version 1.1 2009) http://www.fao.org/fileadmin/templates/nr/documents/ HWSD/HWSD_Documentation.pdf; and soils of Northern Province published by Department of Agriculture Uganda.

Textural classes for the different soil unit were derived using the information provided by SEA soil database, harmonized world soil database and the publication of soil of Northern Province. The textural classes were used to identify the composition of percentage sand, clay and silt using table extracted from Ahuja et al. (1993) (Table 1).

A known correlation between composition of sand, clay and silt in textural class, bulk density and organic matter developed by Saxton and Rawls (2006) was used to estimate available water capacity, and saturated hydraulic conductivity.

2.2.3 Estimation of Universal Soil Lost Equation USLE Erodibility (K) Factor

The USLE erodibility (K) factor was calculated using Williams (1995) equation and the textural classes derived previously as input. That is;

$$K_{USLE} = f_{csand} \times f_{cl-si} \times f_{hisand} \qquad (7)$$

where, f_{csand} is the factor that gives low soil erodibility for soils with high coarse-sand contents and high values for soil with little sand, given by the Eq. (8)

$$f_{csand} = \left(0.2 + 0.3 \times \exp\left[-0.256 \times m_s \times \left(1 - \frac{m_{silt}}{100}\right]\right) \qquad (8)$$

m_s is the percent sand content (0.05–2.00 mm) and m_{silt} is the percent silt content (0.002-0.05 mm),

$$f_{cl-si} = \left(\frac{m_{silt}}{m_c - m_{silt}}\right)^{0.3} \qquad (9)$$

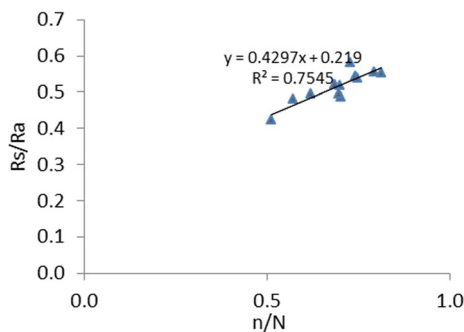**Fig. 2** Calibration of angstrom formula

Table 1 Descriptive statistics for percentage sand and clay content

Soil type	Sand				Clay			
	$\bar{x}$	s	CV	n	$\bar{x}$	s	CV	n
Clay	14.9	10.7	71.6	1177	55.2	10.9	19.7	1177
Clay loam	29.8	5.9	19.7	1317	32.6	3.7	11.4	1317
Loam	40.0	6.5	16.3	1991	19.7	5.2	26.3	1991
Loamy sand	80.9	3.8	4.6	881	6.4	3.2	50.1	881
Silt	5.8	4.5	77.2	115	9.5	2.7	28.9	115
Silt loam	16.6	11.7	70.8	3050	18.5	5.9	31.6	3050
Silty clay	6.1	4.5	73.5	1002	46.3	4.9	10.7	1002
Silty clay loam	7.6	5.3	70.7	1882	33.2	3.7	11.1	1882
Sand	92.7	3.7	4.0	803	2.9	2.0	67.1	803
Sandy caly	47.5	3.9	8.2	74	41.0	4.5	10.9	74
Sandy clay loam	54.3	7.3	13.5	610	27.4	4.0	14.6	610
Sandy loam	63.4	7.9	12.5	2835	11.1	4.8	43.2	2835

Here, $\bar{x}$ is the mean, s, the standard deviation; CV, coefficient of variation (percent); and n, the sample size

m_c is percent clay content (<0.002 mm),

$$f_{org} = \left(1 - \frac{0.25 \times orgC}{orgC + \exp[3.72 - 2.95 \times orgC]}\right) \tag{10}$$

$orgC$ is the percent organic carbon content for the soil layer (%) and

$$f_{hisand} = \left(1 - \frac{0.7 \times \left(1 - \frac{m_s}{100}\right)}{\left(1 - \frac{m_s}{100}\right) + \exp\left[-5.51 + 22.9 \times \left(1 - \frac{m_s}{100}\right)\right]}\right) \tag{11}$$

2.2.4 Estimation of Moist Soil Albedo (r)

The moist soil albedo r was estimated from Landsat 5TM image, using the reflectance corrected values for atmospheric effect (ρ_λ) and weighting coefficient ω_λ according to D'Urso (2001), using the equation;

$$r = \sum_{\lambda} \omega_\lambda \rho_\lambda \tag{12}$$

3 Modelling Methodology

3.1 SWAT Model Setup

The SWAT project was setup using ArcSWAT GIS interface. Watershed delineation and parameterization of stream reaches and sub-basin geomorphology was automatically done

using the interface. DEM based stream definition was used to derive flow direction and accumulation. With a minimum drainage set at 16000 ha or 160Km^2 and watershed outlet at ASWA86202 (Fig. 1), 40 sub-basins were delineated.

When formulating and applying distributed models, the concepts of nonlinearity of hydrologic response must be taken into account (Beven 2001). In nonlinear systems, extremes of any distribution of responses may be important in controlling the observed response. This means that hydrologic model should be described at much smaller scale in order to capture all the local heterogeneities such as infiltration rates, preferential flows, areas of first saturation and others local extremes responses (Beven 1995).

In SWAT model, natural homogenous areas referred to as hydrologic response unit (HRU), that assumes non-variability of the data and parameters within its delineation was introduced as necessary notion in hydrologic modeling (Arnold et al. 1998). The objective of HRU definition was to reduce the heterogeneities due to climate, soil types, topography and geology that influence hydrologic response.

In this study, the HRU definition was done using a combination of 1 % land use area over sub-basin, 1 % soil class over land use area and 1 % slope class over soil area, after the land use and soil were imported, reclassified and overlaid with slope class. With these combinations, a total of 630 HRUs were defined.

The crop database and management file were edited to match the land use type in the watershed, and the management techniques in place. The development of the LAI (Maximum value and the pattern), was used to guide the modification of the crop growth database. Management techniques were scheduled based on the heat unit theory (Boswell 1926; Magoon and Culpepper 1932). The heat unit scheduling was in particular useful in this case study because land use in the basin are generic in nature and there exist dissimilar climate across the watershed.

SWAT weather generator developed by Sharpley and Williams (1990), to simulate missing climatic records during simulations was adopted. The default weather generator parameters were replaced with custom weather parameters. The custom weather parameters generated for three weather stations (Gulu, Lira and Kitgum) using the historical weather records are mean daily maximum and minimum temperature for the month (12 months), standard deviation for daily maximum and minimum air temperature in the month, average or mean total monthly precipitation, standard deviation for daily precipitation in month, skew coefficient for daily precipitation in the months, probability of a wet day following a dry day in the month, probability of wet day following wet day in the month, average numbers of day of precipitation in the month, maximum 0.5 h rainfall in the entire period of record for month, average daily solar radiation for the month, average daily dew point temperature in the month and average daily wind speed in month.

3.2 Sensitivity Analysis of SWAT Parameters

Identifying parameters that do or do not have any significant influence on the model simulation is crucial not only in reducing parameter uncertainty but also in reducing over parameterization of the model, which can destroy its physical representation. In this paper, sensitivity analysis was performed to determine SWAT model parameters that are very sensitive to streamflow prediction. The in-built sensitivity analysis tool in the model interface ArcWAT developed by van Griensven (2005) was used. The method uses dimensionless index to express the sensitivity of a parameter. The equation calculate the ratio between the relative changes of model output to relative change of a parameter.

$$I = \frac{x}{y}\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \tag{13}$$

where x is the parameter and y is the predicted output, x_2, and, y_2, correspond to ±10 % of the initial parameter and corresponding output values, respectively (James and Burges 1982). The greater the value of I, the more sensitive a model output variable is to that particular parameter.

3.3 SWAT Model Calibration

3.3.1 Parameter Selection

Although sensitivity analysis is always performed, no commonly accepted guideline has been established on the sensitivity bound. It is therefore important that sensitivity analysis result be aided by knowledge on key hydrologic process within the watershed in identifying model parameters for calibration. This is important in avoiding uncertainty due to unrepresented hydrologic processes in the watershed. According to Yang et al. (2012), it is important to identify key physical processes and parameter interactions, and examine how different model setups affect model simulation in order to ensure the validity of calibration. In this paper, the following parameters were therefore considered in the model calibration after considering both sensitivity analysis result and knowledge of the hydrologic processes (Table 2): Groundwater

Table 2 Sensitivity output using ArcSWAT sensitivity tool including Parameters definition

Rank	Index	Parameter	Definition	Process
1	1.04	Esco	Soil evaporation compensation factor	Evaporation
2	0.98	Cn2	SCS curve number for moisture condition II	Runoff
3	0.74	Gwqmn	Threshold depth in shallow aquifer required for return flow	Groundwater
4	0.28	Alpha_bf	Base-flow alpha factors	Groundwater
5	0.225	Sol_Awc	Available soil water capacity	Soil
6	0.17	Sol_Z	Soil depth	Soil
7	0.09	GW_Revap	Groundwater 'revap' coefficient	Groundwater
8	0.065	Canmx	Maximum canopy index	Runoff
9	0.0588	Revapmn	Threshold depth of water in shallow aquifer for revap to occur	Groundwater
10	0.057	Ch_K2	Channel effective hydraulic conductivity	Channel
11	0.04	Blai	Leave area index for crops	Crops
12	0.024	GW_Delay	Groundwater delay	Groundwater
13	0.021	Sol_K	Soil conductivity	Soil
14	0.018	Ch_N2	Manning's n value for main channel	Channel
15	0.063	Slope	Average slope steepness	Geomophology
16	0.0059	Epco	Plant evaporation compensation factor	Evaporation
17	0.0023	Slsubbsn	Average slope lenght	Geomophology
18	0.002	Surlag	Surface runoff lag coefficient	Runoff
19	0.0016	Sol_Alb	Soil albedo	Evaporation
20	0	Biomix	Biological mixing effeciency	Soil
21	0	Smtmp	Snow melt base temperature	Snow
22	0	Smfmn	Minimum melt rate for snow during the year	Snow
23	0	Smfmx	Maximum melt rate for snow	Snow
24	0	Timp	Snow pack temperature lag factor	Snow
25	0	Tlaps	Temperature laps rate	Geomorphology

'Revap' coefficient (Gw_Revap) and Maximum canopy index (Canmx), which scored poorly in sensitivity analyses were considered together with five other parameters that scored highly in sensitivity analyses in model calibration. Gw_Revap is known to control movement of water from shallow aquifer into unsaturated layer and this affects evapotranspiration losses, which is an important processes. In addition, in the event that rainfall is under estimated, which can be possible scenario, this parameter compensate for the unavailable water to be evaporated. In highly vegetated watershed, which is the case for the watershed under investigation, canopy storage can be significant hydrologic process, affecting infiltration, surface runoff and evapo-transpiration, which was the bases for choosing maximum canopy index (Canmx). A total of seven model parameters, five parameters with high sensitivity index and two parameters with low sensitivity were calibrated.

SWAT model was manually calibrated. Automatic calibration was considered cumbersome as it require significant runtime. In any case, with the knowledge of the hydrologic processes in the watershed and the ability to make decision on which parameter to introduce in the calibration, manual calibration is handy. Visual analysis of the simulated and measured streamflow hydrograph was used during the manual calibration, the "sensitive parameters" to inform modification of parameter bounds and introduction of new parameter in the calibration.

Streamflow recorded in the period of 1970 and 1974 was used as observed model output in the calibration. The performance of the model in predicting the output during manual calibration was evaluated using both statistical and graphical methods. In particularly, the graphical techniques (streamflow hydrograph), was used to provides a visual comparison of the simulated and measured data, identify model bias, identify the differences in timing and the magnitude of peak flows and shapes of recess curves (Moriasi et al. 2007). It was possible to identify the next parameter to optimize to improve on the predicted streamflow using visual analysis of streamflow hydrograph.

The standard regression with slope and y-intercept of the best fit regression line was used to provide the statistical measure of the fit of the calibration process. In this approach, the slope is used to indicate the relative relationship between simulated and measured values, and the y-intercept to indicate the presence of lag or lead between model prediction and measured data. As the slope approaches 1 and y-intercept approaches 0 the calibration process may be considered to have converged to an optimal parameter set. The statistical coefficient of determination (R^2) was also used. The value of R^2 ranges from 0 to 1, with higher values indicating less error variance, and typical values greater than 0.5 considered acceptable (Santhi et al. 2001 and Van Liew et al. 2003).

Nash-Sutcliffe (Nash and Sutcliffe 1970) efficiency (NSE), was also used to determine the relative magnitude of the residual variance ("noise") to the measured data variance ("information") and how well the plot of the observed data versus the simulated data fits the 1:1 line. One observation was considered and the NSE is computed as;

$$g = 1 - \frac{\sum_{i=1}^{n} (Q_m - Q_s)^2}{\sum_{i=1}^{n} \left(Q_m - \overline{Q_m}\right)^2} \tag{14}$$

where, $= Q_m$ measured discharge, Q_s = simulated discharge

The calibration was then considered successful after no significant improvement could be realised in any of the above indicators measured by the two statistical methods. Also, the subjectivity in the judgment of the goodness-of-fit of the model simulation is reduced.

3.4 SWAT Model Validation

Calibrated model may fail the verification test on some occasions. Reasons may be due to: 1) errors in the data used in calibration, both the data used as input to the model and the data used to check model output should be checked very carefully (data with large errors should not be used for calibration), 2) use of a period of record that does not contain enough events of the physical processes needed to calibrate key parameters, 3) inadequate and or miss-representation by the model of hydrological processes found in the watershed, model results should be compared visually with the recorded data series to look for consistent variations.

Validation procedures are similar to calibration procedures in that predicted and measured values are compared to determine if the objective function is met. However, a dataset of measured watershed response selected for validation preferably should be different from the one used for model calibration, and the model parameters are not adjusted during validation. Validation provides a test of whether the model was calibrated to a particular dataset or the system it is to represent. If the objective function is not achieved for the validation dataset, calibration and/or model assumptions may be revisited. The study uses independent streamflow dataset recorded in the period between 1975 and 1978 in the validation of calibrated SWAT model.

4 Results and Discussion

4.1 Sensitivity Analysis

Twenty five hydrologic parameters that influence streamflow were used in the sensitivity analysis. Table 2 shows the model parameters and the sensitivity analysis result, ranked with most sensitive parameter in the first row. The most sensitive parameters (Index>0.1) using the objective function according to van Griensven (2005) were; soil evaporation compensation factor (Esco), initial SCS curve number II (Cn2), threshold depth of water in shallow aquifer for return flow to occur (Gwqmn), base-flow factors (Alpha_bf), available soil water capacity (Sol_Awc), and soil depth (Sol_Z).

4.2 Calibration

Three groundwater parameters (Alpha_bf, Gwqmn, Gw-Revap), one soil parameter, (Sol_Awc), one evaporation parameter (Esco) and two runoff parameters (Cn2, Canmx) were considered in model calibration (Table 3). Simulated monthly streamflow verses the observed streamflow together with corresponding monthly rainfall is plotted in Fig. 3. Visual analysis of the monthly hydrographs indicates that the calibrated model slightly overestimate the peak runoff. The hydrograph also showed that the model failed to simulate peak flow between May 1974 and November 1974. This can be due to rainfall data not being representative, localized storm having no response or there is malfunctioning of the gauges (rainfall and flow).

Standard regression plot (Fig. 4) evaluates the calibrated model performance with slope of 0.9 indicating a good relative relationship between simulated and measured streamflow and y-intercept of the best fit regression line of +13.616 indicating the presence of lag between model prediction and measured streamflow. This lag can be attributed to; over estimation of surface roughness, or less than actual slope for over land flow.

The coefficient of determination (R^2) describing the proportion of the variance in the measured data explained by the model was obtained as 0.618. The value of R^2 ranges from

Table 3 Adjustment of parameter values

Parameter	Definition	Unit	Default values	Changed values
Esco	Soil evaporation compensation factor		0	1 (replacement)
Cn2	SCS curve number for moisture condition II	–	Relative to soil hydrologic group and land cover	−15 (add)
Gwqmn	Threshold depth in shallow aquifer required for return flow	mm	0	0.95 (replacement)
Alpha_bf	Base-flow alpha factors	–	0.048	0.65 (replacement)
Sol_AWC	Soil available water capacity	mm	Relative to soil type	X 1.2 (relative)
Canmx	Maximum canopy index	mm	0	10 (replacement)
GW_REVAP	Groundwater 'revap' coefficient	–	0.02	0.2 (replacement)

0 to 1, with higher values indicating less error variance. The reported performance rating for R^2 (Santhi et al. 2001 and Van Liew et al. 2003), indicate that typical values of R^2 greater than 0.5 is acceptable. The Nash-Sutcliffe efficiency (NSE) of 0.47 was obtained for monthly calibration. The performance rating of NSE for SWAT model calibration in the ranges of 0.54 to 0.65 was reported as adequate. However, considering that the measured data, (streamflow and climatic data) are highly uncertain, the performance of the calibrated model can be considered good if the rating of NSE is relaxed (Moriasi et al. 2007).

The challenges faced in calibration of SWAT model was in particular the quality of input data, especially precipitation and streamflow. Precipitation data had considerably missing values. The hydrograph of observed and simulated monthly streamflow after model calibration shows inconsistence prediction of streamflow peaks between May and November 1974. The inconsistency in prediction could partly be due to unrepresentativeness of the input precipitation data between these periods. In addition, land use dataset used in calibration was dataset for much later year (1986) compared to the streamflow observation, which was earlier in 1970's. A considerable land use change could have therefore occurred in the 10 year time lag. Land use dataset in the period of 1970's could not be prepared due to lack of data necessary for the preparation. In addition, no streamflow records exist in 1980's to be used in calibration. The best was to use streamflow of 1970's and land use dataset of 1986.
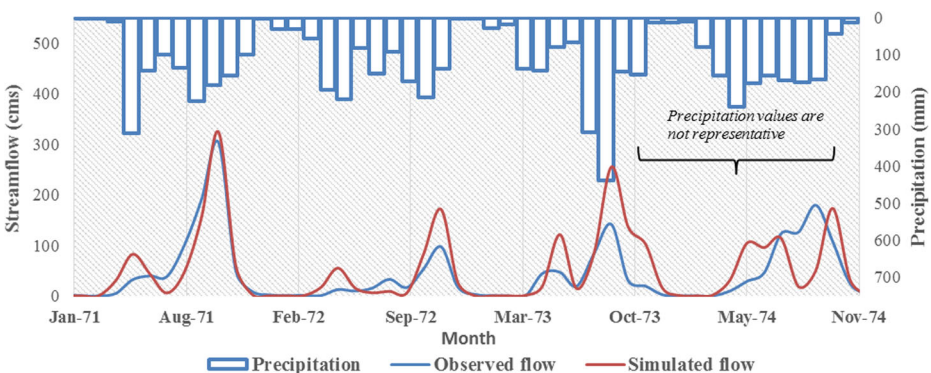

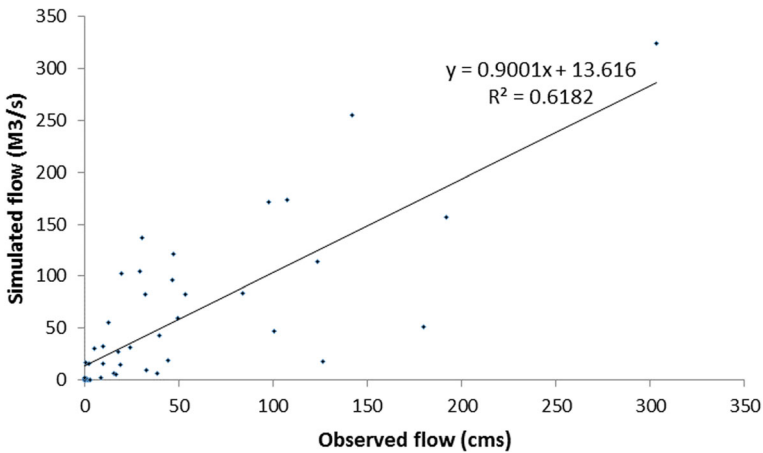**Fig. 3** Hydrograph of observed and simulated monthly streamflow after model calibration

Fig. 4 Regression correlation of observed and simulated monthly streamflow

It also seems that not all processes were being modeled in the basin, especially the processes regulating runoff and evaporation losses. The model seems to be underestimating evapotranspiration losses and over estimating runoff. The underestimation of the evapotrans-piration could also be attributed to the inadequate water available to meet the evapotranspira-tion demand. The real cause is probably underestimation of precipitation. To increase the water available for evapotranspiration, the Gw_Revap coefficient, which controls the water move-ment from the shallow aquifer into unsaturated layer, was adjusted to allow more water from shallow aquifer to flow to unsaturated zone.

4.3 Model Validation Results

The model validation was conducted using climatic data set for the period of 1975 to 1978. Evaluations of model performance during validation are presented in Figs. 5 and 6. The hydrograph (Fig. 6) indicates that the model consistently predicts the measured streamflow, but
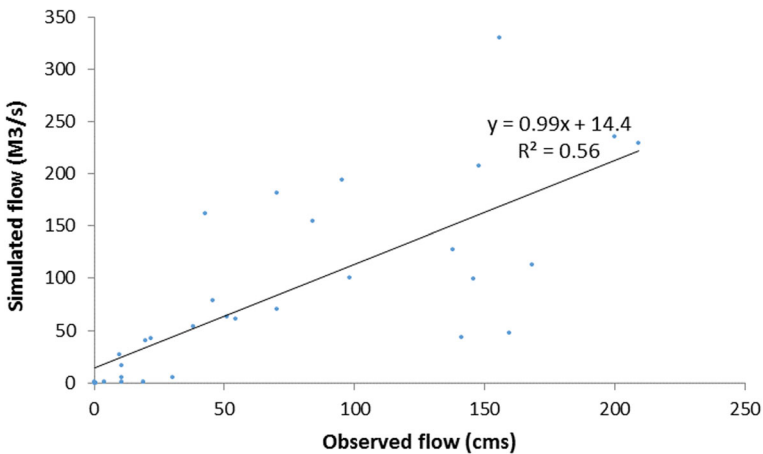

Fig. 5 Regression correlation of observed and simulated monthly streamflow during validation
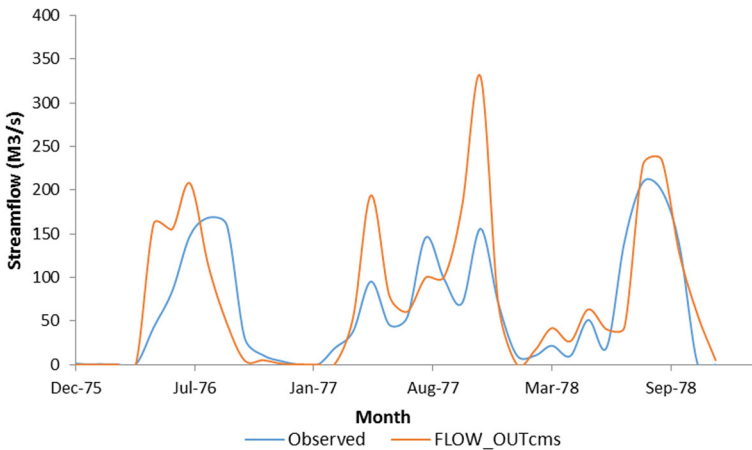
Fig. 6 Hydrograph of observed and simulated monthly streamflow during model validation

with some lags. The visual evaluation of the hydrograph plot showed fairly good model match in validation period and peak flow is still over estimated.

Statistical evaluation of model performance during validation using standard regression plot (Fig. 5) indicates a good relationship between simulated and measured streamflow with the slope of 0.99 and the y-intercept of the best fit regression line of +14.4, which indicate lag between model prediction and measured streamflow. The values of 0.56 for R^2 obtained indicate a good model fit during validation. Above all, the objective function, the Nash-Sutcliffe efficiency (NSE) of 0.64 indicates that the model performance during validation is satisfactory.

The better model performance during validation with NSE =0.64 could partly be due to the representativeness of the dataset including the land use and the complete initialization (warm up) of the model, which was limited during the calibration period. The warm-up period allows the model to "stabilize" or calculate values that become initial values for the period of interest. Therefore, after the warm-up period, the model is considered to represent conditions in the watershed. A warm period that is too short could significantly alter flow regime. No agreed warm up period however has been established, but always it is advisable to complete more than one simulation cycle. In the current study only one simulation cycle, 1970 was considered warm up period.

5 Conclusion

The study adopted different techniques in generating and estimating various missing model parameters and input especially solar radiation, saturated soil hydraulic conductivity, available soil water content, USLE erodibility factor and moist soil albedo. SWAT model was then manually calibrated using monthly historical streamflow records. The calibration was success-ful with coefficient of determination (R^2) value of 0.618 and the Nash and Sutcliffe efficiency value of 0.47. Validation of the calibrated model using independent dataset shows even better model performance with Nash and Sutcliffe efficiency value of 0.64 and coefficient of determination (R^2) value of 0.56. Successful calibration of hydrologic model SWAT under the data scarcity still proves the potential of the application of the model even in data limited

basin, but more especially to water resources managers who needs understanding of existing condition and modelling possible future.

References

Abbott MB, Bathurst JC, Cunge JA, O'Connell PE, Rasmussen J (1986a) An introduction to the European hycirological system-systeme hydrologique European 'SHE' 1: history and philosophy of a physically-based, distributed modeling system. J Hydrol 87:45–59

Abbott MB, Bathurst JC, Cunge JA, O'Connell PE, Rasmussen J (1986b) An introduction to the European hydrological system-systeme hydrologique Europeen, 'SHE' 2: structure of physically-based, distributed modeling system. J Hydrol 87:61–77

Ahuja LR, Brakensiek DL, Shirmohammadi A (1993) Infiltration and soil movement. In D. R. Maidment (Editor), Handbook of Hydrology. McGraw-Hill, New York (U.S.A.)

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. Water Resources Development and Management Service, FAO Rome

Arnold JG, Srinivasan R, Muttiah RS, Williams JR (1998) Large area hydrologic modeling and assessment. Part I. Model development. J Am Water Resour Assoc 34:73–89

Beven KJ (1995) Linking parameters across scales: sub-grid parameterisations and scale dependent hydrological models. Hydrol Process 9:507–526

Beven KJ (2001) Dalton medal lecture: how far can we go in distributed hydrological modelling? Hydrol and Earth Syst Sci 5(1):1–12

Boswell VG (1926) The influence of temperature up on the growth and yield of garden peas. Proc of Am Soc of Hortic Sci 23:162–168

D'urso G (2001) Simulation and management of on-demand irrigation systems: a combined agrohydrological and remote sensing approach. PhD Thesis, Wageningen University, Wageningen, The Nertherlands

Gassman PW, Reyes MR, Green CH, Arnold JG (2007) The soil and water assessment tool: historical development, applications, and future research directions. Trans ASABE 50(4):1211–1250. doi:10.13031/2013.23637

Hamilton S (2007) Completing the loop: from data to decisions and back to data. Hydrol Process 21(22):3105–3106

Hargreaves GH, Samani ZA (1982) Estimating potential evapotranspiration. J. Irrig. and Drain. Engr. ASCE 108(IR3):223–230

James LD, Burges SJ (1982) Selection, calibration, and testing of hydrologic models. In: Haan CT, Johnson HP, Brakensiek DL (eds) Hydrologic modeling of small watersheds. ASAE Monograph, St. Joseph, Michigan

Magoon CA, Culpepper CW (1932) Response of sweet corn to varying temperatures from time of planting to canning maturity, U.S.D. A Technical Bulletin, 312

Moriasi DN et al (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, T. ASABE 50:885–900

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models, part I - a discussion of principles. J Hydrol 10:282–290

Santhi C, Arnold JG, Williams JR, Hauck LM, Dugas WA (2001) Application of a watershed model to evaluate management effects on point and nonpoint source pollution. Trans Am Soc Agric Eng 44(6):1559–1570

Saxton KE, Rawls W (2006) Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Sci Soc Am J 70:1569–1578

Sharpley AN, Williams JR (1990) Erosion/productivity impact calculator, 1. Model documentation. USDA-ARS Technical Bulletin 1768, Washington, DC

Van Griensven A (2005) Sensitivity, auto-calibration, uncertainty and model evaluation in SWAT2005. Unpublished report

Van Liew MW, Schneider JM, Garbrecht JD (2003) Stream flow response of an agricultural watershed to seasonal changes in rainfall. In: Renard KG, McElroy SA Gburek WJ, Canfield HE, Scott RL (eds) Proceedings, 1st interagency conference on research in the watershed. U.S. Department of Agriculture, Agricultural Research Service, Benson, Arizona

Williams JR (1995) The EPIC model. SWAT models, Input/Output file documentation, Version 2005

Yang J, Liu Y, Yang W, Yaning C (2012) Multi-objective sensitivity analysis of a fully distributed hydrologic model WetSpa. Water Resour Manag 26(1):109–128