Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

Improving SWAT model performance in the Upper Blue Nile River Basin using meteorological data integration and catchment scaling

Erwin Isaac Polanco[1], Amr Fleifle[1,2], Ralf Ludwig[3], Markus Disse[1]

[1]Chair of Hydrology and River Basin Management, Faculty of Civil, Geo and Environmental Engineering, Technische Universität München, Arcisstrasse 21, 80333, Munich, Germany.
[2]Irrigation Engineering and Hydraulics Department, Faculty of Engineering, Alexandria University, El-Horia St., 21544, Alexandria, Egypt.
[3]Chair of Geography and Geographic Remote Sensing, Faculty of Geography, Ludwig-Maximilians-Universität München, Luisenstrasse 37, 80333, Munich, Germany.

Correspondece to: E. I. Polanco (erwinisaac.polanco@tum.de)

Abstract. The Upper Blue Nile River Basin is confronted by land degradation problems, insufficient agricultural production, and limited number of developed energy sources. Process-based hydrological models provide useful tools to better understand such complex systems and improve water resources and land management practices. In this study, SWAT was used to model the hydrological processes in the Upper Blue Nile River Basin. The calibration was done in such a way that the parameterization had a realistic representation of the interaction of land cover and soils properties. Comparisons between a Climate Forecast System Reanalysis (CFSR) and a ground weather dataset were done under two subbasin discretizations (30 and 87 subbasins) to create an integrated dataset to improve the spatial and temporal limitations of both datsets. A SWAT Error Index (SEI) was also proposed to compare the reliability of the models under different discretization levels and weather datasets. This index offers an assessment of the model quality based on precipitation and evapotranspiration. SEI demonstrates to be a reliable and useful method to measure the level of error of SWAT and develop better models. The results showed the discrepancies of using different weather datasets with different levels of subbasin discretization. Datasets under 30 subbasins achieved Nash-Sutcliff (NS) values of 0.15, 0.68 and 0.82; while models under 87 subbasins achieved values of 0.05, 0.61, and 0.84, for the CFSR, ground and integrated datasets, respectively. Based on the parameterization, the integrated dataset provided more reliable results and a more realistic representation of the land use and soil conditions of this region.

Keywords. SWAT, subbasins discretization, CFSR, Integrated dataset, SWAT Error Index (SEI).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

1 Introduction

Water resources in the Upper Blue Nile River Basin (UBNRB) are not being managed adequately; land use changes, fast population growth, land erosion and deforestation are some of the causes currently affecting the watershed. Therefore, researchers need to understand and improve the watershed management to provide good land use management practices and mitigate the alarmingly erosion problems. Physically based, distributed models have provided a very efficient alternative for watershed researchers for analyzing the impact of land management practices on soil degradation, agriculture, water allocation and chemical yields (Shimelis, 2008). Due to its versatility and applicability to complex watersheds, researchers have identified the Soil and Water Assessment Tool (SWAT) as one of the most intricate, consistent and computationally efficient models (Neitsch *et* al., 2009; Gassman et al., 2007). Recent studies are a prove that SWAT has become internationally and interdisciplinary accepted for modelling large and small watersheds (Malunjkar *et* al., 2015; Me *et* al., 2015; Emam *et* al., 2016; Wang *et* al., 2016). SWAT provides a wide range of parameters to work with, allowing users to analyze several hydrological processes. It also has the advantage to have been developed to analyze the interaction of several hydrological parameters and the impact of land management practices specifically for large and complex basins, thus a good model to be applied in the UBNRB. However, due to the lack of a unifying theory to accurately model the interaction of the hydrological processes, complex hydrological models suffer from overparameterization and high predictive uncertainty (Sivapalan, 2006). Therefore, it is difficult to simulate the complex interactions of hydrological processes and weather conditions of watersheds without uncertainties.

Among all the input parameters, the meteorological data has the most significant impact on the water balance of a watershed. However, a common problem to set up hydrological models of the UBNRB are related to data limitations. In developing countries the distribution of meteorological stations is irregular and dispersed (Worqlul et al., 2014). Other weather data problems are related to measuring gauges; many weather data parameters contain missing data periods, and in several cases erroneous measurements are also possible. Thus, many models are often set up based on limited and incomplete data, which may lead to less reliable models. Therefore, the first objective of this study has been the comparison of different weather datasets at large scale and under different subbasin discretization levels. This comparison provided a better understanding of the effects of different subbasin discretizations on the total water balance of a watershed. It also helped to identify the temporal and spatial constraints of both datasets. Roth and Lemann (2016) performed a comparison between CFSR and conventional data in small catchments in the Ethiopian highland, where they showed that the CFSR data provided unreliable results. Furthermore, this study proposes an integration of a CFSR and conventional weather data that can more accurately represent the magnitudes and distribution of the different weather variables in the Upper Blue Nile River Basin.

After a hydrological model has been setup, a critical point to determine its quality is the water balance. Therefore, in addition to graphical assessments, other statistical indicators as Nash-Sutcliffe coefficient

(NS), percent bias (PBIAS), and ratio of the root-mean-square error (RSR) to the standard deviation of measured data were proposed by Moriasi *et* al. (2007). Based on these commonly used statistical indicators most of the SWAT models provide very good results for discharge values at the outlet of a basin (Griensven *et* al., 2012). However, problems concerning the evapotranspiration and the water balance of the catchment are not discussed in details. Several published documents could even report unrealistic parameter values (Griensven *et* al., 2012). Therefore, the second objective of this study has been to propose an index, the SWAT Error Index (SEI), to quantify the level of error of a hydrological model. The SWAT Error Index (SEI) uses flexible weighting values for the relative Root Mean Square Error (rRMSE) obtained from measured flow discharge data and satellite evapotranspiration data. Hence, SEI shows to be an useful method to develop models that can provide a better representation of the water balance of a watershed.

Two models (SWAT 30 and SWAT 87) were created using different subbasin discretizations, 30 and 87 subbasins (Figure 1). The time frame of the models was from 1980 to 2004, using a 4 years warm up period (1980-1983), a 14 years calibration period (1984-1997) and a 7 years validation period (1998-2004).

2 Materials and methods

2.1 Study site

The Upper Blue Nile River Basin (UBNRB), also known as Abbay basin, is located in the northwestern highlands of Ethiopia, approximately between Latitude 7 40'N and 12 51'N, and Longitude 34 25'E and 39 49' E, with elevations raging between 483 and 4248 m.a.s.l. The total area of the UBNRB is approximately 199,812 km^2, including two subbasins shared with Sudan in the northern region. The climate in the UBNRB fluctuates from humid to semi-arid and it is mainly dominated by latitude and altitude, with average temperatures ranging from 13°C in the south eastern regions to 26°C in the south western. The lowest rainfall data detected during the current research period (1980-2004) corresponds to the eastern region, for the subbasins of Beshelo, North Gojam, South Gojam, Welaka, Jemma, Muger, Guder and Fincha; where the precipitation drops below 1000 mm per year. While the highest precipitation ranges belong to the western region: Didessa, Wenbera, Anger, Dabus and Beles; with precipitations above 1900 mm per year. The topographic disparity and variations in altitude of the UBNRB have a great impact in the weather, soil and vegetation conditions. Consequently, rainy seasons are very variable in this watershed, for instance the total discharge peaks at the Eldiem gauging station can reach 7,000 m^3/s; and dry seasons can go as low as 100 m^3/s (refer to figures 4 and 5). Soils in the UBNRB are mainly dominated by ten types: Eutric Nitosols, Eutric Cambisols, Humic Fluvisols, Cambic Arenosols, Chromic Vertisols, Dystric Cambisols, Eutric Fluvisols, Eutric Regosols, Orthic Acrisols and Pellic Versitols (FAO, 2015).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

2.2 Datasets

The final quality of a hydrological model is highly dependent on the quality of the input data, which in this case has been assessed and processed before its use. Two types of datasets were used in this study: the input data that was used to set up the model, and the evaluation data that was used to calibrate, validate and determine the quality of the models. SWAT requires four types of input data: a Digital Elevation Model (DEM), a land use map, a soil map and a weather dataset.

A Shuttle Radar Topographic Mission Digital Elevation Model (SRTM DEM) from the Consultative Group on International Agricultural Research-Consortium for Spatial Information (CGIAR-CSI) was used to setup the model. This DEM has a resolution of 90 meters, and was used to perform an automatic watershed delineation of the UBNRB, where the flow direction, flow accumulation and streams network were automatically determined by SWAT.

The second input dataset was a land use map, which was obtained from the GIS Portal of the International Livestock Research Institute (ILRI), and corresponds to the year 2004. (http://data.ilri.org/geoportal/catalog/main/home.page).

The soil map used for these models was developed by the Food and Agriculture Organization of the United Nations (FAO-UNESCO). This world soils map was prepared by FAO and UNESCO at 1:5 000 000 scale (http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/). The information provided by this map was used in combination with the Harmonized World Soil Database v1.2, a database that combines existing regional and national soil information (http://www.fao.org/soils-portal/soil-survey/ soil-maps-and-databases/harmonized-world-soil-database-v12/en/).

The last input dataset was the meteorological information. Two weather datasets from different sources were used to setup the models. The first weather dataset was collected from the National Meteorology Agency of Ethiopia (NMA). The data used for these models correspond to 42 stations distributed in the Upper Blue Nile River Basin (Figure 1). However, only 15 of these stations are capable of measuring all 5 parameters needed to set up SWAT: rainfall, temperature, relative humidity, solar radiation and wind speed. Moreover, few of these 15 station have available complete and continuous data for the entire period under study (1980-2004). For instance, the collected data for solar radiation was limited to 2 stations only, wind speed was available for 4 stations; only maximum temperature was available for 4 stations, relative humidity was available for 3 stations, and precipitation was available for all 42 stations. Additionally, the quality of this observed data is somehow questionable. Many meteorological stations are more than 10 years old, and their constant technical failure due to the lack of continuous expert maintenance also questions the quality of the data. Large part of the available ground data has been collected from old stations that could have in many cases malfunctioning, defected and outdated devices. The second weather dataset was the Climate Forecast System Reanalysis (Figure 1), a dataset

that has been produced by the National Centers for Environmental Prediction (NCEP) (http://globalweather.tamu. edu/). CFSR data brings several uncertainties due to its multiple spatial and temporal interpolations (Dile & Sriniavasan, 2014). It was generated using different assimilation techniques that include satellite radiances, advanced coupled atmospheric, oceanic and land surface modelling components. The global atmosphere resolution of CFSR data is approximately 38km. These atmospheric, oceanic and land surface output products are available at a 0.5°x0.5° latitude and longitude resolution. Both weather datasets used for these models correspond to the period from 1980 to 2004.

For the analysis of the quality of the SWAT models, monthly flow discharge data and evapotranspiration data were used. The flow discharge data was obtained from the Ministry of Water and Energy of Ethiopia and corresponds to the gauging stations at Kessie and Eldiem at the main stream of the UBNRB (Figure 1). For the evapotranspiration analysis, MODIS data for the Upper Blue Nile River Basin was obtained from the MOD16 Global Terrestrial Evapotranspiration Project (http://www.ntsg.umt.edu/project/mod16). The global evapotranspiration data from MOD16 are regular 1 km^2 land surfaces datasets for the 109.03 million km^2 of vegetated area in the whole globe at different time interval: 8 days, monthly and annual, from which monthly data generated specifically for the Nile basin was used.

2.3 Water balance and evapotranspiration processes in SWAT

Water balance in watersheds is one of the most important factors used to determine if a model is good enough for any particular application. Hence, analyses of the processes involved in the estimation of the water balance of a watershed (evapotranspiration, runoff and groundwater) can provide more details about the hydrological behavior of a watershed and can be used to understand the interaction of main hydrological processes (Zhang *et* al., 1999). For the input data processing and hydrological estimation SWAT is using two levels of discretization, i.e., subbasins and Hydrologic Response Units (HRUs). HRUs are contained in the subbasins and are defined based on the land use map, soil map and slope classes. HRUs allow the model to reflect differences in evapotranspiration and other hydrologic conditions for each crop and soil type. The water balance in SWAT is calculated for each HRU using the following formula (Neitsch et al., 2009):

$$SW_t = SW_0 + \sum_{i=1}^{t}\left(R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}\right)$$

Equation (1)

where SW_t is the final soil water content (mm), SW_0 is the initial soil water content on day i (mm), R_{day} is the amount of rainfall on day i (mm), Q_{surf} is the amount of surface runoff on day i (mm), E_a is the amount of evapotranspiration on day i (mm), W_{seep} is the amount of water entering the vadose zone from the soil profile on day i (mm), and Q_{gw} is the amount of return flow on day i (mm).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

SWAT can estimate the evapotranspiration using several methods, from which Hargreaves and Penman-Monteith methods were compared in this study (Figures 8 and 9). The Hargreaves method calculates the potential evapotranspiration using minimum and maximum daily temperature as input data (Hargreaves and Samani, 1982). This method was chosen as a better option for the UBNRB due to the data scarcity of the meteorological stations in the basin. Hargreaves equation can be used with the sole input of temperature data, while Penman-Monteith requires more data, for instance wind speed, solar radiation and relative humidity. Hargreaves method has been recommended for computing potential evaporation in cases when only the maximum and minimum temperatures are available (Allen et al., 1998). A study from Tekleab and Uhlenbrook (2011) was also able to successfully use the Hargreaves equation to calculate the potential evaporation in the Upper Blue Nile River Basin. Several improvements were made to the original equation since 1975 (Hargreaves & Samani, 1982 & 1985). The final form of the Hargreaves equation used in SWAT and published in 1985 (Hargreaves *et* al., 1985) is as follows (Neitsch *et* al., 2009):

$$\lambda E_0 = 0.0023 * H_0 * (T_{mx} - T_{mn})^{0.5} * (\bar{T}_{av} + 17.8)$$

Equation (2)

where λ is the latent heat of vaporization (MJ kg^{-1}), E_0 is the potential evapotranspiration (mm d^{-1}), H_0 is the extraterrestrial radiation (MJ m^{-2}d^{-1}), T_{mx} and T_{mn} are the maximum and minimum air temperature for a given day (°C), respectively, and T_{av} is the mean air temperature for a given day.

The evaporative water demand after the evapotranspiration of free water in the canopy has occurred is calculated with the following formula:

$$E'_0 = E_0 - E_{can}$$

Equation (3)

where E'_0 is the potential evapotranspiration adjusted for evaporation of free water in the canopy (mm H_2O), E_0 is the potential evapotranspiration on a given day (mm H_2O), and E_{can} is the amount of evaporation from free water in the canopy on a given day (mm H_2O).

Following the potential evapotranspiration, the actual evapotranspiration must be calculated. Initially, SWAT calculates the evaporated water intercepted by the canopy, then, maximum transpiration and soil evaporation are calculated. Evaporation from canopy is very significant in forested areas and in several cases can be higher than transpiration. Transpiration for the Hargreaves equation is calculated as (Neitsch *et* al., 2009):

$$E_t = \frac{E'_0 \cdot LAI}{3.0}$$

Equation (4)

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

where E_t is the maximum transpiration on a given day (mm H_2O), E_0' is the potential evapotranspiration adjusted for evaporation of free water in the canopy (mm H_2O), and *LAI* is the leaf area index.

Evaporation from the soil on a given day is calculated with following equation (Neitsch *et* al., 2009):

$$E_s = E_0' \cdot cov_{sol}$$

Equation (5)

where E_s is the maximum soil evaporation on a given day (mm H_2O), E_0' is the potential evapotranspiration adjusted for evaporation of free water in the canopy (mm H_2O), and cov_{sol} is the soil cover index.

The maximum soil evaporation is reduced during periods of high plant water use with the relationship (Neitsch *et* al., 2009):

$$E'_s = \min\left|E'_s, \frac{E_s * E'_0}{E_s + E_t}\right|$$

Equation (6)

where E'_s is the maximum sublimation/soil evaporation adjusted for plant water use on a given day (mm H_2O), E_s is the maximum sublimation/soil evaporation o a given day (mm H_2O), E'_0 is the potential evapotranspiratin adjusted for evaporation of free water in the canopy (mm H_2O), and E_t is the transpiration of a given day (mm H_2O).

2.4 Weather data processing and integration

A common problem in the Upper Blue Nile River Basin is the data scarcity. If input data is used without the respective analyses, models will provide less reliable results. And even small errors in temperature or precipitation can result in considerable inaccuracies and impacts on the models results (Maraun et al., 2010). Therefore, an analysis of the quality of the input data was performed based on comparison graphs between both ground and CFSR datasets (Figure 2 and 3). Tekleab and Uhlenbrook (2011) also checked the data quality of stream flow data in the Upper Blue Nile River Basin based on comparisons graphs and additionally a double mass analysis. In this study the consistency of the time series on monthly basis in terms of magnitude and spatial distribution of the five input variables of SWAT were analyzed through comparison graphs (Figure 3) to determine the deficiencies of the two datasets and to form an integrated dataset.

The ground dataset obtained for the National Meteorology Agency of Ethiopia (NMA) correspond to 42 stations in the basin (Figure 1), where most of the meteorological stations were located in the eastern part of the watershed. Additionally, the data obtained from these stations also had several

months of missing data, leading to temporal uncertainties. The CFSR dataset is evenly distributed at 38 km resolution, with over 100 stations available for the UBNRB, and is temporally continuous. However, after checking the data quality through a comparison between graphs and maps obtained from ground and CFSR data (Figure 2 and 3), it was noticed that not all the weather variables from CFSR are reliable. The precipitation distribution appeared to be underestimated in the eastern region of the basin and overestimated in the western region (Figure 2). Two stations in the eastern part, Alemketema and Adet (Figure 3), showed the underestimation of the CFSR rainfall at the eastern region, and Ayehu (Figure 3) showed the overestimation of the CFSR rainfall in the western region. For this reason, additional CFSR rainfall stations were not used in this study. However, the graphical comparisons of the few available stations for relative humidity, temperature and solar radiation showed a good agreement between ground and CFSR values, their seasonal behavior and the magnitudes of the values are very similar (Figure 3). For instance, the values for relative humidity for Debre Tabor and Aykel with both datasets show very similar values (Figure 3). The comparisons of maximum temperature for Aykel and Bahir Dar from both datasets are also within a good agreement to each other, although CFSR data show a slight overestimation at Aykel station and a slight underestimation at Bahir Dar (Figure 3). The solar radiation comparison at Bahir Dar and Debre Tabor also show similar values in both cases. The exception was the wind speed data, which in both cases at Adet and Ayehu was overestimated by the CFSR data. Therefore, these two datasets were integrated with the objective of overcoming their spatial and temporal limitations. Tekleab and Uhlenbrook (2011) filled in missing stream flow data of the UBNRB using regression analysis, which is also be a good approach to fill in missing meteorological values. However in this study, the missing values refer to complete time series of a specific station and variable. Therefore, these variables were filled in using the values from the CFSR dataset, i.e., the rainfall of the 42 ground stations was used in combination with the weather variables (relative humidity, temperature and solar radiation values) of their respective neighboring CFSR station (Figure 1). The missing weather variables for each of the 42 ground stations, were completed by using the weather data of its nearest CFSR station. However, missing daily values within a variable were completed by the built-in SWAT weather generator. Since these 42 stations are mostly located in the eastern part of the UBNRB (Figure 1), subbasins in the western region used weather data from their nearest stations. Figure 2 shows the precipitation distribution in the UBNRB using CFSR and ground stations. The map created from the ground stations (Figure 2, right) showed a precipitation distribution in the western region that is the result of SWAT using the precipitation values from the nearest stations.

This integrated dataset contained more data for the variables required by SWAT, which provided more reliable values and distribution for the weather variables than those provided by the CFSR and ground datasets separately.

2.5 Parameterization for the calibration and validation of the models

One of the major constrains of hydrological modeling is the difficulty of the parameterization of different variables (Hauhs & Lange, 2008). The correct combination of the values of the parameters

influencing the ground water, runoff and evapotranspiration processes is a key point on a model calibration. The characterization of watersheds considering their most influential variables is a good approach to determine the predictive capabilities of a model (McDonnell *et* al., 2007). Initially, it is recommended to perform calibrations for annual discharge values, once acceptable results are acquired; a calibration based on monthly values can be performed to achieve more detailed results (Neitsch *et* al., 2009).

SWAT divides the watershed into HRUs based on soil types, land coverage and slope. Therefore, a potential value can be assigned for each parameter and for each HRU, which will generate a large number of parameters. However, these values can also be applied as a global modification to estimate parameters by multiplying or adding values. Table 2 shows the parameterization applied to the respective areas in the watershed to calibrate the stream flow at Kessie and Eldiem. The same parameters showed in Table 2 were applied to all the models with different delineation and data sources. Where *r* stand for relative values and *v* for values to be replaced. Land coverage, soil types and slope have great impact on the total watershed discharge. Therefore, the values of the parameters were modified within the ranges specified by the SWAT Input/Output Documentation 2012 (Arnold *et* al., 2012). For instance, the available water content of the soils were calibrated in such a way that they did not change the physical properties of the soils. The Curve Number 2 (CN2) values were defined within different ranges based on the type of land cover. Calibration of models with wrong parameters values will only produce models with good statistical results but with less realistic representation of the actual properties of the watershed.

2.6 Statistical indices and SWAT quality analyses

2.6.1 Calibration and validation with flow discharge

In the case of hydrological modeling, limitation with the data quality and capabilities of the model to represent the complexity of the hydrological process often constitute obstacles. Therefore, models must be calibrated, and a statistical analysis is also required to see how reliable the results of the model are prior to their applications (Bastidas et al 2002). Since sediment data for the UBNRB is very limited, the calibration and validation of the models were done using flow discharge data only. The calibrated stations were Kessie and Eldiem at the mainstream of the Blue Nile River (Figure 1). With the values provided in Table 2 it was possible to obtain good statistical values for the calibrated models (Figures 4, 5, 6 and 7). For the automatic calibration SUFI-2 was used to efficiently calculate the coefficient of determination (R^2) and Nash Sutcliffe coefficient (NS) as likelihood measures, trying to catch the seasonal dynamics and magnitudes of the measured discharge data.

The coefficient of determination (R^2) is a measure of how well the regression line represents the data and gives a measure of the proportion of the fluctuation of a variable that is predictable from another variable. The values for this coefficient denote the strength of the linear relation between *x* and *y*,

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

representing the percentage of the data closest to the line of best fit. The R^2 objective function provided in SWAT-CUP (Abbaspour, *et* al., 2015) is as follows:

$$R^2 = \frac{\left[\sum_{i=1}^{n}(Q_{m,i} - \bar{Q}_m)(Q_{s,i} - \bar{Q}_s)\right]^2}{\sum_{i=1}^{n}(Q_{m,i} - \bar{Q}_m)^2 \sum_{i=1}^{n}(Q_{s,i} - \bar{Q}_s)^2}$$

Equation (7)

where Q are discharge values, m and s stand for observed and simulated values, respectively, and i is the i^{th} measured or simulated data.

Nash-Sutcliffe coefficient (NS), is widely used as goodness-of-fit indicator that expresses the potential predictive ability of a hydrological model (Nash and Sutcliffe 1970). The Nash-Sutcliff objective function provided in SWAT-CUP (Abbaspour, *et* al., 2015) is as follows:

$$NS = 1 - \frac{\sum_{i=1}^{n}(Q_m - Q_s)_i^2}{\sum_{i=1}^{n}(Q_{m,i} - \bar{Q}_m)^2}$$

Equation (8)

where Q are discharge values, m and s stand for observed and simulated data, respectively, and the bar stands for the average values.

The uncertainty in SUFI - 2 is expressed as an uniform distribution of parameters ranges, and parameters uncertainties are considered for any possible source in variables, for instance model inputs, model structure, model parameters and also measured data (Abbaspour, *et* al., 2015). The uncertainties in the outputs are expressed as the 95% prediction uncertainty or 95PPU. The uncertainty analysis in SUFI-2 is based on the concept that a single parameter value generates a single model response, while a parameter range or propagation of the parameter uncertainty leads to the 95PPU. In SUFI-2 the objective is to capture most of the observed values within the 95PPU range at the same time that thinner 95PPU range is preferable. Therefore, the simulation starts assuming large and physically meaningful parameter ranges, so that the measure data falls within the 95PPU, and continuously decreases the ranges of the 95PPU and produces better results. The final 95PPU is the 95% of the observed data captured within the final 95PPU band, which is defined by the final parameters intervals. Therefore, the best simulation is the best iteration within the 95PPU, and considering that is difficult to claim a specific parameter range for a certain watershed, then any solution within the 95PPU should be an acceptable solution.

2.6.2 Actual evapotranspiration analysis

With the objective of verifying and comparing the evapotranspiration results of the models, the actual evapotranspiration data for the UBNRB was obtained from the MODIS Global Terrestrial

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Evapotranspiration Project (MOD16). This is a global estimated data from land surface by using satellite remote sensing data. This data is intended to be used to calculate regional water balances, hence a very important source of data for watershed management and hydrological models analyses. The original MOD16 ET algorithm (Mu *et* al., 2007) was based on the Penman-Monteith equation (Monteith, 1965). However, the current MOD16 ET has used the improved evapotranspiration algorithm (Mu *et* al., 2011).

In this improved algorithm, the sum of the evaporation from the wet canopy surface, transpiration from the dry canopy surface and evapotranspiration from the soil surface constitute the total daily ET (Mu *et* al., 2011). The formulae for the total daily ET (λE) and potential ET (λE_{POT}) are:

$$\lambda E = \lambda E_{wet\_C} + \lambda E_{trans} + \lambda E_{SOIL}$$
$$\lambda E_{POT} = \lambda E_{wet\_C} + \lambda E_{POT\_trans} + \lambda E_{wet\_SOIL} + \lambda E_{SOIL_{POT}}$$

Equation (9)

where λE_{wet_C} is the evaporation from the wet canopy surface, λE_{trans} is the transpiration from the dry canopy surface (plant transpiration), λE_{SOIL} is the evaporation from the soil surface, λE_{POT_trans} is the potential plant transpiration, $\lambda E_{SOIL_{POT}}$ is the potential soil evapotranspiration.

The calibration and validation of the models with flow discharge constitute an important part of the quality analysis of a model. Additionally to this analysis, comparisons with evapotranspiration data can provide more details to quantify the reliability of hydrological models. Previous studies have already shown that the annual ET derived from the MOD16 algorithm are lower than those provided by hydrological models. For instance, Ruhoff (2013) detected an underestimation of 21% in the evapotranspiration provided by MOD16 in the Rio Grande basin, Brazil, where the underestimation was mainly caused by the misclassification of the land use. Sun Z. *et* al. (2007) also identify certain disadvantages in the MOD16 evapotranspiration. Nevertheless, a comparison of the SWAT models with satellite evapotranspiration data could help to more accurately identify the level of reliability of the models and also to show the performance of the proposed SWAT Error Index (SEI). The models under analysis correspond to the period 1980-2004, however MOD16 ET data is available only for the period 2000-2010. Therefore, the comparison was done only for 5 years, from 2000-2004, corresponding to a part of the validation period of the SWAT models.

2.6.3 SWAT Error Index (SEI)

Several previous studies have modeled the entire and also small catchments of the Nile Basin (Tibebe *et* al. 2011; Setegn *et* al. 2009a; Setegn *et* al. 2010; Swallow *et* al. 2009; Mulungu *et* al. 2007) providing good and meaningful results. However, common problem of hydrological models is the wrong combination of the values of the calibrated parameters, which can also lead to good graphical results, consequently good statistical values, but wrong water balance values. Consequently, good R^2 and NS values do not always denote the reliability of a model. R^2 and NS are common statistical

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

parameters used to evaluate and compare time series in hydrological models (Abbaspour, 2015; De Almeida Bressiani *et* al., 2015: Dile and Srinivasan, 2014; Gebremicael *et* al., 2013). Additionally in this study, after good calibration and validation values for R^2 and NS were achieved, and after a comparison between the SWAT ET and MOD16 ET values was completed, an index to quantify the models reliability has been introduced, the SWAT Error Index (SEI). This index provides an assessment of the quality of the SWAT model, where the relative Root Mean Square Error (rRMSE) was chosen as fitting function. SWAT Error Index (SEI) has been proposed as a tool to identify the quality of the models considering weighting values for the rRMSE results achieved from the calibration and validation of the models with discharge data, and the results obtained from the analysis performed on evapotranspiration.

Since SEI is using rRMSE values for discharge and evapotranspiration, a model with a good SEI results represents a model with a good agreement between these two hydrological processes, which are two important processes influencing the water balance. Therefore, by evaluating the SEI results, better models with a more reliable representation of the actual runoff and evapotranspiration processes of a watershed can be created. The proposed equation for SEI is as follows:

$$SEI = W_1\left(\frac{\left(\sqrt{\frac{\sum_{i=1}^{n}(Q_{oi}-Q_{si})^2}{N}}\right)}{(Q_{o\,max}-Q_{o\,min})}\right) + W_2\left(\frac{\left(\sqrt{\frac{\sum_{i=1}^{n}(ET_{oi}-ET_{si})^2}{N}}\right)}{(ET_{o\,max}-ET_{o\,min})}\right)$$

Equation (10)

The first part of the equation corresponds to the rRMSE of the values obtained from the discharge data, where, Q_{oi} is the observed discharge data (m^3/s), Q_{si} is the simulated discharge data (m^3/s), Q_{omax} is the maximum value of the observed discharge data and Q_{omin} is the minimum value of the observed discharge dataset. The second part of the formula corresponds to the rRMSE achieved from the evapotranspiration data that was obtained from MODIS16, where, ET_{oi} is the MOD16 evapotranspiration values, ET_{si} is the SWAT simulated evapotranspiration data, ET_{omax} and ET_{omin} are the maximum and minimum values of the MOD16 evapotranspiration data, respectively. W_1 and W_2 are the assigned weighted values for discharge and evapotranspiration, respectively.

Nowadays, several reliable measured flow discharge datasets are available for rivers, but that is not the case for evapotranspiration data. However, satellite evapotranspiration data is available for most watersheds in the world. Furthermore, the measured discharge dataset and the satellite estimated evapotranspiration dataset, do not have the same level of reliability. Therefore, different weighting values (W_1 and W_2) were assigned to define differences in the level of reliability of the datasets, 0.7 for flow discharge and 0.3 for evapotranspiration. *SEI* ranges from 0 to $+\infty$, with 0 corresponding to the ideal value. The closer the SEI value of the model is to 0, the model will have a better match with the flow discharge and the evapotranspiration data.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

3 Results and discussions

3.1 Impact of different discretization levels of a catchment and rain gauge combinations

After analyzing the different datasets under different discretization levels, it was detected that not only the input data and the parameterization have a critical impact on the water balance, but also the subbasins distribution. The water balance analysis was done for two calibrated stations, three datasets, and two different basins distributions. Water balance results for the Upper Blue Nile River Basin (UBNRB) at Eldiem and Kessie, and also the values for the different hydrological processes and models are given on table 3, referenced values for these hydrological processes are also given on table 1 (Cherie N. Z., 2013). The uncertainty of the rainfall data in the UBNRB basin is noticeable when models with different subbasins delineations are compared and show different values (see figures 4-5 for Eldiem and 6-7 for Kessie, under 30 and 87 subbasins, respectively). Based on other literature reviews, the average annual precipitation in the UBNRB is estimated to be approximately 1300 mm/year (Table 1). Results from the SWAT models (Table 3) show that the average annual precipitation given by the three datasets have values within acceptable ranges, with only a small overestimation for the CFSR dataset under 87 subbasins.

Figures 4 and 5 shows the magnitude and dynamics of the measured and estimated monthly discharge flow at Eldiem. The integrated dataset provided good statistical values for R^2 and NS (Table 4) under both discretization levels. The other models using the ground and CFSR datasets also showed good R^2 results, but very low NS values under both discretization levels (Table 4, Fig. 4 and 5). Although R^2 is always high in all the models, R^2 is a coefficient that measures only the dynamic of a model. Meaning that the models behave with accuracy matching the seasonality of the rainfalls and dry periods. However, NS is probably a more important factor to be considered as it can be used to quantitatively describe the accuracy of models outputs. Calibrations and validations at Kessie were good for the models using the ground and integrated datasets, achieving good R^2 and NS values (Table 4, Fig. 6 and 7), and also good water balance results (Table 3).

CFSR data provides a trustable annual average precipitation of 1221 mm under the model with 30 subbasins (Table 3). Under the model with 87 subbasins the average annual precipitation increases to 1427 mm. The rainfall overestimation provided by the CFSR dataset is caused by the number of subbasins, the models under 87 subbasins considered more stations than the model with 30 subbasins. However, this precipitation overestimation compared to the other two datasets is still within acceptable ranges for the watershed based on other studies (Table 1), and it is not the main factor affecting the water balance, but its distribution in the watershed. CFSR dataset is significantly underestimating the precipitation in the eastern region of the UBNRB (Figure 3). While observed values for Kessie (Table 3) show an average annual precipitation of 1082 mm and 1090 mm, CFSR only provided 631 mm and 811 mm, for the 30 and 87 subbasins models, respectively. Figures 6 and 7 showed how CFSR data is underestimating the precipitation compared to that provided by the ground data. Figure 6 and 7 also

showed the effect of the number of subbasin on the simulated discharge flow. The flow discharge provided by the CFSR is higher for the model with 87 subbasins, although in both cases this dataset continues to underestimate the flow discharge. The differences in the results of the models with different datasets showed that the number of subbasins can affect the water balance of a model. The difference is more significant when using CFSR data. As the precipitation in the watershed changes in magnitude and distribution, the parameterization for the calibration of the model will be different. Therefore, in order to meet good R^2 and NS for the model with a wrong precipitation distribution (in this case the CFSR data), the values of the parameters need to be modified to unrealistic values.

3.2 Average annual evapotranspiration and the impact of different data sources and PET methods

The evapotranspiration has been another critical factor subject to analysis in this study. Precipitation percentage values of evapotranspiration are frequently given for the Upper Blue Nile River Basin (UBNRB). But as it was shown in the water balance analyses (Table 3), the percentages would yield totally different amounts, for instance CFSR data defers significantly. The MOD16 evapotranspiration data used to evaluate the results of the SWAT models are estimated using the Mu *et* al.'s improved ET algorithm (2011) based on the Penman-Monteith equation (Monteith, 1965). However, the SWAT models were calibrated under the Hargreaves evapotranspiration method. And as it was previously mentioned, the MOD16 algorithm tends to underestimat the evapotranspiration in this case. Therefore, a comparison of the actual evapotranspiration data provided by MOD16 with the values calculated by SWAT under Hargreaves and Penman-Monteith equations was done to show the level of discrepancy between data sets (Figures 8 and 9). Models using Hargreaves equation showed acceptable discharge values and trends compared to those of measured discharge data (Figures 4 and 5). However, the models overestimated the evapotranspiration values compared to those provided by MOD16 (Figure 8). Nevertheless, when using the Penman-Monteith method, the SWAT models provided more similar evapotranspiration values, better R^2 and NS values compared to the values given by the MOD16 evapotranspiration data (Figure 9). It is important to mention that the SWAT models were not calibrated to match those values of the MOD16 evapotranspiration dataset. The best evapotranspiration values are obtained using the CFSR dataset (Figure 9). This model provided low evapotranspiration (Figure 9) values consequently overestimated the flow discharge values.

3.3 SWAT Error Index (SEI) evaluation

Representing the dynamics and magnitudes of the measured and simulated values by models, NS and R^2 have been considered the most significant parameters to determine the reliability of hydrological models. Additionally, rainfall distribution, parameterization and evapotranspiration are also crucial points to be considered in any hydrological model. Therefore, this study proposed the SWAT Error Index (SEI), which in addition to the flow discharge data also considered the behavior of the evapotranspiration to define the level of error of a model. The underestimation of the MOD16

evapotranspiration values for the Upper Blue Nile River Basin was shown in Fig. 8, therefore it is not a trustable source in terms of magnitude. Nevertheless the dynamics of the evapotranspiration seems to have similar behavior to the evapotranspiration obtained from SWAT. Therefore, a relative comparison between both dataset is able to provide a measure of fit, and will help to understand the behavior of SEI. This index intends to define the level of error of a model, and also its level of realistic representation of the actual characteristics of the watershed. By including flow discharge data and evapotranspiration data, the level of accuracy in the combination of the values of the parameters in SWAT increases, and is less expectable to have a wrong parameterization. SEI values closer to 0 mean that the model have no error, which will be a perfect match of the model to the flow discharge and evapotranspiration data.

SEI results (Table 5) show that the behavior and capability of SEI to quantify the level of error of a model is good. The behavior of SEI in representing both flow discharge and evapotranspiration is good. For instance, the values in the tables show that the lower the index value of the discharge data is, the value for evapotranspiration tends to increase. This is because the flow discharge data is being matched, however the evapotranspiration increases and tends to overestimate those value provided by MOD16 ET. If MOD16 ET had a good representation of the evapotranspiration data of a watershed, then SEI values for both discharge and evapotranspiration values should be closer to 0. SEI results showed the accuracy of the models; which in this case is low due to the discrepancy of the evapotranspiration values used as a reference to compare the results of the models. However, SEI (Table 5) showed that the models using the integrated datasets are more reliable than the other two datasets, achieving a SEI of 0.28 and 0.26 for the models with 30 and 87 subbasins, respectively. It also demonstrated that the CFSR dataset is less accurate, with SEI values of 0.43 and 0.45 for the models with 30 and 87 subbasins, respectively.

4 Conclusions

Two weather data sources: CFSR and observed ground data were analyzed in terms of statistical results, water balance and precipitation distribution in the Upper Blue Nile River Basin. After detecting their limitations and disadvantages, an integration of both datasets was proposed with the purpose of overcoming their uncertainties and limitations.

A comparison of the three datasets under different discretization levels was also performed. This comparison was important to obtain a better understanding of how crucial the subbasin discretization process is during a SWAT model setup. The two different subbasins distributions included a model under 30 subbasins with 242 HRUs, and another model under 87 basins and 3227 HRUs, which increased the models complexity and level of details. The comparisons showed that the three input datasets, under models with different number of subbasins, yield different results. The number of subbasins in a SWAT model will affect the magnitude of the flow discharge, hence the total water balance of a watershed.

The comparison of the results of the model with 30 subbasins demonstrates that the values for the total annual average precipitation at Eldiem are similar for the three datasets (Table 3). Nevertheless, only the model using the CFSR dataset was not able to achieve good water balance results under similar parameterization. The magnitude of the precipitation is good, however its distribution is not very reliable. CFSR data underestimates the precipitation in the eastern region of the Upper Blue Nile River Basin and overestimates it in the western region. Table 3 shows the annual average precipitation values at Kessie, where the CFSR data is providing a value of 631 mm/year, while the ground dataset gave a value of 1082 mm/year.

For the second case, the three datasets were analyzed in more details using 87 subbasins. The ground and integrated datasets showed good results using a realistic parameterization, and also achieved slightly better statistical results than those obtained from the 30 subbasins models (Table 4 and 5). However, CFSR data showed an overestimation of approximately 200 mm at Eldiem, this was caused by the number of subbasins. As it was experience in the first case, the distribution continued to show an underestimation of precipitation at eastern region and was not possible to obtained good statistical results using a realistic parameterization. The annual average precipitation at Kessie provided by the CFSR data was 811 while the ground dataset showed average values of 1090 mm/year (Table 3).

This paper also proposed a SWAT Error Index (SEI), which was very helpful to obtain a more accurate concept of the reliability of a SWAT model. This index used the weighted relative Root Mean Square Error (rRMSE) of the discharge and evapotranspiration data. Furthermore, by performing an analysis of the SEI results it was possible to identify which hydrological processes need to be improved in order to reduce the level of error of a model. Therefore, SEI demonstrated to be a reliable and useful tool to develop better SWAT models. SEI also showed that the integrated dataset successfully achieved better and more reliable results than the ground and CFSR datasets.

The integrated dataset improved the results of the model, obtaining better R^2, NS and SEI values. The model under 87 subbasin was the model that provided more details in terms of land use, soil and slope classes; and also achieved better statistical values. Therefore, this model is more suitable to perform several types of hydrological analyses and propose watershed management practices in the UBNRB.

5 Acknowledgements

Authors are very grateful with the Deutscher Akademischer Austauschdienst and the International Graduate School of Science and Engineering of the Technische Universität München for their financial support. Also to the National Meteorology Agency (NMA) and the Ministry of Water and Energy of Ethiopia for providing weather and flow discharge data, respectively.

6 References

Abbaspour, K. C.: SWAT-CUP: SWAT calibration and uncertainty programs-A user manual, Tech. rep., Swiss Federal Institute of Aquatic Science and Technology, Eawag, Dübendorf, Switzerland, 2015.

Allen RG, Pereira LS, Raes D & Smith M. 1998: Crop Evapotranpiration: Guildlines for computing crop water requirements, FAO Irrigation and Drainage Paper No. 56, Food and Agriculture Organization, Land and Water, Rome, Italy.

Arnold JG, Kiniry JR, Srinivasan R, Williams JR, Haney EB & Neitsch SL. 2012: Soil & Water Assessment Tool, Input/Output documentation, version 2012, Texas Water Resources Institute, 246-248.

Bastidas LA, Gupta HV & Sorooshian S. 2002: Emerging paradigms in the calibration of hydrologic models, Mathematical Models of Large Watershed Hydrology, Water Resources Publications, LLC, Englewood, CO, USA, 1, pp.25-56.

Cherie NZ. 2013: Downscaling and Modeling the Effects of Climate Change on Hydrology and Water Resources in the Upper Blue Nile River Basin, Ethiopia.

De Almeida Bressiani, D., Srinivasan, R., Jones, C. A., and Mendiondo, E. M.: Effects of different spatial and temporal weather data resolutions on the stream flow modeling of a semiarid basin, Northeast Brazil, Int. J. Agr. Biol. Eng., 8, 1–16, doi:10.3965/j.ijabe.20150803.970, 2015.

Dile, YT & Srinivasan, R. 2014: Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. JAWRA Journal of the American Water Resources Association, 50(5), 1226-1241, doi:10.1111/jawr.12182, 2014.

FAO. 2015: World reference base for soil resources 2014: International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Report, 106.

Gassman WP, Reyes MR, Green CH & Arnold JG. 2007: The soil and water assessment tool: historical development, applications, and future research direction, T. ASABE, 50, 1211-1250.

Gebremicael, T. G., Mohamed, Y. A., Betrie, G. D., van der Zaag, P., and Teferi, E.: 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, J. Hydrol., 482, 57–68, doi:10.1016/j.jhydrol.2012.12.023, 2013.

Griensven A, Ndomba P, Yalew S, and Kilonzo F. 2012: Critical review of SWAT applications in the upper Nile basin countries, Hydrol. Earth Syst. Sci., 16, 3371–3381, 2012.

Hargreaves GH & Samani ZA. 1982: Estimating potential evaporation. J. Irrig. Drain. Eng.-ASCE, 108(3), 225-230.

Hauhs M & Lange H. 2008: Classification of runoff in headwater catchments: A physical problem?, Geography Compass, 2(1), 235–254.

Malunjkar V. S., Shinde M. G., Ghotekar, S. S. & Atre. A. A.: Estimation of Surface Runoff using SWAT Model. International Journal of Inventive Engineering and Sciences (IJIES) ISSN: 2319–9598, Volume-3 Issue-4, March 2015.

McDonnell JJ, Sivapalan M, Vache K, Dunn S, Grant G, Haggerty R, Hinz C, Hooper R, Kirchner J, Roderick ML, Selker J & Weiler M. 2007: Moving beyond heterogeneity and process complexity: a new vision for watershed hydrology, Water Resour. Res., 43, 1-6.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Reviews of Geophysics, 48, doi:10.1029/2009RG000314, 2010.

Me, W., Abell, J. M., and Hamilton, D. P.: Effects of hydrologic conditions on SWAT model performance and parameter sensitivity for a small, mixed land use catchment in New Zealand, Hydrol. Earth Syst. Sci., 19, 4127-4147, doi:10.5194/hess-19-4127-2015, 2015.

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

Mu, Q., Heinsch, F. A., Zhao, M., & Running, S.W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sensing of Environment, 111, 519–536.

Mu, Q., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sensing of Environment, 115, 1781–1800.

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.

Nash JE & Sutcliffe JV. 1970: River flow forecasting through conceptual models, Part I: A discussion of principles, Journal of Hydrology 10(3), pp.282–290.

Neitsch SL, Arnold JG, Kiniry JR, & Williams JR, 2009: Soil and Water Assessment Tool Theoretical Documentation Version 2009, Grassland, Soil and Water Research Laboratory and Black land Research Center, Texas.

Rafiei Emam, A., Kappas, M., Hoang Khanh Nguyen, L., and Renchin, T.: Hydrological Modeling in an Ungauged Basin of Central Vietnam Using SWAT Model, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-44, in review, 2016.

Roth, L., & Lemann, T. 2016: Comparing CFSR and conventional weather data for discharge and soil loss modelling with SWAT in small catchments in the Ethiopian Highlands, Hydrol. Earth Syst. Sci., 20, 921–934, 2016.

Ruhoff A.L., Paz A.R., Aragao L.E.O.C., Mu Q., Malhi Y., Collischonn W., Rocha H.R. & Running S.W. 2013: Assessment of the MODIS global evapotranspiration algorithm using eddy covariance measurements and hydrological modelling in the Rio Grande basin, Hydrological Sciences Journal – Journal des Sciences Hydrologiques, 58 (8) 2013.

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., 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.

Shimelis GS, Ragahavan Srinivasan & Bijan Dargahi. 2008: Hydrological Modelling in the Lake Tana Basin, Ethiopia Using SWAT Model, The Open Hydrology Journal, 2, 49-62.

Sivapalan M. 2006: Pattern, process and function: Elements of a unified theory of hydrology at the catchment scale, Encyclopedia of Hydrological Sciences, edited by: Anderson, M. G., John Wiley & Sons, Ltd.

Sun Z, Wang Q., Ouyang Z., Watanabe M, Matsushita B & Fukushima T. 2007: Evaluation of MOD16 algorithm using MODIS and ground observational data in winter wheat field in North China Plain, HYDROLOGICAL PROCESSES Hydrol. Process. 21, 1196–1206 (2007).

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. Pol., 12, 504–519, doi:10.1016/j.envsci.2008.11.003, 2009.

Tekleab S., Uhlenbrook S., Mohamed Y., Savenije H. H. G., Temesgen M., and Wenninger J.: Water balance modeling of Upper Blue Nile catchments using a top-down approach, Hydrol. Earth Syst. Sci., 15, 2179–2193, 2011.

Tibebe, D. and Bewket, W.: Surface runoff and soil erosion estimation using the SWAT model in the Keleta watershed, Ethiopia, Land Degrad. Develop.,22, 551–564, doi:10.1002/ldr.1034, 2011.

Wang, H. and Sun, F.: Impact of LUCC on Streamflow using the SWAT Model over the Wei River Basin on the Loess Plateau of China, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-332, in review, 2016.

Worqlul, A. W., Maathuis, B., Adem, A. A., Demissie, S. S., Langan, S., and Steenhuis, T. S.: Comparison of rainfall estimations by TRMM 3B42, MPEG and CFSR with ground-observed data for the Lake Tana basin in Ethiopia, Hydrol. Earth Syst. Sci., 18, 4871–4881, doi:10.5194/hess-18-4871-2014, 2014.

Zhang L, Dawes WR & Walker GR. 1999: Predicting the effect of vegetation changes on catchment average water balance, Tech. Rep.99/12 Coop. Res. Cent. Catch. Hydrol., Canbera, ACT.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

7 List of tables

Table 1: Average annual water balance components of the Upper Blue Nile River Basin for a 30 years period (1971-2000), calibration and validation (N. Z. Cherie, 2013).

Hydrologic parameters	1971-2000 (mm/year)	Calibration period 1976-1982 (mm/year)	Validation period 1992-1995 (mm/year)
Precipitation	1358	1338	1348
Evapotranspiration	941	962	960
Revap/shallow aquifer	63	59	58
Surface runoff	165	143	151
Return flow	107	70	38
Recharge to deep aquifer	10	9	9

Table 2: Parameterization of the models using the ground dataset and SUFI-2 algorithm.

No.	Parameter	Description	Type of change	Fitted absolute values	
				Eldiem	Kessie
1	CN2	Curve number for moisture condition II	r	72	77
2	SOL_AWC	Available water capacity of the soil	r	0.14-0.47	0.14-0.47
3	ESCO	Soil evaporation compensation factor	v	0.1	0.05
4	ALPHA_BF	Baseflow alpha factor	v	0.05	0.09
5	EPCO	Plant uptake compensation factor	v	0.6	0.8
6	RCHRG_DP	Deep aquifer percolation fraction	v	0.01	0.01

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Table 3: Water balance analysis in the Upper Blue Nile River Basin.

Hydrological Parameters	Model with 30 subbasins			Model with 87 subbasins		
	CFSR Data	Ground Data	Integrated Data	CFSR Data	Ground Data	Integrated Data
Water balance at Eldiem (All values in mm/year)						
Precipitation	1221	1190	1190	1427	1211	1211
Evapotranspiration	806	813	823	817	807	837
Revap/shallow aquifer	57	61	64	33	42	41
Surface runoff	152	156	162	207	160	145
Return flow	112	107	116	269	133	141
Recharge to deep aquifer	9	8	8	16	8	9
Water balance at Kessie (All values in mm/year)						
Precipitation	631	1082	1082	811	1090	1090
Evapotranspiration	435	778	787	520	801	806
Revap/shallow aquifer	14	63	61	15	25	23
Surface runoff	102	129	134	137	117	109
Return flow	61	97	91	103	37	48
Recharge to deep aquifer	2	9	4	8	10	9

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

Table 4: Statistical results for the calibrations and validations at Eldiem and Kessie.

		CFSR dataset		Ground dataset		Integrated dataset	
Subbasins		30	87	30	87	30	87
Eldiem							
Calibration	**R^2**	0.92	0.91	0.88	0.90	0.87	0.91
	NS	0.15	0.05	0.68	0.61	0.82	0.84
Validation	**R^2**	0.90	0.89	0.92	0.92	0.90	0.90
	NS	0.05	0.10	0.73	0.69	0.87	0.89
Kessie							
Calibration	**R^2**	0.80	0.79	0.74	0.75	0.77	0.79
	NS	0.45	0.38	0.68	0.72	0.70	0.71
Validation	**R^2**	0.82	0.75	0.74	0.80	0.78	0.79
	NS	0.52	0.47	0.70	0.68	0.74	0.75

Table 5: SWAT Error Index results for the UBNRB.

SWAT with 30 subbasins				
	Weighting	CFSR Dataset	Ground Dataset	Integrated Dataset
Discharge Data rRMSE	(0.7)	0.37 (0.259)	0.13 (0.091)	0.12 (0.084)
Evapotranspiration rRMSE	(0.3)	0.58 (0.174)	0.63 (0.189)	0.67 (0.201)
SEI	0 best	**0.43**	**0.28**	**0.28**
SWAT with 87 subbasins				
	Weighting	CFSR Dataset	Ground Dataset	Integrated Dataset
Discharge Data rRMSE	(0.7)	0.45 (0.315)	0.22 (0.154)	0.13 (0.091)
Evapotranspiration rRMSE	(0.3)	0.46 (0.138)	0.47 (0.141)	0.57 (0.171)
SEI	0 best	**0.45**	**0.29**	**0.26**

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

8 List of figures

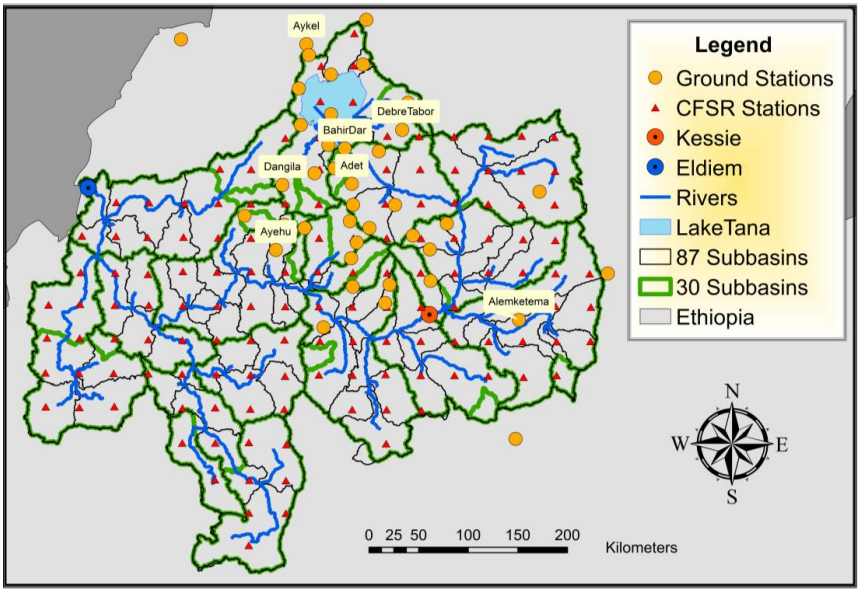

Figure 1: Weather and hydrometric gauging stations in the Upper Blue Nile River Basin under two discretization levels, 30 and 87 subbasins.

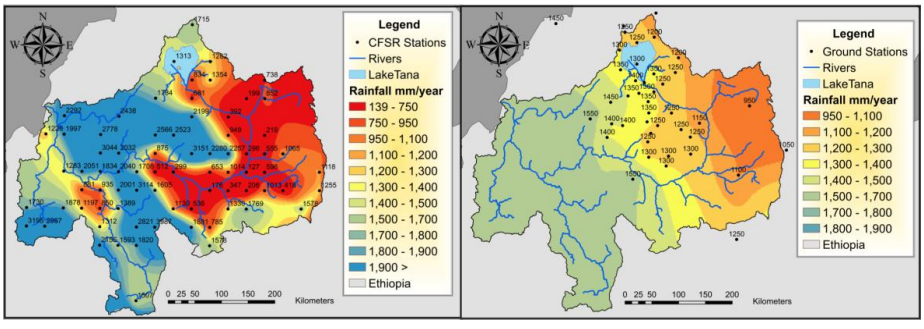

Figure 2: Spatial annual rainfall variation in the Upper Blue Nile River Basin using two different data sources: CFSR data (Left) and ground data (Right).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

Figure 3: Comparisons of ground and CFSR weather datasets.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

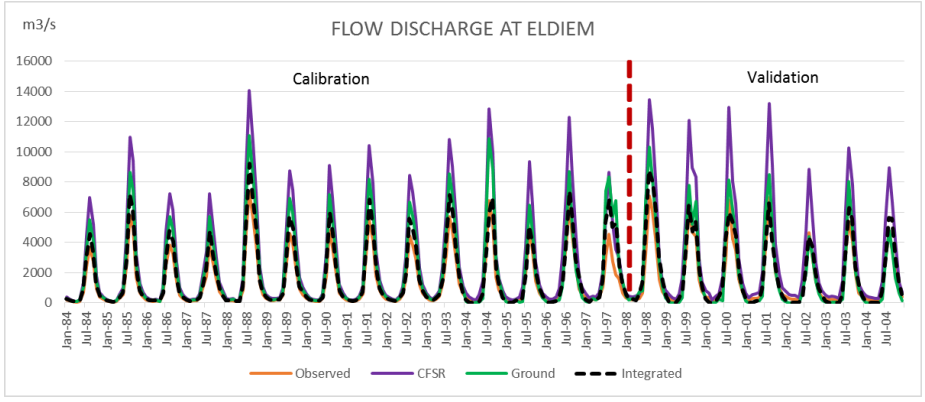

Figure 4: Calibration and validation results at Eldiem under 30 subbasins. Calibration results achieved R^2 and NS values of: Integrated data: 0.87, 0.82; ground data: 0.88, 0.68; CFSR data: 0.92, 0.15; respectively. Validation results achieved R^2 and NS of: Integrated data: 0.90, 0.87; ground data: 0.92, 0.73; CFSR data: 0.90, 0.05; respectively.

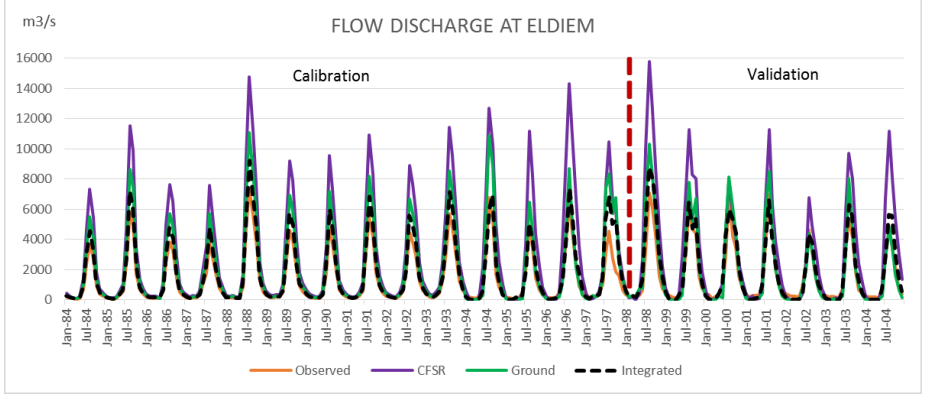

Figure 5: Calibration and validation results at Eldiem under 87 subbasins. Calibration results achieved R^2 and NS values of: Integrated data: 0.91, 0.84; ground data: 0.90, 0.61; CFSR data: 0.91, 0.05; respectively. Validation results achieved R^2 and NS of: Integrated data: 0.90, 0.89; ground data: 0.92, 0.69; CFSR data: 0.89, 0.10; respectively.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions

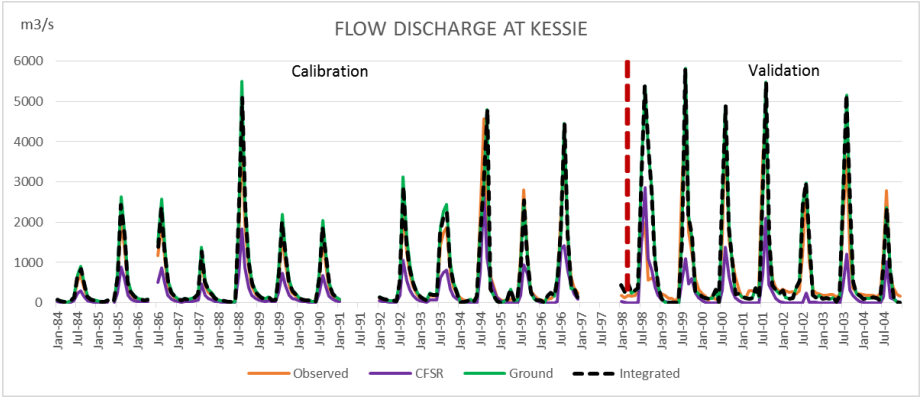

Figure 6: Calibration and validation results at Kessie under 30 subbasins. Calibration results achieved R^2 and NS values of: Integrated data: 0.77, 0.70; ground data: 0.74, 0.68; CFSR data: 0.80, 0.45, respectively. Validations results achieved R^2 and NS values of: Integrated data: 0.78, 0.74; ground data: 0.74, 0.70; CFSR data 0.82, 0.52; respectively.

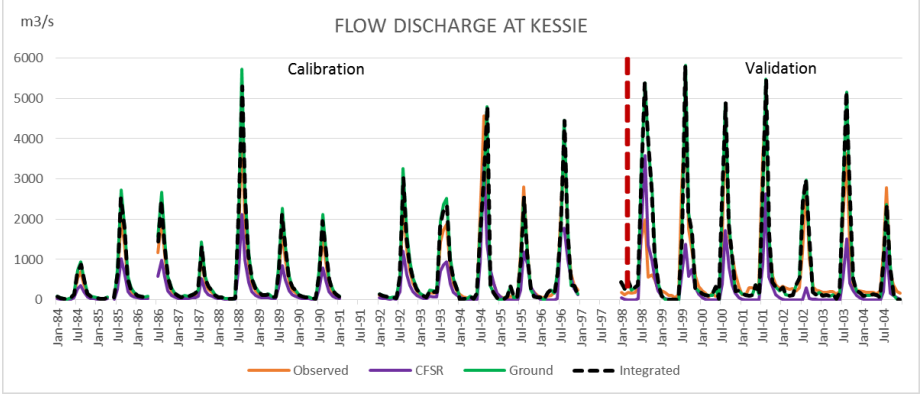

Figure 7: Calibration and validation results at Kessie under 87 subbasins. Calibrations results achieved R^2 and NS values of: Integrated data: 0.79, 0.71; ground data: 0.75, 0.72; CFSR data 0.79, 0.38; respectively. Validations results achieved R^2 and NS values of Integrated data: 0.79, 0.75; ground data: 0.80, 0.68; CFSR data 0.75, 0.47; respectively.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-664, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 16 February 2017

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

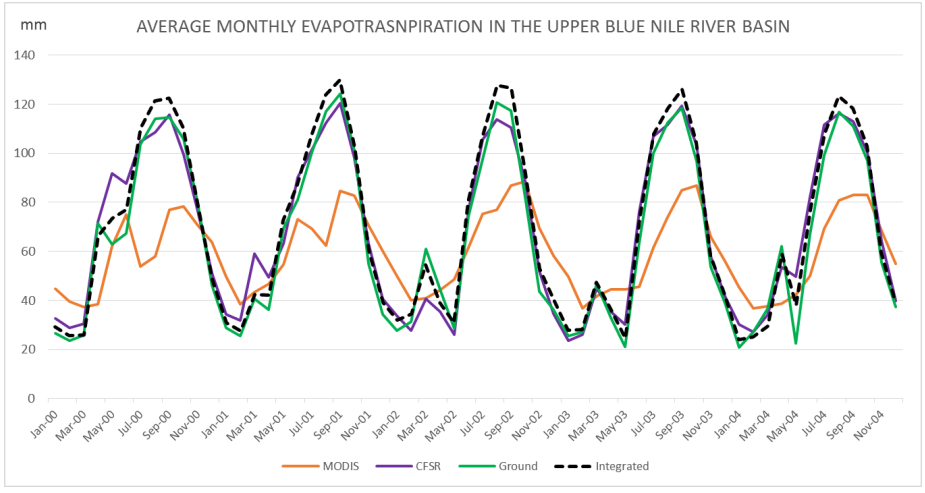

Figure 8: Average monthly evapotranspiration analysis using SWAT with 87 subbasins and the Hargreaves method, with R^2 and NS values of Integrated data: 0.63, -2.32; ground data: 0.60, -1.32; CFSR data: 0.63, -1.20; respectively, compared to the MODIS data.

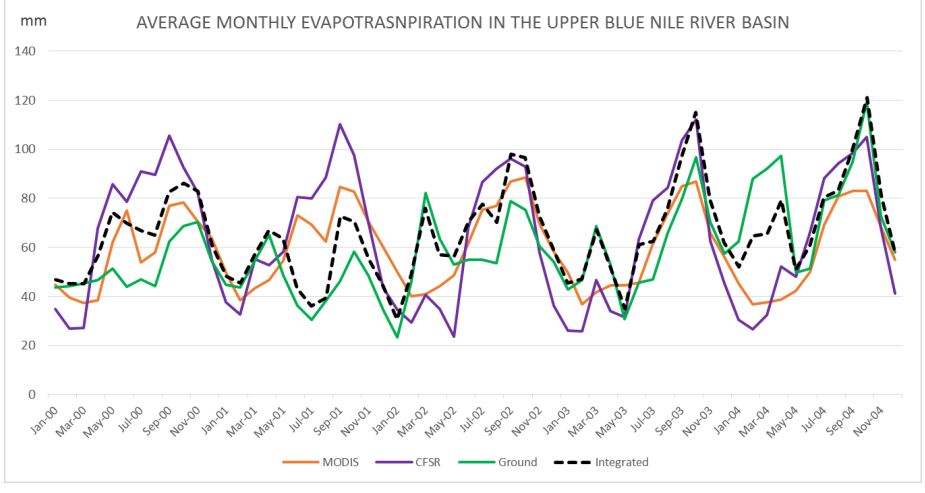

Figure 9: Average monthly evapotranspiration analysis using SWAT with 87 subbasins and the Penman-Monteith method, with R^2 and NS values of Integrated data: 0.36, -0.02; ground data: 0.34, -0.10; CFSR data: 0.74, 0.03; respectively, compared to the MODIS data.