Hydrol. Earth Syst. Sci. Discuss., 9, 3761–3788, 2012
www.hydrol-earth-syst-sci-discuss.net/9/3761/2012/
doi:10.5194/hessd-9-3761-2012

Hydrology and Earth System Sciences Discussions

Critical review of the application of SWAT in the upper Nile Basin countries

A. van Griensven[1,2], P. Ndomba[3], S. Yalew[1], and F. Kilonzo[1,2,4]

[1]UNESCO-IHE Institute of Water Education, P.O. Box 3015, Delft, The Netherlands
[2]Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
[3]University of Dar es Salaam, P.O. BOX 35131, Dar es Salaam, Tanzania
[4]Kenyatta University, P.O. BOX 43844, Nairobi, Kenya

Received: 15 February 2012 – Accepted: 19 February 2012 – Published: 20 March 2012

Correspondence to: A. van Griensven (a.vangriensven@unesco-ihe.org)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The Soil and Water Assessment Tool (SWAT) is a hydrological simulation tool that is widely applied within the Nile basin. Up to date, more than 20 peer reviewed papers describe the use of SWAT for a variety of problems in the upper Nile basin countries, such as erosion modeling, land use modeling, climate change impact modeling and water resources management. The majority of the studies are clustered in the tropical highlands in Ethiopia and around Lake Victoria. The popularity of SWAT is attributed to the fact that the tool is freely available and that it is readily applicable through the development of Geographic Information System (GIS) based interfaces and its easy linkage to sensitivity, calibration and uncertainty analysis tools. The online and free availability of basic GIS data that are required for SWAT made its applicability more straight forward even in data scarce areas. However, the easy use of SWAT may not always lead to knowledgeable models. In this paper, we aim at critically reviewing the use of SWAT in the context of the modeling purpose and problem descriptions in the tropical highlands of the Nile Basin countries. A number of criteria are used to evaluate the model set-up, model performances, physical representation of the model parameters, and the correctness of the hydrological model balance. On the basis of performance indicators, the majority of the SWAT models were classified as giving satisfactory to very good results. Nevertheless, the hydrological mass balances as reported in several papers contained losses that might not be justified. Several papers also reported unrealistic parameter values. More worrying is that many papers lack this information. For this reason, it is difficult to give an overall positive evaluation to most of the reported SWAT models. An important gap is the lack of attention that is given to the vegetation and crop processes. None of the papers reported any adaptation to the crop parameters, or any crop related output such as leaf area index, biomass or crop yields. A proper simulation of the land cover is important for obtaining correct runoff generation, evapotranspiration and erosion computations. It is also found that a comparison of SWAT applications on the same or similar case study but by different research teams and/or

model versions resulted in very different results. It is therefore recommended to try to find better methods to evaluate the representativeness of the distributed processes and parameters, especially when land use studies are envisaged or predictions of the future through environmental changes. The main recommendation is that more details on the model set-up, the parameters and outputs should be provided in the journal papers or supplementary materials in order to allow for a more stringent evaluation of these models.

1 Introduction

The Soil and Water Assessment Tool (SWAT) is a physically based, spatially distributed, continuous time hydrological model (Arnold et al., 1998). Major modules in the model include hydrology, erosion/sedimentation, plant growth, nutrients, pesticides, land management, stream routing, and pond/reservoir routing. The SWAT modeling tool simulates, among others, climate changes, hydrologic processes, land use changes, water use management, water quality and water quantity assessments. SWAT requires a number of basin specific input data encompassing different components such as weather, hydrology, erosion/sedimentation, plant growth, nutrients, pesticides, agricultural management, channel routing, and pond/reservoir routing. Weather inputs (i.e. precipitation, maximum and minimum temperature, relative humidity, wind speed, solar radiation) are required on a daily temporal resolution, although recent versions of the model allow hourly input files. SWAT is imbedded in several GIS interfaces (e.g. ArcGIS, OpenMap, Grass, etc.) that allow to discretise a basin into sub-basins. Each subbasin contains river reaches and one set of weather inputs. The sub-basin is further subdivided into Hydrological Response Units that are identified on the basis of similar land use, soil type and slope classes.

Over 600 peer-reviewed journal papers related to the SWAT model have been reported (Gassman et al., 2010). Besides its obvious advantage as a hydrological modelling tool that includes modularity, computational efficiency, ability to predict long term

impacts as a continuous model, and ability to use readily available global datasets, availability of a reliable user and developer support has contributed to its acceptance as one of the most widely adopted and applied hydrological models worldwide (Gassman et al., 2010). The Nile countries are no exceptions in adopting SWAT as a hydrological modeling tool. As much as authors advocate the use of SWAT as a modeling tool, they have concerns on whether the reported methods and approaches in fact, help achieve their reported goals. The purpose of this review, therefore, is to evaluate various models that have been reported in peer-reviewed journal papers in the upper Nile countries by looking at their used approaches and methods with respect to what they state to achieve. In order to do so, the authors follow several fit-for-purpose (how usefull is the model for its purpose), fit-to-observation (how well do the model outputs fit to field observations), and fit-to-reality (how well do the models represent the physical processes) evaluation criteria designed for measuring strength/weakness of the various SWAT models the journals were based on.

2 Case study and model descriptions

The Nile river drains an area of 2.9 million km^2 that covers 10 % of the African continent with its spread over 11 “Nile countries”: Egypt, Sudan, South-Sudan, Ethiopia, Eritrea, Uganda, Tanzania, Kenya, Burundi, Rwanda and DR Congo. With a course of 6695 km it is the longest river in the world. The two major tributaries are the Blue Nile, stemming from Lake Tana in Ethiopia and flowing to Sudan, and the White Nile, from Lake Victoria in the East African Community. Lake Victoria is fed by several tributaries: Kagera, Yala, Sondu, Nyando, Mara, Mbalageti, Simiyu and Konga rivers.

The Victoria Nile leaves Lake Victoria at the site of the now-submerged Owen Fall in Uganda and rushes for 483 km over rapids and cataracts until it enters Lake Albert. The river leaves Lake Albert as the Albert Nile through northern Uganda and at the South-Sudanese border it becomes the Bahr al Abyad or the White Nile.

The Blue Nile is locally called Abbay river when it leaves Lake Tana and flows through the Ethiopian plateau in an immense curve and pours itself out of the mountains in the hot plain of South Sudan where it is called the Bahr al Azraq. The Blue Nile and the White Nile join each other in Khartoum to form the Nile river that flows Northeast. After 322 km the Nile river is joined by the Atbara river and continues its course up to Egypt where it enters Lake Nasser and flows further downstream to enter the Nile Delta before reaching the Mediterranean sea.

This paper reviews the applications of SWAT within the tropical highlands of the Nile countries (i.e. Tanzania, Uganda, Kenya, Rwanda and Ethiopia) and includes river basins that are not located within the upper Nile watershed. For that purpose, peer reviewed papers have been reviewed in the previously mentioned countries, up to the year 2011. More than 20 peer-reviewed papers were identified out of which more than half are located in Ethiopia which are listed in Table 1 according to their topic. The main results of these papers are summarized below per topic.

2.1 Model calibration, parameterization and validation

Jayakrishnan et al. (2005) modeled the hydrology of the 3050 km^2 Sondu river basin in Kenya using land use, soil and elevation data with limited spatial resolution (1–10 km^2). The objective was to assess impacts of land use changes as a result of changes to intensive dairy farming. The simulation Nash-Sutcliffe efficiency coefficient (NSE) of <0.1 was attributed to inadequate rainfall and other model input data. The use of one rain gauge station situated at the upper end of the catchment was not representative of the basin.

Mulungu and Munishi (2007) calibrated the SWAT model for the 11 000 km^2 Simiyu catchment in Tanzania using improved spatial inputs for land use and soil. The study used land use map developed from Landsat thematic mapper images to coincide with the period of available flow data. Local soil and geological maps were used to augment the SOTER 1:2 000 000 global database. The sensitivity analysis was performed for 16 parameters with the Latin hypercube one at a time (LH-OAT), and autocalibration

with the shuffled complex evolution (SCE) algorithm. Although resulting total water yield and surface runoff fractions of the water balance were within ±1 % of the observed flow, the base flow fraction was off target by 50 %. Improving the spatial resolution of the soil and land use inputs did not improve the model performance, which resulted in an NSE of 0.4. Although no particular factors were attributed to this poor performance, the authors recommended the use of improved spatial distribution of rainfall.

In modeling the hydrology of the Mitano river basin in Uganda, Kingston and Taylor (2010) used the gridded 0.5° CRUTS3.0 database as the climatic input. Although there was a good agreement between observed and simulated monthly means and flow duration curves, the model performance after calibration was poor, resulting in an NSE of −0.09. According to the author, the poor performance in the hydrological modeling was attributed to "model-observation divergences with the calibration period that are simply too large to be resolved by an auto-calibration routine".

Setegn et al. (2008) used SWAT to model the hydrological water balance of the Lake Tana Basin in Ethiopia with the objective of testing the performance of the SWAT model for stream flow prediction. These authors calibrated and validated on four tributaries of Lake Tana using SUFI-2, GLUE and ParaSol algorithms. This paper reported that the SWAT model was more sensitive to HRU definition thresholds than to sub-basin discretization. Further, the paper reported that uncertainties were captured by bracketing more than 60 % of the observed river discharge.

Mekonnen et al. (2009) with the intention of developing a generic rainfall-runoff model that would be better suited to Ethiopian catchments, used a spectrum analysis method to extract evidence regarding the relationships between different temporal scales of available daily rainfall and runoff series which reflect the temporal and spatial scales of discharge in two watersheds in Ethiopia. The paper reported that frequencies in rainfall and stream discharge longer than 50 days had a sufficient coherence to warrant model calibration.

Tibebe and Bewket (2010) assessed surface runoff generation and soil erosion rates for a small watershed in the Awash River basin of Ethiopia using the SWAT model.

Comparing monthly predicted runoff against the measured values, the study demonstrates that distribution of observed and simulated runoff was quite uniform throughout the simulation period, with a high correlation value ($r = 0.831$). It further reports a NSE of 0.789 to show that the model was able to generate monthly runoff close to the observed.

Notter et al. (2011) applied the SUFI2 algorithm within SWAT-CUP to perform the calibration and validation in 8 and 6 stations respectively. NSE values were higher than 0.5 for 7 out of the 8 gauges during calibration and 4 out of the 6 gauges during validation.

2.2 Land use change

In analyzing the impacts of land cover change on runoff for the Nzoia basin in Kenya, Githui et al. (2009) used plausible "worst" (scenario 1) and "best" (scenario 2) case scenarios. The emphasis was on "reforestation and sustainable agriculture" for the best-case scenario, and "deforestation and expansion of unsustainable agriculture" for the worst-case scenario. Using the CLUE-S model (Verburg and Veldkamp, 2004) land cover scenarios were generated by using a baseline map as the dependent variable and location factors such as population, elevation, slope, distances to rivers and towns, and lithology in logistic regression (see Fig. 1).

Mango et al. (2011) used three hypothetical scenarios; partial deforestation, complete deforestation to grassland, complete deforestation to agriculture to analyse the sensitivity of model outputs to land use change for a small subbasin (700 km^2) on the Nyangores tributary of the Mara river basin, Kenya. Simulations under all land use change scenarios where forest is converted to agricultural land indicated an increased surface flow and a decreased subsurface flow and average flow over the period of simulation, while evapotranspiration shows a small positive increase. These results are contrary to the results obtained by Githui et al. (2009), where a reduction in forest cover led to a decrease in evapotranspiration, an increase in both surface and base flow and a large increase in water yield.

2.3 Climate change uncertainty and impacts analyses

Kingston and Taylor (2010) explored the impacts of projected climate change on water resources of the Upper Nile basin and the uncertainty associated with such projections of the hydrological change on the 2098 km^2 Matano River basin in Uganda. The assessment included the evaluation of the range of uncertainty due to climate sensitivity, choice of Global Circulation Models (GCMs), and hydrological model parameterization. The authors found an overwhelming dependence upon the GCMs used for climate projections and showed that single GCM evaluations of climate change impacts are likely to be wholly inadequate and potentially misleading as a basis for the analysis of climate change impacts on freshwater resources. On the hydrology, the study found that the proportion of precipitation that contributes to the Mitano river discharge via groundwater will decrease as a result of increasing temperature. The increasing evapotranspiration due to increasing global temperatures (rather than reduced precipitation) limits the amount of water penetrating the soil profile and replenishing the shallow groundwater store during the wet season.

Githui et al. (2009) used the monthly change fields of rainfall and temperature instead of mean annual perturbations to the historical time series or hypothetical scenarios for the 12 709 km^2 Nzoia basin in Kenya, since the region has distinct wet and dry seasons. They used the MAGICC and Scenario Generator (SCENGEN) from the Climatic Research Unit (CRU) of the University of East Anglia to construct climate change scenario based on IPCC A2 and B2 scenarios, for two selected 30-yr periods: 2010–2039 centred on 2020 and 2040–2069 centred on 2050. Five GCMs (CCSR, CSIRO, ECHAM4, GFDL, and HADCM3) selected based on a correlation of greater than 0.7 between the observed and the simulated rainfall and temperature and a small root mean square error were used in this study. Scenarios of future climate were obtained by adjusting the baseline observations by the difference for temperature or percentage change for rainfall between period-averaged results for the GCM experiments (30-yr period) and the simulated baseline period (1981–2000). All the scenarios indicated that temperature

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

would increase in this region, with the 2050s experiencing much higher increases than the 2020s. While the models were consistent with respect to changes in both runoff and base flow, average stream flow seen to increase with rainfall increase, relatively higher amounts were observed in the 2050s than in 2020s. All scenarios indicated higher probabilities to exceed the bankfull discharge than the observed time series.

Mango et al. (2011) developed the regional averages of temperature and precipitation projections from a set of 21 global models in the MMD (multi-model dataset) for the A1B scenario for East Africa. Based on the reported changes in temperature and precipitation, the hydrological model was run for minimum, median and maximum change scenarios. The mean for all projections is a 7 % increase in annual precipitation by 2099, with projections ranging from −3 % to 25 %. Notable is the disproportionately nonlinear response of a large stream flow change occurred by a small change in precipitation. A combined decrease in precipitation and an increase in temperature led to increased evapotranspiration and reduced runoff.

Whereas Githui et al. (2009) argues that stream flow response was not sensitive to changes in temperature, Kingston and Taylor (2010), and Mango et al. (2011) postulated that increases in temperature lead to an increase in evaporation and hence a change in the water balance reducing the stream flow. Interestingly, both Kingston and Taylor (2010), and Mango et al. (2011) used satellite derived climatic data for their input into the hydrological model and baseline, while Githui et al. (2009) built their model on observed climatic data. Another difference between the two sides is the size of the catchments under consideration. On the one hand, the small size of the Mitano and Nyangores catchments at 2098 km^2 and 700 km^2, respectively means that all the components of the hydrological cycle may not be fully reflected, especially the loss of groundwater to shallow and deep aquifer and transfer to downstream sub basins. On the other hand, Githui et al. (2009) simulated a large and complex catchment (>12 000 km^2) which compounds the interactions in the processes and reduces the transfers to other basins.

2.4 Erosion modeling

Swallow et al. (2009) used the SWAT model to estimate sediment yields and changes in sediment yield for the Yara and Nyando basins draining into the Lake Victoria from the Mau region in Kenya. A spatial analysis of tradeoffs and synergies between sediment yield and agricultural production for the year 2005 was generated through a spatial overlay of results on sediment yields and value of agricultural production at the sub-basin level. The Yala and Nyando basins measuring 4000 km^2 and 3000 km^2 respectively have a mix of land tenure types. The authors noted the inability of the SWAT model to consider gully in the Modified Unified Soil Loss Equation, as a potential cause of underestimation of sediment yield especially for soil prone to gully erosion.

Setegn et al. (2010) used SWAT to simulate the sediment yield simulations for the Anjeni, a small watershed (1.35 km^2) in the northern highlands of Ethiopia, using different slope classifications. The annual sediment yields were around 27.8 and 29.5 $t\,ha^{-1}$. The paper showed that the results are highly sensitive to the size of the sub-basins. The obtained erosion parameters were used to model sediment transport in the Lake Tana basin in Ethiopia and gave annual sediment yields that varied spatially between 0 and 65 $t\,ha^{-1}$. Betrie et al. (2011) used SWAT to evaluate effects of several Best Management Scenarios (filter strips, stone bunds, and reforestation) for the Upper Blue Nile Basin in Ethiopia. The results showed a very high spatial variability for the obtained annual sediment yields, ranging from 0 to more than 150 $t\,ha^{-1}$. Easton et al. (2010) simulated the hydrologic balance and sediment loss for the Blue Nile watershed that lies mainly in Ethiopia using SWAT-WB, a modified SWAT model that captures variable source area hydrologic phenomena. Predicted runoff losses (averaged across the entire subbasin) varied from as low as 13 $mm\,yr^{-1}$ for the entire Blue Nile Basins to 44 $mm\,yr^{-1}$ in Anjeni. Very large spatial variations in the computed erosion rates where reported (10 % of the area contributes to 75 % of the total sediment yield).

Tibebe and Bewket (2010) used SWAT for hydrologic and soil erosion predictions for the Keleta watershed in central Ethiopia after calibrating the model against surface runoff that was obtained from flow separation techniques. The annual sediment yield varied between 1.57 and 7.57 $t\,ha^{-1}\,yr^{-1}$ with a long term average of 4.26 $t\,ha^{-1}\,yr^{-1}$.

Muvundja et al. (2009) used an un-calibrated SWAT model to estimate flows and pollutant loads for the 127 streams draining to Lake Kivu; SWAT was used in a supporting role to other techniques that were used for the primary analysis; problems regarding the un-calibrated SWAT results are discussed.

3 Evaluation criteria

The appropriateness of the models is evaluated based on three criteria. The evaluation is done on so called performance indicators (fit-to-observations) as well as on evaluation of to what extent the hydrological and agricultural processes are realistically represented by means of parameter and mass balance evaluations (fit-to-reality) and to what extent the models are able to tackle the problem (fit-to-purpose).

3.1 Fit-to-observations

A fit-to-observations criteria computes the error between the model outputs and observations for the same variable and is the most typical evaluation criteria to evaluate the performance of hydrological modeling. Moriasi et al. (2007) proposed model evaluation guidelines by assessing the accuracy of simulations compared to measurements. Quantitative statistics Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and ratio of the root mean square error to the standard deviation of measured data (RSR) were recommended for model evaluation in addition to graphical assessment through hydrographs and percent excedence probability curves. The guidelines proposed that model results can be judged as satisfactory if $NSE > 0.5$ and $RSR \leq 0.7$, and if PBIAS $\pm 25\,\%$ for stream flow, PBIAS $\pm 55\,\%$ for sediment for a monthly time step. For hydrological modeling, the NSE is the most frequently used indicator in the assessment of model performance.

3.2 Fit-to-reality

The aim of a conceptual model is to represent the physical processes whereby the observed processes should be well described in the coded model equations while the assessed or calibrated parameters should maintain their physical meaning. The obtained mass balances should be in equilibrium (e.g. inputs minus outputs should be explained by the change in the state variables) and the hydrological mass balance should be in line with the knowledge from the field.

3.2.1 Process representation

The popularity of the SWAT model is largely due to the multi-disciplinary coverage of processes representing the hydrology, soil science, erosion/sediment transport, crop growth, in-stream water quality and the agricultural management. Even though SWAT contains many processes, certain processes may still not be well represented. E.g. Ndomba and van Griensven (2011) indicated in their paper that certain landscape elements, such as wetlands, are not well represented in the SWAT model while they may have a huge impact on the hydrological and nutrient cycle. One may also wonder whether the concepts behind the way the processes are represented in the SWAT model are generally applicable all over the globe. Several of these processes have an empirical background whereby the equations were derived from large data sets in the US. The used Curve Number approach and the USLE soil loss equations are good examples.

As the SWAT model is open-sources, it allowed some users to redefine these processes for specific regional needs. One of these developments that has been applied within the Nile basin is the SWAT-WB model that represents a hydrology that is driven by saturation excess processes as an alternative to the curve number of the SWAT model that represents infiltration-excess processes (Easton et al., 2010).

3.2.2 Parameter value evaluations

After a calibration process, the parameters should maintain their physical meaning when looking at the absolute values as well as how they relate to each other in a relative way (for the distributed parameters). With regard to the parameter values, it is important to see to what extent the default parameters that have been identified for the USA were adjusted towards the African conditions during the calibration process. Special attention should be given to the parameters listed in Table 2 where wrong parameter values may lead to unrealistic simulations. These parameters govern processes that result in a loss of water from the system. Modelers should hence take care that they do not use these parameters to match the water balance for the wrong reasons. "CH_K" describes infiltration in the river bed which occurs in "hanging" rivers where the river bed is higher than the groundwater level. Normally, this water is not lost but should reach (partially) the aquifer below the river. "RCHRG_DP" simulates the water that is going to deep water storage that will not discharge towards the river. Such deep groundwater losses might be significant in small catchments but should not dominate in large river basins. "GW_REVAP" describes the process of capillary rise but the equation rather describes evapotranspiration from the shallow aquifer which is controlled by the potential evapotranspiration. The "revap" water volume is not moving to the soil profile, but is lost from the system and should not become too large. "GWQMN" defines a threshold in the shallow aquifer, and recharge will only occur when the aquifer level goes beyond GWQMN. Since a SWAT model will start with an empty shallow aquifer, it may take several years before the GWQMN level is reached. In that case, the model will build up water in the shallow aquifer whereby the input (rainfall) might not equal to the output (flow + losses).

3.2.3 Mass balance evaluations

Models are always simplifications of reality and the degree that a model is representing the physical processes within a catchment cannot be accurately quantified. However,

there are a couple of checks that can be done with regard to the mass balance. To close the hydrological mass balance, it is expected that the flows and the losses equal the rainfall (P) on the long term, as there should not be a trend in the storage. This may not be true in case that the model is building up storage in the shallow aquifer (ΔGW Storage) or the soil profile (ΔSW Storage), all expressed in mm yr^{-1}.

$$P = \text{ET} + \text{Water Yield} + \Delta\text{SW Storage} + \Delta\text{GW Storage} + \text{Losses}. \quad (1)$$

The water yield (mm yr^{-1}) is computed as the sum of surface flow (Surf$Q$), the laterfal flow (Lat$Q$) and the shallow groundwater flow (GW_Q) diminished by the pond abstractions.

$$\text{Water Yield} = \text{Surf}\,Q + \text{Lat}\,QQW_Q^2 - T\,\text{Loss} - \text{Pond Abstractions}. \quad (2)$$

In addition, certain losses should not be over-estimated in order to compensate for an underestimation of the evapotranspiration. Within the SWAT model, there are a couple of "black holes" where the water might be trapped. An example is the losses to the deep aquifer. Such losses should not be too high for large basins. Also the capillary raise (GW_REVAP) is a loss component that should not be too high. Since this water is not going to the stream flow, it is simply disappearing from the system. Another loss component is the river bed infiltration (controlled by the parameter CH_K), which should not happen in streams where the shallow aquifer is higher than the river bed, and hence producing a groundwater flow. Table 2 summarizes parameters that control the previously mentioned losses.

3.3 Fit-to-purpose

Certain characteristics should be taken into account for certain model applications. For a good land use analysis, it is important, first of all, that all the land uses are included in the model. This means that during a model set-up, one should not use the option in the SWAT interface to exclude some marginal land uses whose percentage within the sub-basin is below a certain threshold. It is difficult to judge whether the models used in the papers properly represent the land covers as no information has been given on how the default land management and crop parameters have been adjusted to local land use practices. A land use modeling study requires more stringent evaluations of the model than a good NSE value. It is important that the hydrological processes (i.e. surface runoff, infiltration, groundwater discharge, evapotranspiration) are properly simulated for the different land uses. It may also be important to properly validate the spatial variability by means of internal observations or other spatial observations such as remote sensing.

4 Review results and discussion

4.1 Fit-to-observations

Table 3 gives an overview of the evaluations including NSE values, comments on reported parameter values and baseflow factors (subsurface flow divided by total runoff) are provided based on what has been reported for the different case studies. In several cases, the performance indexes of the model are not reported. Besides, the cases reviewed have a wide range of spatial representation with the catchment size varying from 1.1 to 184 560 km^2. Some authors calibrated against monthly data whereas others against daily data. Sometimes several calibrations where done with different sources of input data. There are differences in climate zones, where in Ethiopia there is a very distinct rainy season and a very high variability in the stream flows which tends to favor the obtained values of performance indexes. So, a complete fair comparison is not possible. Our evaluation is based on the NSE. When several values were reported, the overall evaluation was based on the highest value (Table 3). Using the classification as proposed by Moriasi et al. (2007) 15 catchment models were classified as very good, 2 as good, 6 as satisfactory 3 as poor and 5 studies did not report any NSE value at all.

4.2 Fit-to-reality evaluations

Under the fit-to-reality evaluation criteria, we assessed hydrological mass balances, the way processes are formulated in different model versions and the parameter values in the case study applications both in the Blue Nile and the Lake Victorian countries.

4.2.1 Mass balance evaluation

Two different SWAT-based modeling concepts are used in the modeling of the Blue Nile basin: the original SWAT model that uses the curve number (SWAT-CN) and a modified version that contains a newly developed water balance concept (SWAT-WB) that uses the topographic index to define the generation processes of the surface runoff (Easton et al., 2010; White et al., 2011).

Both concepts seem to give extremely different results, not only in their spatially distributed outputs, for instance, when looking at the major contributing areas towards the surface runoff (see Fig. 2), but also in the estimation of the base flow factor (% base flow of the total discharge) (Table 3). The reason for the differences are the basic hydrological concepts. The original SWAT-CN model uses the Curve Number concept that is built on the assumption that runoff is generated by means of infiltration excess processes. SWAT-WB simulates the effects of saturated excess phenomenon. Since the regions close to river get saturated with shallow groundwater table or more often saturated than the upland areas, the low regions are generating more runoff compared to the upper regions.

The original model gives a base flow factor between 0.5–0.65 (Setegn et al., 2008) while the modified SWAT-WB computes a base flow factor between 0.9 and 0.95 (Easton et al., 2010). The base flow factors that are obtained after using base flow filter programs on observed flow time series range between 0.49 and 0.6 and are more in line with the results of the original SWAT model (see Fig. 2).

Few papers describe the hydrological mass balance of the simulated results. Mango et al. (2011) provide the different hydrological components that are computed by SWAT. Out of the mass balance components, it could be derived that the groundwater flow (GQ_Q) should be 480 mm yr^{-1} instead of the reported 48 mm yr^{-1}. Taken this into account, it appears that the model has unexplained losses of 110 mm yr^{-1} for Nyangores-RG while 32 mm yr^{-1} for the Nyangores-RFE gauge. These gaps are likely introduced by errors in rainfall and cause that in the model there is an increase of storage of the same amounts. This is most likely happening in the shallow aquifer, since the parameter "gwqmn" has been reported as being incorporated into the calibration process what may result in simulations where the shallow aquifer volume might be much larger at the end of the simulation compared to the beginning of the simulation (up to 1500 mm). Mulungu and Munishi (2007) reported a very low water yield of 74 mm for the Simiyu catchment. The low water yield and the very low groundwater contributions are probably obtained after simulating very large losses to the deep aquifer through the parameter "RCHRG_DEEP = 0.9", which means that 90 % of the water that is recharging to the groundwater is lost to the deep aquifer.

4.2.2 Parameter values

In total, values for 29 parameters have been reported, of which 19 affecting the hydrological processes while the remaining are related to the sediment transport processes. Curve numbers have been reported covering their full range from 34.5 to 98, with an average value of 61.8. We give more focus on the parameters that are reported in Table 1 as these parameters may lead to errors in the mass balance. The very high values of the "RCHR_DP" parameter in the Tanzanian case studies (Ndomba et al., 2008; Mulungu et al., 2007), as well as in the Blue Nile basin (Betrie et al., 2011) seem to be unrealistic as losses of water to the deeper groundwater layers are not expected to be significant in large basins. However, Ndomba et al. (2008) attributed it to the presence of a huge groundwater reservoir in the basin. Also Notter et al. (2010) obtained a high "RCHR_DEEP" parameter value of 0.75 during the calibration of the Pangani river

basin. Setegn et al. (2010) reported an out of range "GW_REVAP" value which is likely to be a typing error as the value is out of the physical range.

4.3 Fit-to-purpose

Many models have been reported for different purposes. Several models have been used to run scenarios such as climate change and land use change. Even though all these cases used are tributaries of Lake Victoria that have similar climate zones and vegetation, the models seem to give different results. Here, we are focusing on the models that are used for land use change studies, since such studies would need a very good representation of the spatial variability with special attention to the land use/land cover related inputs and processes. There are currently only 2 papers that used SWAT for land use change analysis, namely on the upper Mara (Mango et al., 2011) and the Nzoia (Githui et al., 2008). None of the studies used internal flow gauges. The papers did not describe neither crop input parameters, nor provided any outputs on the computed vegetation/crop variables. One may wonder the correctness of the forest simulations in the model that predicts only marginal effect of deforestation on the evapotranspiration component (Mango et al., 2011). The study in the Nzoia shows a stronger difference in evapotranspiration between scenarios of a continuation of deforestation versus where degraded forest would be replanted. Still, forest land seems to have less evapotranspiration compared to grassland and shrubland. So far, none of the reviewed papers discussed details on the simulations of crops and/or evergreen forests.

5 Conclusions and recommendations

Data availability is a general problem within the Nile countries and the lack of data is often mentioned as a problem. Nevertheless, most of the models seem to perform quite well in representing the temporal dynamics within the catchment. On the basis of performance indicators, the SWAT models were in general produced satisfying or good results. However, little confidence can be given to the degree that the models are able to represent the processes in a spatially distributed way and hence to properly represent the spatial heterogeneity. The models tend to lack a method of validation for a spatially distributed representation of the processes. Very few of the studies included some internal calibration points or other distributed data (e.g. remote sensing data, tracer data, groundwater data etc.) to check the distributed predictions, even though they might exist. None of the studies reported the used crop parameters or how the land covers in the basin are represented in the SWAT model. Neither did any of the papers report the crop related outputs such as leaf area index, biomass or crop yields. A proper simulation of the land cover is important for obtaining correct evapotranspiration, runoff generation and erosion computations. It is therefore recommended to try to evaluate the representativeness of the distributed processes and parameters, especially when land use studies are envisaged. A validation of the crop processes could be achieved through comparison with remote sensing data. For that reason, the models may not always be adequate for land use analysis studies.

When different studies in the same or similar catchments are compared, the differences in the results are often striking. There are different responses to climate change and land use change in the Lake Victoria basins whereas one would expect that they respond similarly. SWAT-CN and SWAT-WB versions give very different results when the hydrological responses are plotted spatially and they also show very different base flow factors.

In several papers, the reported hydrological mass balances encompassed several losses that might not be justified or some papers reported parameter values that might not be realistic. More worrying, however, is the fact that many papers lack this type of information. For that reason, it is difficult to give an overall positive evaluation to most of the reported SWAT models.

The following recommendations could lead to better model practices in the Nile basin and beyond:

1. Improving the spatial variability computations by the inclusion of internal calibration points.

2. Crop and vegetation database for Africa should be built and shared among the African SWAT user community.

3. Crop outputs should be evaluated, especially when land use/land cover studies are aimed for. The information obtained from remote sensing should be used in this context.

4. The hydrological water balance, as well as parameter values should be checked and compared with knowledge from the field and with field observations.

5. Special attention should be given to the computed hydrological losses in the catchment. They should not be used to make the model fit and to account for incorrect input variables.

6. More attention should be given to the dominating hydrological processes and their representativeness in the SWAT model. A catchment might not have infiltration excess or saturation excess exclusively, but these may happen at the same place at different moments in time, or, at the same time both processes might happening depending on the position of a place within the landscape. It is also important to better represent spatial dynamics of the subsurface storage (often depending on the position in the hill slope) and the routing of the sub-surface flow from one landscape element to the other or from one sub-basin to the other.

7. An overall recommendation is that the journal papers should be more complete in reporting model performances, computed mass balances and the calibrated parameter values in order to allow for a better evaluation as well to allow for a reproduction of the studies by others.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Acknowledgements. We are very thankful to the UPARF ACCION project support, the UNESCO/FRIEND Nile project and the EU FP7 AFROMAISON project that financially supported the review of SWAT modeling in the Nile basin. Also thanks to Stefan Uhlenbrook, Pieter van der Zaag and Getnet Dubale Betrie for their comments and feedbacks on the manuscript.

References

Betrie, G. D., Mohamed, Y. A., van Griensven, A., and Srinivasan, R.: Sediment management modelling in the Blue Nile Basin using SWAT model, Hydrol. Earth Syst. Sci., 15, 807–818, doi:10.5194/hess-15-807-2011, 2011.

Dargahi, B. and Setegn, S. G.: Combined 3D hydrodynamic and watershed modeling of Lake Tana, Ethiopia, J. Hydrol., 398, 44–64, 2011.

Easton, Z. M., Fuka, D. R., White, E. D., Collick, A. S., Biruk Ashagre, B., McCartney, M., Awulachew, S. B., Ahmed, A. A., and Steenhuis, T. S.: A multi basin SWAT model analysis of runoff and sedimentation in the Blue Nile, Ethiopia, Hydrol. Earth Syst. Sci., 14, 1827–1841, doi:10.5194/hess-14-1827-2010, 2010.

Gassman, P. W., Reyes, M. R., Green, C. H., and Arnold, J. G.: The Soil and Water Assessment Tool: historical development, applications, and future research directions, T. ASABE, 50, 1211–1250, 2007.

Gassman, P. W., Arnold, J. G., Srinivasan, R., and Reyes, M.: The worldwide use of the SWAT Model: Technological drivers, networking impacts, and simulation trends, in: Proceedings of the Watershed Technology Conference, 21–24 February, American Society of Agricultural and Biological Engineers, Earth University, Costa Rica, St. Jospeh, MI, ASABE, 2010.

Gessese, A. and Yonas, M.: Prediction of Sediment Inflow to Legedadi Reservoir Using SWAT Watershed and CCHEID Sediment Transport Models, Nile Basin Water Engr. Sci. Magazine, 1, 65–74, 2008.

Githui, F., Gitan, W., Mutua, F., and Bauwens, W.: Climate change impact on SWAT simulated streamflow in western Kenya, Int. J. Climatol., 29, 1823–1834, doi:10.1002/joc.1828, 2009a.

Githui, F., Mutua, F., and Bauwens, W.: Estimating the impacts of land-cover change on runoff using the soil and water assessment tool (SWAT): Case study of Nzoia catchment, Kenya, Hydrolog. Sci. J., 54, 899–908, 2009b.

Jayakrishnan, R., Srinivasan, R., Santhi, C., and Arnold, J. G.: Advances in the application of the SWAT model for water resources management, Hydrol. Process., 19, 749–762, 2005.

Kingston, D. G. and Taylor, R. G.: Sources of uncertainty in climate change impacts on river discharge and groundwater in a headwater catchment of the Upper Nile Basin, Uganda, Hydrol. Earth Syst. Sci., 14, 1297–1308, doi:10.5194/hess-14-1297-2010, 2010.

Mekonnen, M. A., Worman, A., Dargahi, B., and Gebeyehu, A.: Hydrological modelling of Ethiopian catchments using limited data, Hydrol. Process., 23, 3401–3408, doi:10.1002/hyp.7470, 2009.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, T. ASABE, 50, 885–900, 2007.

Mulungu, D. M. M. and Munishi, S. E.: Simiyu river catchment parameterization using SWAT model, J. Phys. Chem. Earth A/B/C, 32, 1032–1039, 2007.

Muvundja, F. A., Pasche, N., Bugenyi, F. W. B., Isumbisho, M., Müller, B., Namugize, J., Rinta, P., Schmid, M., Stierli, R., and Wüest, A.: Balancing nutrient inputs to Lake Kivu, J. Great Lake. Res., 35, 406–418, 2009.

Ndomba, P. M. and Birhanu, B. Z.: Problems and Prospects of SWAT Model Applications in NILOTIC Catchments: A Review, Nile Basin Water Engr. Sci. Magazine, 1, 41–52, 2008.

Ndomba, P. M. and van Griensven, A.: Suitability of SWAT model in sediment yields modeling in the Eastern Africa, in: Chapter 13: Advances in Data, Methods, Models and Their Applications in Geosciences, edited by: Chen, D., ISBN 978-953-307-737-6, INTECH open access publisher, available at: www.intechweb.org, 2011.

Ndomba, P., Mtalo, F., and Killingtveit, A.: SWAT model application in a data scarce tropical complex catchment in Tanzania, Phys. Chem. Earth, 33, 1–19, doi:10.1016/j.pce.2008.06.013, 2008.

Setegn, S. G., Srinivasan, R., and Dargahi, B.: Hydrological Modelling in the Lake Tana Basin, Ethiopia Using SWAT Model, Open Hydrol. J., 2, 49–62, 2009a.

Setegn, S. G., Srinivasan, R., Dargahi, B., and Melesse, A. M.: Spatial delineation of soil erosion vulnerability in the Lake Tana Basin, Ethiopia, Hydrol. Process., 23, 3738–3750, doi:10.1002/hyp.7476, 2009b.

Setegn, S. G., Darfahi, B., Srinivasan, R., and Melesse, A. M.: Modeling of sediment yield from Anjeni-gauged watershed, Ethiopia using SWAT model, J. Am. Water Resour. Assoc., 46, 514–526, doi:10.1111/j.1752-1688.2010.00431.x, 2010.

Swallow, B. M., Sang, J. K., Nyabenge, M., Bundotich, D. K., Duraiappah, A. K., and Yatich, T. B.: Tradeoffs, synergies and traps among ecosystem services in the Lake Victoria basin of East Africa, Environ. Sci. Poll., 12, 504–519, doi:10.1016/j.envsci.2008.11.003, 2009.

Tibebe, D. and Bewket, W., Surface runoff and soil erosion estimation using the SWAT model in the Keleta watershed, Ethiopia, Land Degrad. Develop., 22, online first: doi:10.1002/ldr.1034, 2010.

White, E. D., Easton, Z. M., Fuka, D. R., Collick, A. S., Adgo, E., McCartney, M., Awulachew, S. B., Selassie, Y. G., and Steenhuis, T. S.: Development and application of a physically based landscape water balance in the SWAT model, Hydrol. Process., 25, 915–925, doi:10.1002/hyp.7876, 2010.

Table 1. Overview of papers describing SWAT applications in the tropical highlands of the Nile countries.

	Hydrology/ water balance	Calibration uncertainty	Erosion	Land manage-ment	Land use change	Climate change	SWAT development	Data	Water quality
Jayakrishnan (2005)		X						X	
Mulungu and Munishi (2007)	X	X							
Ndomba and Birhanu (2008)		X						X	
Ndomba et al. (2008)		X					X		
Gessese and Yonas (2008)			X						
Setegn et al. (2008)	X	X							
Setegn et al. (2009)			X						
Mekonnen et al. (2009)	X	X							
Githui et al. (2009a)						X			
Githui et al. (2009b)					X				
Swallow et al. (2009)			X	X					
Muvundja et al. (2009)		X							X
Setegn et al. (2010)			X						
Kingston and Taylor (2010)		X				X			
Tibebe and Bewket (2010)	X		X						
Easton et al. (2010)	X		X				X		
White et al. (2011)	X						X		
Dargahi and Setegn (2011)			X						
Betrie et al. (2011)			X	X					
Bitew and Gebremichael (2011)	X							X	
Mango et al. (2011)	X		X		X	X			
Notter et al. (2011)	X						X		

Table 2. Parameters that are controlling losses (in addition to evapotransiration).

Parameters	Description	Potential concerns	Evaluation	Reported values	Mean value	Number of reported values
CH_K	Effective hydraulic conductivity [mm h^{-1}]	The parameter should only get a value higher than 0 for channels where the groundwater is below river bed.	TRANSMISSION LOSSES (printed in output.std or output.rch) should not be too high, and should be 0 in rivers with all year round baseflow contributions.	0.7 to 150	17.7	16
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur [mm].	SWAT will build up groundwater over the years in case the parameter is high, and the initial value is zero.	SA_ST (mm) = storage in the shallow aquifer printed in output.hru.	0 to 1500	307.8	5
RCHRG_DP	Deep aquifer percolation fraction	These losses should be small in most of the catchment, definitely at a larger scale.	DEEP AQ RECHARGE output in output.std.	0 to 1.1	0.6	5
GW_REVAP	Groundwater revap coefficient (0–1)	These losses should not be too high, certainly not in humid and/or cold regions and for deep aquifers.	REVAP (SHAL AQ => SOIL/PLANTS) output in the output.std file.	0.0	20.0	2.6

Table 3. Overview of evaluations for parameters and performance index (NSE). Black refers to "missing data", grey to "incomplete data", green to "OK", yellow "slightly worrying", red "worrying". The underscored papers are using SWAT-WB version. "D" refers to daily performance, "M" to Monthly performance, "r2D" to the daily correlation factor.

Ethiopian cases	River	Area (km^2)	Performance for flow	Baseflow factor	Evaluation of parameter values
Easton et al. (2010)	Anjeni	1.134	0.84 (D)	0.91	incomplete
Setegn et al. (2010)	Anjeni	2.134	0.89 (M)		
Easton et al. (2010)	Gumera	1286	0.81 (D)	0.93	incomplete
Setegn et al. (2008)	Gumera	1286	0.61 (D)	0.64	
White et al. (2011)	Gumera	1296	0.77 (D)		
White et al. (2011)	Gumera	1296	0.64 (D)		
Mekonnen et al. (2009)	Gumera	1286	0.84 (r2D)		Reasonable
Easton et al. (2010)	Ribb	1295	0.77 (D)	0.92	incomplete reporting
Setegn et al. (2008)	Ribb	1295	0.55 (D)		
Easton et al. (2010)	North Marawi	1658	0.75 (D)	0.94	incomplete reporting
Easton et al. (2010)	Jemma	5429	0.92 (M)	0.90	incomplete reporting
Easton et al. (2010)	Angar	4674	0.79 (M)	0.91	incomplete reporting
Easton et al. (2010)	Blue Nile@Kessie	65 385	0.53 (M)	0.93	incomplete reporting
Easton et al. (2010)	Abay Ethiopian@El Diem	174 000	0.87 (D)	0.95	incomplete reporting
Betrie et al. (2011)	Abay Ethiopian@El Diem	184 560	0.68(D), 0.82(M)		high RCHR_DP
Easton et al. (2010)	Megesh			0.60	incomplete reporting
Setegn et al. (2009)	Lake Tana	15 096		0.50	rather high gwqmn
Setegn et al. (2008)	Gilgel Abay		0.73 (D)	0.54	high GW_REVAP value
Mekonnen et al. (2009)	Gilgel Abay		0.84 (r2D)		high CN
Bitew and Gebremichael (2011)	Gilgel Abay	299			High CH_K and high gwqmn
Bitew and Gebremichael (2011)	Koga	1656			
Tibede et al. (2010)	Keleta	1060	0.789 (M)		
Lake Victoria region					
Githui et al. (2009)	Nzoia	12 709	0.71 (D) 0.76 (M)	0.77	
Mango et al. (2011)	Mara-Nyaragores RG	700	−0.53 (M)	0.82	incomplete reporting
Mango et al. (2011)	Mara-Nyaragores RFE	700	0.43 (M)	0.90	incomplete reporting
Mulungu and Munishi (2007)	Simiyu-Ndagalu	5320	0.1373 (M)	0.15	worrying
Jayakrishnan et al. (2004)	Sondu	3050	−0.72 (D)		
Kingston and Taylor (2010)	Mitano	2098	0.06 (M), −0.09 (D)	0.99	
Swallow et al. (2009)	Nyando	4000			
	Yala	3000			
Ndomba et al. (2008)	Pangani	7280	0.54 (D), −0.65 (M)		high RCHR_DP, high gwq_mn, high ch_k
Notter et al. (2010)	Pangani	43 000			high RCHR_DP

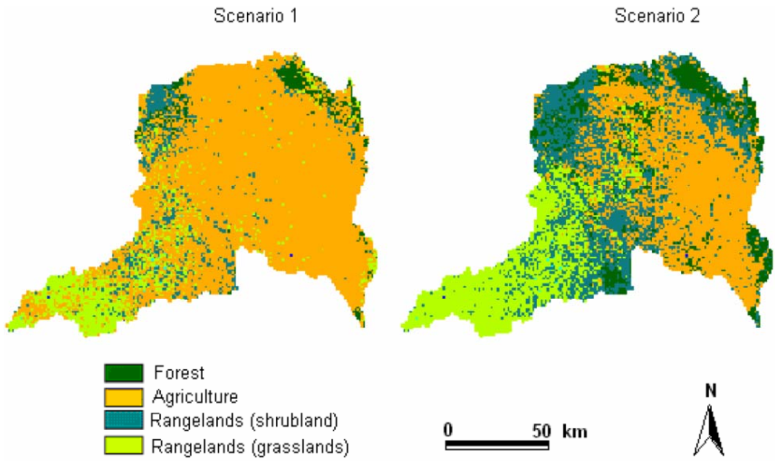

Fig. 1. Land cover scenarios (Githui et al., 2009).

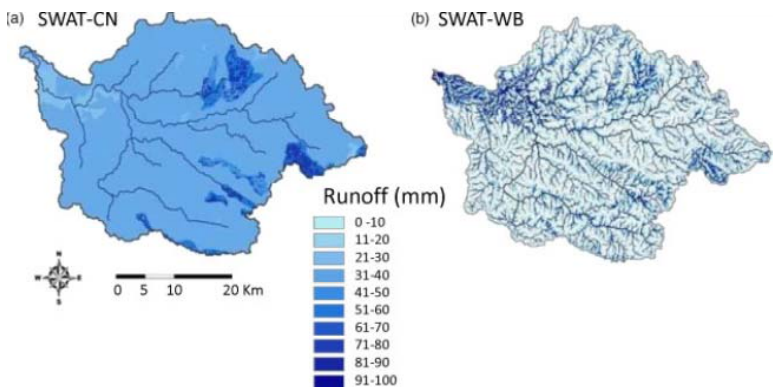

Fig. 2. Spatial distribution of surface runoff in Gumera modeled with **(a)** SWAT-CN and **(b)** SWAT-WB (White et al., 2010).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |