Analysis of the combined and single effects of LULC and climate change on the streamflow of the Upper Blue Nile River Basin (UBNRB): Using statistical trend tests, remote sensing landcover maps and the SWAT model

5   Dagnenet F. Mekonnen[1, 2], Zheng Duan[1], Tom Rientjes[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]Amhara Regional State Water, Irrigation and Energy Development Bureau, Bahirdar, Ethiopia
[3]Department of Water Resources, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twenty,
10 Enschede, Netherlands
Correspondence to: Dagnenet F. (dagnfenta@yahoo.com)

Abstract: Understanding the response of land use/land cover (LULC) and climate change has become a priority issue for water management and water resource utilization of the Nile basin. This study assesses the long-term trends of rainfall and

15   streamflow to analyse the response of LULC and climate changes on the hydrology of the UBNRB. The Mann-Kendal (MK) trend tests showed no statistically significant changes in daily, monthly and annual rainfall. Tests for mean annual and seasonal streamflow showed a statistically significant and increasing trend. Landsat satellite images for 1973, 1985, 1995 and 2010 were used for LULC change detection. The LULC change detection findings indicate the conversion of forest land to cultivated land during the period 1973-2010. Natural forest decreased from 17.4% to 14.4%, 12.2% and 15.6% while

20   cultivated land increased from 62.9% to 65.6%, 67.5% and 63.9% from 1973 to 1985, 1995 and 2010 respectively.

The hydrological SWAT model result showed that mean annual streamflow increased by 15.6% between the 1970s and the 2000s due to the combined effect of LULC and climate change. The single effect of LULC change on streamflow analysis suggested that LULC change significantly affects surface run-off and base flow. This could be attributed to the 5.1%

25   reduction in forest coverage and 4.6% increase in cultivated land. Effects of climate change revealed that increased rainfall intensity and number of extreme rainfall events from 1971 to 2010 have greatly affected the surface run-off and base flow of UBNRB.

1. Introduction

The Nile basin, having the world's largest river (6700 km) and catchment area of 3.3 million km^2, is characterized by the

30   limited water resource where climatic and hydrological extremes such as floods and droughts hit the basin population

severely and regularly associated with scarce hydro-climatic data (Gebrekristos, 2015). Over 200 million people are estimated to rely directly on the Nile river for their food and water supply with projected increases on water demands and water uses. The direct and indirect impacts brought by both LULC and climate change exacerbate the water scarcity of the Nile basin as they are the key factors that can modify the hydrology and water availability of the basin. Furthermore,

5 unbalanced water utilization of the downstream countries 94% (Egypt and Sudan) remained the crucial sociopolitical issue for many years. To date, Ethiopia has utilized insignificant amount less than 5% of the Blue Nile water, as compared to downstream countries Sudan and Egypt.

Meanwhile, the Ethiopian government has planned and carried out studies to significantly increase large reservoir for water

10 storage in the Blue Nile basin both for irrigation and hydropower in order to support national development and get rid of poverty (BCEOM, 1998). However, as UBNR is a transboundary river, its development and management should be agreed and reached consensus between shared countries. Tackling all these complexities and developing the better water resource development strategies is only possible by understanding the hydrological processes of the basin. Therefore, scientific research on climatic and hydrological processes is needed to maximize water development activities and benefit from it. .

A literature review shows that there are few sub-basin and basin level studies carried out in the UBNRB, with most studies focusing on trend analysis of precipitation and flow. Considering precipitation, most studies e.g.,(Bewket and Sterk, 2005; Cheung *et al.*, 2008; Conway, 2000; Gebremicael *et al.*, 2013; Melesse *et al.*, 2009; Rientjes *et al.*, 2011; Seleshi and Zanke, 2004; Teferi *et al.*, 2013; Tekleab *et al.*, 2014; Tesemma *et al.*, 2010) report no significant trend in annual and seasonal

20 precipitation totals within the Lake Tana sub-basin, where there are relatively better hydro-meteorological data, while Mengistu *et al.* (2014) reported statistically non-significant increasing trends at annual and seasonal except Belg season.

For streamflow from the UBNRB (Gebremicael *et al.*, 2013) reported statistically significant increasing long-term mean annual flow at the El Diem gauging station. However, (Tesemma *et al.*, 2010) reported no significant long-term trend in

25 annual streamflow from the UBNRB at ElDiem gauging station, but significantly increasing at Bahirdar and Kessie stations. At the sub-basin scale, (Rientjes *et al.*, 2011) reported that low flows in the Gilgel Abay sub-basin decreased during the period (1973–2001), specifically 18.1% and 66.6% decrease for the periods 1982–2000 and 2001–2005, respectively. However, for the same periods, the high flows show an increase in 7.6% and 46.6% due to LULC change and seasonal variability of rainfall.

30

Although, substantial progress has been made in assessing the impacts of LULC and climate change on the hydrology of UBNRB, most studies focused on single aspects i.e., either analysing the statistical trend of precipitation and streamflow or analysing impacts of single factor LULC or climate change on the flow (Gebremicael *et al.*, 2013; Rientjes *et al.*, 2011; Tekleab *et al.*, 2014) Impacts by combined effects of LULC and climate changes are not well understood because their

contributions are difficult to separate and vary regionally (Yin *et al.*, 2017). However, proper water resource management requires an in-depth understanding on the aggregated and disaggregated effects of LULC and climate changes on streamflow as they are the most significant drivers of environmental change in the Nile basin.

5 Therefore, the objectives of this study are as follows (i) assess the long-term trend of rainfall and streamflow (ii) analyse the LULC change (iii) examining the streamflow responses to combined and isolated effects of LULC and climate changes in the UBNRB through a combined analysis of statistical trend test, satellite remote sensing LULC map and SWAT hydrological model during the period 1971-2010.

2. Study area

10 The UBNRB is located in the northwest of Ethiopia, between longitudes 34.30^0 and 39.45^0E and latitudes 7.45^0 and 12.45^0N, with an approximate area of 172,760 km². Topography of the basin is typically characterized by highlands, hills, valleys and occasional rock peaks with elevations that range from 500 m.a.s.l to above 4000 m.a.s.l (Figure 1). According to BCEOM (1998), the larger portion of the basin (2/3) lies in the highlands of Ethiopia with annual rainfall ranging from 800mm to 2,200 mm. Mekonnen and Disse (2016) showed that the UNBRB has a mean areal annual rainfall of 1452 mm, and a mean

15 minimum and maximum temperature of 11.4^oC and 24.7^oC respectively.

According to the classification of NMA (2013), there are three seasons in Ethiopia; namely, Belg (short rainy season), Kiremt (main rainy season) and Bega (dry season). Belg is a short rainy period from February to May, Kiremt is the period from June to September and Bega is the period from October to January. According to BCEOM (1998), the average annual

20 discharge is estimated about 49.4 BCM, with the low flow month (April) equivalent to less than 2.5 % of that of the high flow month (August), at the Ethio-Sudan border (El Diem). The analysis of this study revealed that the long-term (1971-2010) mean annual volume of flow at El Diem is 50.7 BCM, with the low flow (dry season) contributing 21.1% and the short rainy season accounting for about 6.2%, while most flow occurred during the rainy season, contributing about 73% (Table 1).

The land cover of the basin essentially follows the divide between highland and lowland. The highlands are predominantly covered by farmlands (about 90%), bush and shrubs. The lowlands, in contrast, are still largely untouched by development, so that as a result, woodlands, bush and shrub lands are the dominant forms of land cover (BCEOM, 1998).

3. Input data sources

In this study, long time series (1971-2011) hydro-meteorological data data were used for trend analysis. The streamflow data set was collected from the Federal Ministry of Water Irrigation and Electricity of Ethiopia. Daily precipitation, minimum and maximum air temperature, relative humidity, hours of sunshine and windspeed data at 15 stations from 1971-2010 were

5   obtained from the Ethiopian National Meteorological Agency (NMA). The monthly, seasonal and annual hydro-meteorological data was aggregated from the daily time series data.

The spatially distributed (GIS input) data for the SWAT hydrological model includes Digital Elevation Model (DEM), soil data and LULC maps. A Shuttle Radar Topographic Mission Digital Elevation Model (SRTM DEM) of 90 metres resolution

10 from the Consultative Group on International Agricultural Research-Consortium for Spatial Information (CGIAR-CSI; http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp) was used to delineate the watershed and to analyse the drainage patterns of the land surface terrain. The soil map developed by the Food and Agriculture Organization of the United Nations (FAO-UNESCO) at a scale of 1:5,000,000 downloaded from (http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/faounesco-soil-map-ofthe-world/en/) was used for SWAT model. The soil information such as soil textural and

15 physiochemical properties needed for the SWAT model was extracted from 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-databasev12/en/) in combination with information provided by FAO-UNESCO soil map. The LULC maps, representing one of the most important driving factors to affect surface run-off and evapo-transpiration in a basin were produced from satellite remote sensing Landsat images for 1973, 1985, 1995 and 2010.

After the raw rainfall and discharge data had been collected, the data was screened and corrected. In many stations, pronounced length of time series data was missed and hence stations with long time series and relatively little time missing records were selected. Filling missed or gap records was the first task for any further meteorological data analysis. This task was performed using the inverse distance weighing (IDW) and regression methods, the best performed method was chosen. .

25 ## 4. Methodology

4.1 Trend Analysis

The non-parametric Mann-Kendal (MK) (Kendall, 1975; Mann, 1945) statistics is chosen to detect trends for hydrologic time series data as it is widely used for effective water resource planning, design and management (Yue and Wang, 2004). MK tests do not require any assumptions about the distribution of the variables and effective when the sample data are

30 serially independent. Its advantage over the parametric tests, such as t-test, is that the non-parametric tests are more suitable

non-normally distributed, censored, missing data, which are frequently encountered in hydrological time series (Yue *et al.*, 2004). All the trend results in this paper have been evaluated at the 5% level of significance to ensure an effective exploration of the trend characteristics within the study area. However, the MK test result is affected by a serial correlation of the time series, if there is a positive serial correlation in the time series. According to von Storch (1995), the test will

5   suggest that a significant trend in a time series leads to a disproportionate rejection of the null hypothesis of no trend, which is actually true.

In order to limit the influence of serial correlation on the MK test, pre-whitening was proposed by von Storch (1995) and the Effective or Equivalent Sample Size (ESS) method developed by Hamed and Rao (1998) has also been proposed to modify

10   the variance. However, the study by (Yue *et al.*, 2002) reported that von Stroch's pre-whitening is effective only when no trend exists and the rejection rate of the ESS approach after correction is much higher than that it should be (Yue *et al.*, 2004). Then, Yue *et al.* (2002) proposed trend-free pre-whitening (TFPW) prior to applying the Mann-Kendall trend test in order to minimize the limitation of the MK test. This study therefore employed TFPW to remove the serial correlation, in order to detect a significant trend in a time data series with significant serial correlation. Further details can be found in (Yue

15   *et al.*, 2002).

MK calculates Kendall's statistics (S), the number of positive differences minus the number of negative differences, , to indicate an increasing or decreasing trend(Gilbert, 1987) using eqn.1-eqn.4. Positive values of those parameters indicate a general tendency towards an increasing trend, while negative values show a decreasing trend. Furthermore, a two-tailed

20   probability (p-value) was computed and compared with the user-defined significance level (in this study 5%) in order to identify the trend of variables. When the calculated P-value is greater than the defined significance level (5%), then it indicates acceptance of null hypothesis (no trend) and the reverse is true.

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sgn(k_j - k_x) \tag{1}$$

25   If S is a positive number, it indicates, increasing trend. If S is a negative number, then observations made later tend to decrease. Where n is the data record length, k_j and kx are data sequential values, and the function sgn(x) the sign of all n(n - 1)/2 possible differences of K_j - K_x where j > x is defined as

$$sgn(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x\ 0 = \\ -1, & \text{if } x < 0 \end{cases} \tag{2}$$

30   According to Mann (1945) and Kendall (1975), when n > 8, the statistic S is approximately normally distributed with the mean and the variance as follows:

The mean of S is E(s)=0, and the variance δ^2

$$\delta^2 = \frac{\left\{ n(n-1)(2n+5) - \sum_{j=1}^{p} t_j(t_j-1)(2t_j+5) \right\}}{18} \qquad (3)$$

Where p is the number of tied groups in the data set and t_j is the number of data points in the jth tied group. The statistic S is

5   approximately normally distributed provided that the following Z transformation is employed

$$Z = \begin{cases} \frac{s-1}{\delta}, & \text{if } S > 0 \\ 0, & \text{if } S = 0 \\ \frac{s+1}{\delta}, & \text{if } S < 0 \end{cases} \qquad (4)$$

A positive (negative) value of Z indicates that the data tend to increase (decrease) with time. If the computed value of $|Z| >$ $Z_{1-(\alpha/2)}$, the null hypothesis (H0) is rejected at α level of significance in a two-sided test.

10   The Pettitt test is used to identify if there is a point change or jump in the data series (Pettitt, 1979). This method detects one unknown change point by considering a sequence of random variables X_1, X_2, …, X_T that may have a change point at N if Xt for t = 1, 2, …, N has a common distribution function F1(x) and Xt for t = N + 1, …, T has a common distribution function F2(x), and F1(x) ≠ F 2(x).

4.2 Land use/cover map analysis

4.2.1. Landsat image acquisition

Landsat images of the year 1973, 1985, 1995 and 2010 were accessed free-of-charge from the US Geological Survey (USGS) Center for Earth Resources Observation and Science (EROS) via http://glovis.usgs.gov. The Landsat image scenes were selected based on the criteria of acquisition period, availability and percentage of cloud cover.

According to the recommendation of (Hayes and Sader, 2001), images needed to be acquired for the same acquisition period, in order to reduce scene-to-scene variation due to sun angle, soil moisture, atmospheric condition and vegetation phenology differences. Hence, cloud free images were collected for the dry months of January and May. However, as the basin covers large area, each period of LULC map composed of 16 Landsat scenes, therefore, it was difficult to get all the scenes in a dry

25   season of a single year. Hence, images were acquired of ±1 year for each time period. For 1973, for example, 16 Landsat MSS image scenes were acquired in 1973 (±1 years) and merged to arrive at one LULC representation for selected years.

4.2.2 Pre-processing and processing images

Several standard pre-processing methods including geometric correction and radiometric correction were implemented to prepare the LULC maps from Landsat images. Even though there are many different classification methods, supervised and unsupervised classification are the two widely used methods for landcover classification from remote sensing images. In this
5   study, a hybrid supervised/unsupervised classification approach was carried out to classify the images from 2010 (LandsatTM). First, Iterative Self-Organizing Data Analysis (ISODATA) clustering was performed to determine the spectral classes or land cover classes of the image. Secondly, polygons for all of the training samples based on the identified LC classes were digitized using ground truth data and then the samples for each land cover type were aggregated. Finally, a supervised classification was performed using a maximum likelihood algorithm in order to extract four LULC classes.

A total of 488 ground truth data (GCPs Ground Control Points) regarding landcover types and their spatial locations were collected in March and April 2017 using a Global Positioning System (GPS), assuming that there had not been any significant landcover change between 2017 and 2010. Locations where ground truth data were taken, were selected carefully by interviewing local elderly people and using the prior knowledge of the first author. As many as 288 points were used for
15   accuracy assessment and 200 points were used for developing training sites to generate a signature for each land cover type. The accuracy of the classifications was assessed by computing the error matrix (also known as confusion matrix) that compares the classification result with ground truth information as suggested by DeFries and Chan (2000). A confusion matrix lists the values for known cover types of the reference data in the columns and for the classified data in the rows (Banko, 1998). From the confusion matrix four measures of accuracy are estimated such as overall accuracy, user's
20   accuracy, producer's accuracy and the kappa statistic.

Once the landcover classification of the year 2010 Landsat image is completed and its accuracy is checked, the NDVI differencing technique (Mancino *et al.*, 2014) was applied to classify the images of 1973, 1985 and 1995. This technique was chosen to increase the accuracy of classification as it is hard to find an accurately classified digital or analogue LULC map
25   of the study area during the period of 1973, 1985 and 1995 and also the information obtained from the elders are more subjective and its reliability is questionable. In order to increase the accuracy, we first calculated the NDVI from the Landsat MSS (1973) and three pre-processed Landsat TM images (1985, 1995, 2010) following the general normalized difference between band TM4 and band TM3 images eqn. 5. The resulting NDVI images for the time periods were subtracted to assess the ΔNDVI image with positive (vegetation increase), negative (vegetation cleared) and no change on a 30 m x 30 m pixel
30   resolution (eqn.6-8). The Landsat MSS 60m x 60m pixel size data sets were resampled to a 30 m x 30 m pixel size using the 'nearest neighbour' resampling technique in order to have similar pixel sizes for the different images without altering the original pixel values of the image data.

$$NDVI = \frac{(TM4 - TM3)}{(TM4 + TM3)} \text{ or } \frac{(MSS_3 - MSS_2)}{(MSS_3 + MSS_2)} \qquad (5)$$

$$\Delta NDVI_{1995/2010} = NDVI_{1995} - NDVI_{2010} \qquad (6)$$

$$\Delta NDVI_{1985/1995} = NDVI_{1985} - NDVI_{1995} \qquad (7)$$

$$\Delta NDVI_{1973/1985} = NDVI_{1973} - NDVI_{1985} \qquad (8)$$

5   The difference image $\Delta NDVI$ was then reclassified using a threshold value calculated as $\mu \pm n*\sigma$; where μ represents the $\Delta NDVI$ pixels value mean, and σ the standard deviation. The threshold identifies three ranges in the normal distribution: (a) the left tail ($\Delta NDVI < \mu - n*\sigma$); (b) the right tail ($\Delta NDVI > \mu + n*\sigma$); and (c) the central region of the normal distribution ($\mu - n*\sigma < \Delta NDVI < \mu + n*\sigma$). Pixels within the two tails of the distribution are characterized by significant landcover changes, while pixels in the central region represent no change.

The standard deviation (σ) is one of the most widely applied threshold identification approaches for different natural environments based on different remotely sensed imagery (Hu *et al.*, 2004; Jensen, 1996; Lu *et al.*, 2004; Mancino *et al.*, 2014; Singh, 1989) as cited by Mancino *et al.* (2014). To be more conservative n=1 was selected for this study to narrow the ranges of the threshold for reliable classification.

$\Delta NDVI$ pixel values (2010-1995) in the central region of the normal distribution ($\mu - n\cdot\sigma < \Delta NDVI < \mu + n\cdot\sigma$) represent an absence of landcover change between two different periods (i.e. 1995 and 2010), therefore, pixels of 1995 corresponding to no landcover change can be classified as similar to the 2010 landcover classes. Pixels with significant NDVI change are again classified using supervised classification, taking signatures from the already classified no change pixels. Likewise,

20   landcover classification of 1985 and 1973 images was performed based on the classified images of 1995 and 1985 respectively.

Finally, after classifying the raw images of Landsat into different landcover classes, change detection which requires the comparison of independently produced classified images (Singb, 1989) was performed by the post-classification method.

25   The post-classification change detection comparison was conducted to determine changes in LULC between two independently classified maps from images of two different dates. Although this technique has some limitations, it is the most common approach as it does not require data normalization between two dates (Singh, 1989) because data from two dates are separately classified, thereby minimizing the problem of normalizing for atmospheric and sensor differences between two dates.

30

4.3 SWAT hydrological model

The Soil and Water Assessment Tool (SWAT) is an open-source-code, physically-based, semi-distributed model with a large and growing number of model applications in a variety of studies ranging from catchment to continental scales (Allen *et al.*, 1998; Arnold *et al.*, 2012; Neitsch *et al.*, 2002). It evaluates the impact of LULC and climate change on water resources in a

5   basin with varying soil, land use and management practices over a set period of time (Arnold *et al.*, 2012).

In SWAT, the watershed is divided into multiple sub-basins, which are further subdivided into hydrological response units (HRUs) consisting of homogeneous land-use management, slope and soil characteristics (Arnold *et al.*, 2012; Arnold *et al.*, 1998). HRUs are the smallest units of the watershed in which relevant hydrologic components such as evapo-transpiration,

10   surface run-off and peak rate of run-off, groundwater flow and sediment yield can be estimated. Water balance is the driving force behind all the processes in the SWAT calculated using eqn.9.

$$SW_t = SW_o + \sum_{I=1}^{t}(R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}) \qquad (9)$$

15   where SW_t is the final soil water content (mm H_2O), SW_o is the initial soil water content on day i (mm H_2O), t is the time (days), R_{day} is the amount of precipitation on day i (mm H_2O), Q_{surf} is the amount of surface run-off on day i (mm H_2O), E_a is the amount of evapo-transpiration on day i (mm H_2O), W_{seep} is the amount of water entering the vadose zone from the soil profile on day i (mm H_2O), and Q_{gw} is the amount of return flow on day i (mm H_2O).

20   Run-off is calculated separately for each HRU and routed to obtain the total run-off for the watershed using the soil conservation service (SCS) curve number (CN) method (USDA, 1972). The CN method was used in this study because of its ability to use daily input data (Arnold *et al.*, 1998; Neitsch *et al.*, 2011; Setegn *et al.*, 2008). This study focused on the effects of LULC and climate change on the water balance of the basin, which includes the component of inflows, outflows and the change in storage. Precipitation is the main inflow, while evapo-transpiration (Et), surface run-off (Qs), lateral flow (Ql), and

25   base flow (Qb) are the outflows. SWAT has three storages, namely, soil moisture (SM), shallow aquifer (SA) and deep aquifer (DA). Water movement from the soil moisture storage to the shallow aquifer is due to percolation, whereas, water movement from the shallow aquifer reverse upward to the soil moisture storage is Revap. For a more detailed description of the SWAT model, reference is made to Neitsch *et al.* (2011).

30

The SWAT model setup and data preparation can be done using arcSWAT tools in the arcGIS environment, while parameter sensitivity analysis, model calibration and validation was performed using the SWAT-CUP (Calibration and Uncertainty Procedures) interface Sequential Uncertainty Fitting (SUFI-2) algorithm (Abbaspour, 2008). During model setup, the

observed monthly discharge of the given period was divided in to three separate data sets, the first to warm up the model, the second to calibrate the model and the third to validate the model.

The first step in SWAT is the determination of the most sensitive parameters for a given watershed using global sensitivity
5   analysis option (Arnold *et al.*, 2012). The second step is the calibration process adjusting the model input parameters necessary to match model output with observed data, thereby reducing the prediction uncertainty. Initial parameter estimates were taken from the default lower and upper bound values of the SWAT model database and from earlier studies in the basin e.g.(Gebremicael *et al.*, 2013). The final step, model validation involves running a model using parameters that were determined during the calibration process, and comparing the predictions to independent observed data not used in the
10 calibration.

In this study both manual and automatic calibration strategy was applied to attain the minimum differences between observed and simulated flows in terms of surface flow, peak and total flow following the steps recommended by Arnold *et al.* (2012). In this study, we divided the simulation periods of (1971-2010) in to four periods, namely the 1970s, 1980s,
15 1990s and 2000s, as shown in Table 2, in order to analyse the combined and isolated impacts of LULC and climate changes for the basin. The models performance for the streamflow were then evaluated using statistical methods (Moriasi *et al.*, 2007) such as the Nash–Sutcliffe coefficient of efficiency (NSE), the coefficient of determination (R^2) and the relative volume error (RVE %), which are shown by eqn.10-12. Furthermore, graphical comparisons of the simulated and observed data, as well as water balance checks were used to evaluate the model's performance.

$$R^2 = \frac{[\sum(Q_{m,i}-\bar{Q}_m)(Q_{s,i}-\bar{Q}_s)]^2}{\sum(Q_{m,i}-\bar{Q}_m)^2 \sum {Q_{s,i}-\bar{Q}_s}^2}$$ (10)

$$NSE = 1 - \frac{\sum(Q_m-Q_s)_i^2}{(Q_{m,i}-\bar{Q}_m)^2}$$ (11)

$$RVE\ (\%) = 100 * \frac{\sum_{i=1}^{n}(Q_m-Q_s)_i}{\sum_{i=1}^{n}Q_{m,i}}$$ (12)

where $Q_{m,i}$ is the measured flow data in $m^3 s^{-1}$, Q_m is the mean n values of the measured data, $Q_{s,i}$ is the simulated flow data in$m^3 s^{-1}$, and Q_s is the mean n values of simulated data.

4.4 SWAT simulations

In this study, three different approaches were used for SWAT simulation aimed at assessing the individual and combined effects of LULC and climate change on streamflow and water balance components. The first approach included simulations to attribute changes in streamflow to combined LULC and climate change, the simulation results represented "real runoff"

affected by the combination of LULC and climate changes. We divided the simulation period (1971–2010) into four equal periods to accommodate gradual changes in land use. The first period (1971–1980) was regarded as the baseline period and the other periods (1981–1990, 1991–2000, and 2001–2010) were regarded as altered periods. LULC maps of 1973, 1985, 1995 and 2010 produced from Landsat images were used to represent the LULC patterns of 1970s (1971-1980), 1980s (1981-1990), 1990s (1991-2000) and 2000s (2001-2010), respectively. The LULC and climate data for each model ran were in the same period (i.e 1970s, 1980s, 1990s and 2000s) see Table 2. The SWAT model has been calibrated/validated and optimal parameter values for the 6 most sensitive parameters has been fixed for each four periods of model run. Once the SWAT models are calibrated and validated for the four-decadal time period, streamflow and water balance components are simulated. This approach enables us to investigate how the parameter values alter following the combined LULC and climate changes and also help us analyse the combined impacts of LULC and climate change on streamflow and water balance components

The second approach included simulations to attribute only for LULC changes, aimed to investigate whether LULC change is the main driver for changes in water balance components. The method assumes that applying the calibrated model parameters values for the altered periods using updated LULC maps will enable the simulation of impacts of LULC change on mean annual streamflow and water balance components (Hassaballah et al.; Marhaento et al., 2017). We ran the calibrated SWAT model for the baseline period (1970s) using a LULC map of the year 1973 and a precipitation data set of the 1970s, then we applied the calibrated model parameters for the altered periods using updated LULC maps but with the same precipitation data(Table 2) in order to investigate the attribution of changes to mean annual water balance component ratios, as a result of LULC change. In this case, the values of those 6 most sensitive parameters remained constant for both baseline and altered periods of SWAT simulation processes

The third approach is similar to the second approach but the simulations are attributed only for climate changes. A model was run again four times, corresponding to the LULC periods using a unique LULC map of the year 1973 and its parameter values but four different periods of climate data sets (1970s, 1980s, 1990s and 2000s).

5. Results and discussions

5.1 Trend test

5.1.1 Rainfall

The summary results of the MK tests for the 15 selected stations located inside and around the UBNRB revealed a mixed trend (increasing, decreasing and no change) for the rainfall over the UBNRB. For daily time series, the computed

probability values (P-value) for seven stations was greater while for eight stations it was less than the given significance level (α=5%), which means that no statistically significant trends existed in seven stations, while in eight other a monotonic trend was displayed. On a monthly basis, the p-value for all 15 stations was larger than the given significance level, which showed that no statistically significant trend existed in every station. On an annual time scale, 11 stations showed a p-value

5   larger than the significance level, whereas four stations (Alemketema, Debiremarkos, Gimijabet and Shambu) exhibited a p-value less than the significance level (all showed an upward trend except Alemketema). This result tallies well with earlier studies in the basin at station level such as that of (Gebremicael *et al.*, 2013) who analysed for nine stations of UBNRB on an annual basis and the result of eight stations were similar, except for the Debire markos station.

10   The basin wide rainfall trend analysis was again carried out at daily, monthly, seasonal and annual time scale as computed by the MK and Pettitt tests as summarized in Table 3 and Figure 3 respectively. Both MK test and Pettitt test, indicate that there was no statistically significant trend change at basin-wide rainfall at monthly, annual and seasonal time scales after applying TFPW, however, at daily time scale, Pettitt test showed jump point with increasing trend. This result is in line with the earlier studies in the basin such as (Conway, 2000; Gebremicael *et al.*, 2013; Tesemma *et al.*, 2010). Those studies

15   reported that there was no significant change of annual and seasonal rainfall over the Upper Blue Nile.

5.1.2 Streamflow

The trend analysis of daily, monthly, seasonal and annual streamflow was computed by the MK and Pettitt tests summarized in Table 3 after applying TFPW. The flow of El Diem station at daily, annual and long rainy season time series showed a

20   significant increase over the last 40 years while the mean monthly streamflow at El Diem did not show any clear pattern. This result agreed with the study carried out by Gebremicael *et al.* (2013), which reported an increase in the observed annual flow at the El Diem, Kessi and Bahirdar sub-basins but disagreed with the result of (Tesemma *et al.*, 2010), who reported that there has been no significant pattern in the observed annual flow at El Diem. Since the seasonal and annual rainfall over the basin during the 1971–2010 period did not show any significant changes, the increasing annual flow of the UBNRB at El

25   Diem could be attributed to an LULC change within the basin over the last 40 years (1971-2010).

5.2 LULC change analysis

The confusion matrix is used to measure the accuracy of the classified images by comparing spatially coincident ground control points and pixels of the classified image. It was established using 288 ground control points (GCPs) which are not

30   used in the classification of the 2010 image. According to the confusion matrix report, 80% overall accuracy, producer's accuracy values for all classes ranged from 75.4% to 100 %, user's accuracy values ranged from 83.7% to 91.7% and the

kappa coefficient (k) of 0.77 were attained for the 2010 classified image as shown in Table 5. Monserud (1990) as cited by Rientjes *et al.* (2011) suggested a kappa value of <40% as poor, 40–55% fair, 55–70% good, 70–85% very good and >85% as excellent. According to these ranges, the classification in this study has very good agreement with the validation data set and met the minimum accuracy requirements to be used for the change detection.

The classified images of the basin have shown different LULC proportion at four different time periods as shown in Figure 5. In 1973, the UBNRB was dominated by cultivated land (62.9%), followed by bushes & shrubs (18%), forest (17.4%), and water (1.74%). In 1985, the cultivated land increased (to 65.6%), followed by bushes & shrubs (18.3%), while forest decreased (to 14.4%), and water remained unchanged (at 1.7%). In 1995, cultivated land further increased to (67.5%),

10 followed by bushes & shrubs (18.5%), forest further decreased (to 12.2%), and water remained unchanged (1.7%). In 2010, cultivated land decreased (to 63.9%), bushes and shrubs increased to 18.8 %, forest increased to 15.6 % and water remained unchanged at 1.7%. During the entire 1973–2010 period, cultivated land, along with bushes & shrubs remained the major proportions as compared to the other LULC classes. The highest gain (2.7%) and the largest loss (-3.6%) in cultivated land occurred during the 1973–1985 and 1995-2010 periods respectively. The highest gain in bushes and shrubs was (0.3%) from

15 1973 to 1985, while the highest gain in forest coverage (3.4%) was recorded during the period 1995–2010. Water coverage remained unchanged from 1973 to 2010.

The increased forest coverage and the reduction in cultivated land over the period 1995 to 2010 shows that the environment was recovering from the devastating drought and forest clearing for firewood and cultivation due to population growth. This

20 could be due to the afforestation programme initiated by the Ethiopian government. During the period from 1995 to 2010 eucalyptus tree plantation expanded significantly across the country. To summarize, during the period from 1973 to 2010, forest coverage declined by 1.8%, with both bushes and shrubs, as well as cultivated land increasing by 0.8% and 1% respectively from the original 1973 level. This result agrees well with other local level studies (Gebremicael *et al.*, 2013; Rientjes *et al.*, 2011; Teferi *et al.*, 2013), which reported the dramatic changes in the natural vegetation cover resulting from

25 the agricultural land.

5.3 SWAT model calibration and validation

The most sensitive parameters of the SWAT model to simulate streamflow were identified using global sensitivity analysis of SWAT-CUP and their optimized values were determined by the calibration process recommended by Arnold *et al.* (2012). Parameters such as SCS curve number (CN2), base flow alpha factor (ALPHA_BF), soil evaporation compensation factor

30 (ESCO), threshold water depth in the shallow aquifer required for return flow to occur (GWQMN), groundwater "revap" coefficient (GW_REVAP) and the available water capacity (SOL_AWC) were found to be the most sensitive parameters for the flow predictions.

Figure 6 showed the calibration and the validation results of monthly streamflow hydrographs and this result revealed the model well captured the monthly hydrographs. This was again verified by the statistical performance measures of R^2, NSE and RVE (%) as presented in Table 6. For the calibration period, the values of R^2, NSE and RVE (%) from the four model is ranged from 0.78 to 0.91, 0.73 to 0.91 and 0.7% to 4%, and for the validation period it ranged from 0.84 to 0.94, 0.84 to 0.92 and -7.5% to 7.4% respectively. According to the rating of Moriasi *et al.* (2007), the performance of the SWAT model over UBNRB can be categorized as very good, although underestimation was observed in the base flow simulation.

The optimal parameter values of the calibrated model for the four model runs are shown in Table 7. A change was obtained for CN2 parameter values, such change indicates changes of the catchment response behaviour. For instance, an increase in the CN2 value in the 1980s and 1990s from 0.89 to 0.91 and 0.92 as compared to 1970s respectively, indicate a reduction in forest coverage and expansion of cultivated land which align with the result of LULC classification maps. In contrary, a decrease in CN2 value was observed during the period 1990s to 2000s from 0.92 to 0.9, attributed to the increase in forest coverage and reduction in cultivated land.

5.4 Effects of combined LULC and climate change on streamflow and water balance components

The simulation of the four SWAT model run result indicates that streamflow and water balance is changed due to the combined effect of both LULC and climate change during the last 40 years' time. At basin level, the MK trend analysis showed increasing trend for the mean annual and long rainy season streamflow but no trend for short rainy season streamflow at 5% significance level. This was also confirmed by SWAT simulation result. The mean annual streamflow and water balance components simulated by SWAT for the four-simulation period is shown in Table 8. In general, mean annual streamflow increased by 15.6% between the period 1970s and the 2000s. However, mean annual streamflow changed differently in different decades. For example, it increased by 2.1%, 6.8% and 6% during the period 1980s, 1990s and 2000s respectively from the baseline period 1970s.

The ratio of mean annual streamflow to mean annual precipitation (Qt/P) increased from 19.4% to 22.1%, and actual evaporation to precipitation (Ea/P) decreased from 61.1% to 60.5% during 2000s from the 1970s. Moreover, ratio of surface run-off to streamflow (Qs/Qt) has significantly increased from 40.7% to 55.4% while base flow to streamflow (Qb/Qt) has significantly decreased from 17.1% to 10.3% and 3.2% respectively during the period 1980s and 1990s. This result can be attributed to the dimension of the combined effect of the LULC and climate change on the streamflow. 1990s was the period when the highest deforestation and expansion of cultivated land was reported; meanwhile the rainfall intensity and number of rainfall events has significantly increased compared to the 1970s and 1980s, as shown in Table 4. However, it is difficult to identify which variable (the LULC or the climate) most significantly affects the streamflow and its water balance

components at this stage. Consequently, further analysis was carried out, by changing the LULC and holding climate data constant and vice versa. The results of this simulation are discussed in sections 5.5 and 5.6 below.

5.5 Effects of a single change in LULC on streamflow and water balance components

Once the SWAT model had been calibrated and validated for the baseline period, the SWAT model again ran four times for the baseline period and for three altered periods using updated LULC maps. Firstly, with the LULC map of 1973; secondly with LULC map of 1985; thirdly with LULC map of 1995; and fourthly with LULC map of 2010. Then the outputs from the four different LULCs were compared. We note that the climate data for the period 1973-1980 and calibrated parameter values for the 6 sensitive parameters remained constant while the LULC was changed for all four models to identify hydrological impacts of changes in LULC explicitly as suggested by (Hassaballah *et al.*). The Qs/Qt ratio changed from 40.7% to 47.7%, 53.1% and 39% respectively by using the LULC maps from 1973, 1985, 1995 and 2010 whereas the Qb/Qt ratio changed from 17.1% to 10.1%, 3.3% and 23.4% respectively. The highest Qs/Qt ratio (53.1%) and the lowest Qb/Qt ratio (3.3%) was recorded with the LULC map of 1995. This could be attributed to the 5.1% reduction in forest coverage and 4.6% increase in cultivated land with the 1995 LULC map as compared to the 1973 LULC map. This deforestation may cause a reduction in canopy interception and plant transpiration which ultimately reduce evapo-transpiration. In the other hand, expansion of cultivated land and reduction in forest coverage affects the properties of top soil that cause a lower permeability and less infiltration as a result fraction of precipitation converted to surface run-off is increasing while the fraction of base flow is getting reduced. Based on the SWAT model result, this study provides a strong indication that changes in LULC altered the water balance in the UBNR basin. Findings show that LULC change due to deforestation and expansions of cultivated area has increased surface run-off but reduced base flow.

5.6 Effects of single climate change on streamflow and water balance components

The impacts of climate change are analysed by running the four models using a unique LULC map of 1973 with its model parameters but changing the four different data sets of precipitation (1970s, 1980s, 1990s and 2000s). The simulated water balance components shown in Figure 7, indicate that the Qs/Qt ratio increased from 40.7% to 45.2%, 45.6% and 46.2% during the period 1970s, 1980s, 1990s and 2000s respectively, while, the Qb/Qt ratio changed from 17.1% to 13.5%, 14.9% and 12.7% during the same simulation period. The highest surface run-off fraction and lowest base flow fraction was recorded with climate data of 2000s. The increasing of surface run-off and decreasing of base flow during the simulation period in this study is attributed to increasing of rainfall intensity and extreme rainfall events in the UBNRB as can be seen in

Table 4. The 99-percentile precipitation increased from 17.3 mm to 19.6 mm and R20mm increased from 15 days to 35 days during the period from 1970s to 2000s.

6. Conclusions

The objectives of this study were to understand the long-term variations of climate and hydrology of the UBNRB using statistical techniques (MK and Pettitt tests), and to assess the combined and single effects of climate and LULC change using a semi-distributed hydrological model (SWAT). The MK and Pettitt tests showed no statistically significant change of the annual and seasonal rainfall over the UBNRB between 1971 and 2010. However, both tests showed a statistically significant increasing trend of streamflow for annual, long and short rainy season but no trend during the dry season. The LULC change detection was assessed by comparing the classified images and the result showed that the dominant process is largely the expansion of cultivated land and decrease in forest coverage. The rate of deforestation is high during the period 1973-1995, this is probably due to severe droughts occurred in 1984/85, large population increase as a result expansion of agricultural land. On the other hand, forest coverage increased by 3.4% during the period 1995 to 2010. This indicates that the environment was recovering from the devastating drought and forest clearing as the result of afforestation programme initiated by the Ethiopian government. During the period from 1995 to 2010 the planting of multipurpose eucalyptus trees expanded significantly across the country in order to generate income, as well as to produce fire wood, charcoal and construction materials.

The SWAT model was used to simulate the combined and single effects of LULC and climate changes on the monthly streamflow at the basin outlet (El Diem station, located on the Ethiopia-Sudan border). The result showed that the combined effects of the LULC and climate changes increased the mean annual streamflow by 15.5% from the 1970s to the 2000s. The high reduction in forest coverage and expansion of cultivated land during the 1973 to 1995 period caused a larger fraction of rainfall to be transformed to surface run-off and led to a reduction in the ratio of base flow. Similarly, the increase in rainfall intensity and extreme precipitation events led to a substantial increase in Qs/Qt and a substantial decrease in Qb/Qt and ultimately increases in the streamflow during the 1971-2010 simulation period. The smaller contribution of LULC change may be due to the fact that the SWAT model does not adjust CN2 for slope, which might be significant in areas where the majority of the area has a slope greater than 5%, such as the UBNRB.

The combined results from three different approaches, namely statistical trend test, semi-distributed SWAT modelling and LULC change analysis, are consistent with the hypothesis that LULC change has modified the run-off generation process, which has caused the increase in streamflow of the UBNRB while the climate has remained unchanged. These findings can be useful for basin-wide water resources management in the Blue Nile basin, as it provides a better understanding of the trends of rainfall, and streamflow, as well as the combined and single effects of climate and LULC change on the streamflow for the UBNRB. Hence, protecting and conserving the natural forests is highly recommended, not only for maintaining the

streamflow but also reducing soil erosion because soil erosion is a function of surface run-off, which further increases the productivity, livelihoods and regional water resource use cooperation. The limitation of this study could be due to the uncertainty of the SWAT model, as the SWAT model does not adjust CN2 for slopes greater than 5%, which could be significant in areas where the majority of the area has a slope greater than 5%, such as UBNRB. Therefore, we suggest

5   adjusting the CN2 values for slope > 5% outside of the SWAT model for further research. Finally, the authors would like to point out that the impacts of current and future water resource developments should be investigated in order to establish comprehensive and holistic water resource management in the Nile basin.

References

Abbaspour, C.K., 2008. SWAT Calibrating and Uncertainty Programs. A User Manual. Eawag Zurich, Switzerland.

10 Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300(9): D05109.

Arnold, J.G. *et al.*, 2012. SWAT: Model use, calibration, and validation. Transactions of the ASABE, 55(4): 1491-1508.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment
15   part I: Model development1. Wiley Online Library.

Banko, G., 1998. A review of assessing the accuracy of classifications of remotely sensed data and of methods including remote sensing data in forest inventory.

BCEOM, 1998. Abbay river basin integrated development master plan project.

Bewket, W., Sterk, G., 2005. Dynamics in land cover and its effect on stream flow in the Chemoga watershed,
20   Blue Nile basin, Ethiopia. Hydrological Processes, 19(2): 445-458.

Cheung, W.H., Senay, G.B., Singh, A., 2008. Trends and spatial distribution of annual and seasonal rainfall in Ethiopia. International Journal of Climatology, 28(13): 1723-1734.

Conway, D., 2000. The climate and hydrology of the Upper Blue Nile River. The Geographical Journal, 166(1): 49-62.

25 DeFries, R., Chan, J.C.-W., 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sensing of Environment, 74(3): 503-515.

Gebrekristos, S.T., 2015. Understanding Catchment Processes and Hydrological Modelling in the Abay/Upper Blue Nile Basin, Ethiopia, TU Delft, Delft University of Technology.

Gebremicael, T., Mohamed, Y., Betrie, G., van der Zaag, P., Teferi, E., 2013. Trend analysis of runoff and
30   sediment fluxes in the Upper Blue Nile basin: A combined analysis of statistical tests, physically-based models and landuse maps. Journal of Hydrology, 482: 57-68.

Gilbert, R.O., 1987. Statistical methods for environmental pollution monitoring. John Wiley & Sons.

Hamed, K.H., Rao, A.R., 1998. A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology, 204(1-4): 182-196.

35 Hassaballah, K., Mohamed, Y., Uhlenbrook, S., Biro, K., Analysis of streamflow response to land use land cover changes using satellite data and hydrological modelling: case study of Dinder and Rahad tributaries of the Blue Nile (Ethiopia/Sudan).

Hayes, D.J., Sader, S.A., 2001. Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series. Photogrammetric engineering and remote sensing, 67(9): 1067-1075.

Hu, Y., De Jong, S., Sluiter, R., 2004. A modeling-based threshold approach to derive change/no change information over vegetation area, Proceedings of the "12 International Conference on Geoinformatics-Geospatial Information Research: Bridging the Pacific and Atlantic". University of Gävle (Sweden), pp. 647-654.

Jensen, J.R., 1996. Introductory digital image processing: a remote sensing perspective. Prentice-Hall Inc.

Kendall, M., 1975. Rank correlation methods.

Lu, D., Mausel, P., Batistella, M., Moran, E., 2004. Comparison of land-cover classification methods in the Brazilian Amazon Basin. Photogrammetric engineering & remote sensing, 70(6): 723-731.

Mancino, G., Nolè, A., Ripullone, F., Ferrara, A., 2014. Landsat TM imagery and NDVI differencing to detect vegetation change: assessing natural forest expansion in Basilicata, southern Italy. iForest-Biogeosciences and Forestry, 7(2): 75.

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometrica, 13(3): 245-259.

Marhaento, H., Booij, M.J., Rientjes, T., Hoekstra, A.Y., 2017. Attribution of changes in the water balance of a tropical catchment to land use change using the SWAT model. Hydrological Processes, 31(11): 2029-2040.

Mekonnen, D.F., Disse, M., 2016. Analyzing the future climate change of Upper Blue Nile River Basin (UBNRB) using statistical down scaling techniques. HESSD.

Melesse, A., Abtew, W., Dessalegne, T., Wang, X., 2009. Low and high flow analyses and wavelet application for characterization of the Blue Nile River system. Hydrological processes, 24(3): 241.

Mengistu, D., Bewket, W., Lal, R., 2014. Recent spatiotemporal temperature and rainfall variability and trends over the Upper Blue Nile River Basin, Ethiopia. International Journal of Climatology, 34(7): 2278-2292.

Monserud, R.A., 1990. Methods for comparing global vegetation maps.

Moriasi, D.N. *et al.*, 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. Asabe, 50(3): 885-900.

Neitsch, S., Arnold, J., Kiniry, J.e.a., Srinivasan, R., Williams, J., 2002. Soil and water assessment tool user's manual version 2000. GSWRL report, 202(02-06).

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009, Texas Water Resources Institute.

NMA, 2013. Annual climate buletien for the year 2013.

Pettitt, A., 1979. A non-parametric approach to the change-point problem. Applied statistics: 126-135.

Rientjes, T. *et al.*, 2011. Changes in land cover, rainfall and stream flow in Upper Gilgel Abbay catchment, Blue Nile basin-Ethiopia. Hydrology and Earth System Sciences, 15(6): 1979.

Seleshi, Y., Zanke, U., 2004. Recent changes in rainfall and rainy days in Ethiopia. International journal of climatology, 24(8): 973-983.

Setegn, S.G., Srinivasan, R., Dargahi, B., 2008. Hydrological modelling in the Lake Tana Basin, Ethiopia using SWAT model. The Open Hydrology Journal, 2(1).

Singb, A., 1989. Digital change detection techniques using remotelyWsensed data. International Journal of Remote Sensing, 10(6): 989L1003.

Singh, A., 1989. Review article digital change detection techniques using remotely-sensed data. International journal of remote sensing, 10(6): 989-1003.

Teferi, E., Bewket, W., Uhlenbrook, S., Wenninger, J., 2013. Understanding recent land use and land cover dynamics in the source region of the Upper Blue Nile, Ethiopia: Spatially explicit statistical modeling of systematic transitions. Agriculture, ecosystems & environment, 165: 98-117.

Tekleab, S., Mohamed, Y., Uhlenbrook, S., Wenninger, J., 2014. Hydrologic responses to land cover change: the case of Jedeb mesoscale catchment, Abay/Upper Blue Nile basin, Ethiopia. Hydrological Processes, 28(20): 5149-5161.

Tesemma, Z.K., Mohamed, Y.A., Steenhuis, T.S., 2010. Trends in rainfall and runoff in the Blue Nile Basin: 1964–2003. Hydrological processes, 24(25): 3747-3758.

USDA, 1972. SCS national engineering handbook, section 4: hydrology. The Service.

von Storch, H., 1995. Misuses of Statistical Analysis in Climate Research, Analysis of Climate Variability. Springer, pp. 11-26.

Yin, J., He, F., Xiong, Y.J., Qiu, G.Y., 2017. Effects of land use/land cover and climate changes on surface runoff in a semi-humid and semi-arid transition zone in northwest China. Hydrology and Earth System Sciences, 21(1): 183-196.

Yue, S., Pilon, P., Phinney, B., Cavadias, G., 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrological Processes, 16(9): 1807-1829.

Yue, S., Wang, C., 2004. The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. Water Resources Management, 18(3): 201-218.

Table 1: Areal long term (1971-2010) mean annual and seasonal rainfall and streamflow of UBNRB

| Station | Amount | | | | Contribution (%) | | | | |
	Kiremit	Belg	Bega	Total	Kiremit	Belg	Bega	Mean	Area (km^2)
Flow (m^{3}s^{-1})	3506.3	300.4	1018.4	4825.1	72.7	6.2	21.1	1608	172,254
Flow (BCM)	36.4	3.1	10.6	50.7					
Rainfall (mm)	1070.1	140.8	238.9	1449.8	73.8	9.7	16.5		

Kiremit: long rainy season, Belg: Short rainy season, Bega: Dry season

5 Table 2: Data sets for the baseline and altered periods for the SWAT simulation used to analyse the combined and single effect of LULC and climate changes on streamflow and water balance components

| Model run No. | Combined effect | | Isolated LULC change effect | | Isolated climate change effect | | |
	Climate data set	LULC map	Climate data set	LULC map	Climate data set	LULC map	Remark
1	1970s	1973	1970s	1973	1970s	1973	Base period
2	1980s	1985	1970s	1985	1980s	1973	altered period1
3	1990s	1995	1970s	1995	1990s	1973	altered period2
4	2000s	2010	1970s	2010	2000s	1973	altered period3

Table 3: MK and Pettitt test for the rainfall and streamflow of UBNRB after TFPW at different time scale

| Time scale | Rainfall | | | | Flow | | | |
	P	Sen's slope	r1	Pettitt test	P	Sen's slope	r1	Pettitt test
Daily	0.88	0.000	-0.24	Increase	< 0.0001	0.009	-0.264	Increasing
Monthly	0.148	0.077	0.45	No change	0.736	0.042	0.455	No change
Annually	0.348	1.432	0.003	No change	0.031	9.335	-0.039	Increasing
Kiremit	0.427	1.249	-0.01	No change	0.05	19.466	0.001	Increasing
Belg	0.792	-0.209	0.012	No change	0.289	1.239	0.003	Increasing
Bega	0.631	0.332	0.039	No change	0.219	4.746	0.051	No change

r1:lag1 auto correlation coefficient, p: probability at 5% significance level

Table 4: Summary of precipitation indices of the UBNRB at decadal time series

Indices	1970s	1980s	1990s	2000s
Mean (mm)	4.17	4.05	4.42	4.16
95 percentile (mm)	12.57	12.52	13.66	13.31
99 percentile (mm)	17.34	17.77	19.44	19.65
1-day max (mm)	27.15	25.67	32.24	32.38
R20mm (days)	16	15	30	35
SDII (mm)	7.22	7.38	7.66	7.77

SDII is the ratio of total precipitation to R1mm precipitation.

Table 5: Confusion (error) matrix for the 2010 land use/cover classification map

LULC class	Water	Forest	Cultivated	Bushes and Shrubs	Row total	Producers' accuracy
Water	44	0	0	0	44	100
Forest	1	46	6	8	61	75.4
Cultivated land	2	3	77	15	97	79.4
Bushes and shrubs	1	3	9	73	86	84.9
Column total	48	52	92	86	288	
User's accuracy	91.7	88.5	83.7	84.9		
Over all accuracy	0.8					
Kappa	0.77					

Table 6: Statistical performance measure values of the SWAT model

Period		R2	NSE	RVE(%)
	Calibration (1973-1977)	0.81	0.77	-2.4
1970s	Validation (1978-1980)	0.82	0.89	10.5
	Calibration (1983-1987)	0.78	0.74	0.9
1980s	Validation (1988-19909	0.83	0.84	-6.8
	Calibration (1993-1997)	0.91	0.91	-0.85
1990s	Validation (1998-2000)	0.87	0.88	-6.6
	Calibration (2003-2007)	0.86	0.86	4.2
2000s	Validation (2008-2010)	0.87	0.86	-18.0

Table 7: SWAT sensitive model parameters and their (final) calibrated values for the four model runs.

Parameter	Optimum value			
	1970s	1980s	1990s	2000s
R-CN2	0.88	0.91	0.92	0.9
a-Alpha-BF	0.028	0.028	0.028	0.028
V-GW_REVAPMN	0.7	0.45	0.7	0.34
V-GWQMN	750	750	750	750
V-REVAPMN	550	450	425	550
a-ESCO	-0.85	-0.85	-0.85	-0.85
R-SOL_AWC	6.5	6.5	6.5	6.5

R- value from the SWAT database is multiplied by a given value, V- Replace the initial parameter by the given value, a- Adding the given value to initial parameter value.

5   Table 8: Water balance analysis in the Upper Blue Nile River Basin (mm/year) by considering LULC and climate change over respective periods. All streamflow estimates are for El Diem station.

Water balance component	1970s	1980s	1990s	2000s
Surface flow (Qs)	100.2	143.4	168.6	141.4
Lateral flow (Ql)	117.3	113.35	125.9	117.6
Base flow (Qb)	56.42	29.6	9.8	64.7
Revap	270.8	257.2	310.6	241
PET	1615.1	1627.3	1614.7	1732.9
Ea	871.9	852.6	904.3	885
Precipitation (P)	1428.1	1397.1	1522.2	1462.5
Qt	273.96	286.3	304.3	323.7
Qs/Qt(%)	36.6	50.1	55.4	43.7
Qb/Qt(%)	20.6	10.3	3.2	20.0
Ea/P(%)	61.1	61.0	59.4	60.5
Qt/P(%)	19.2	20.5	20.0	22.1
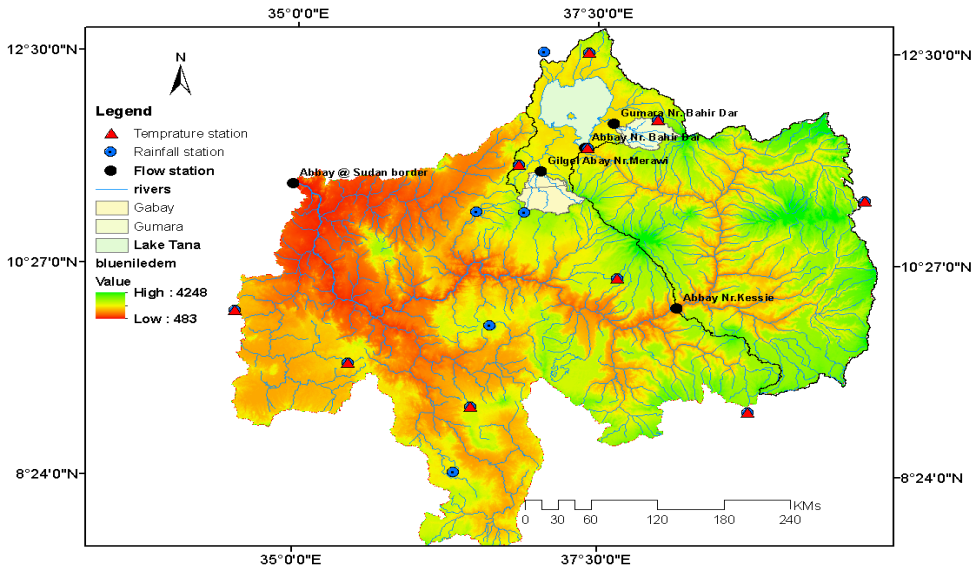
Figure 1 : Locations of study area and meteorological and discharge stations, with the Digital Elevation Model (DEM) data as the background
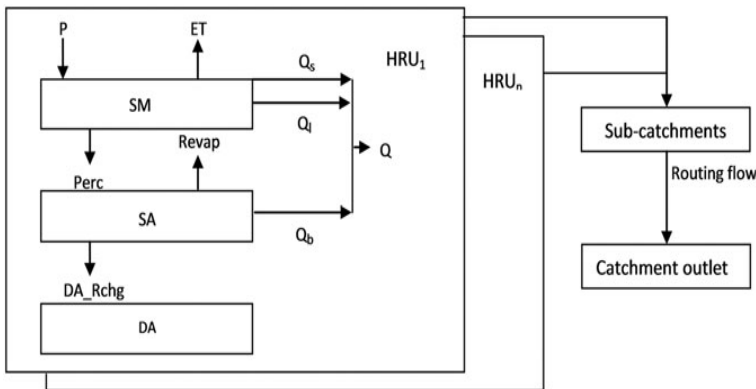
Figure 2: Schematic representation of the SWAT model structure from (Marhaento *et al.*, 2017)
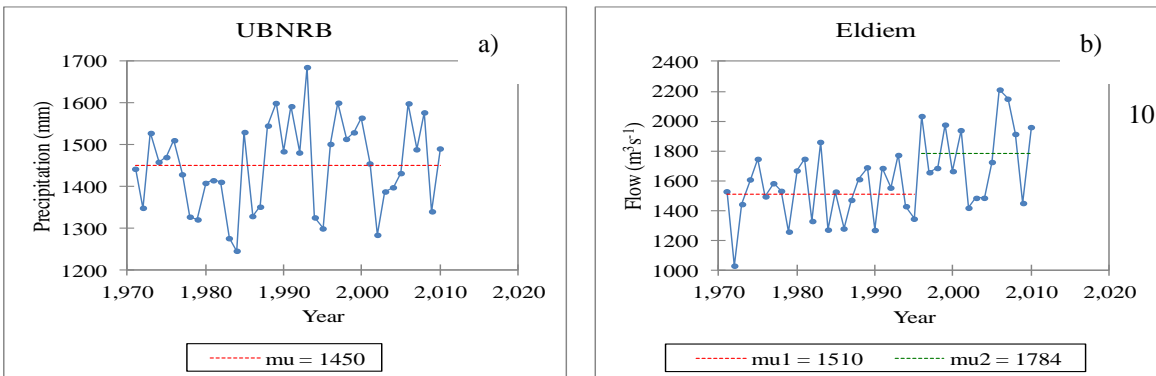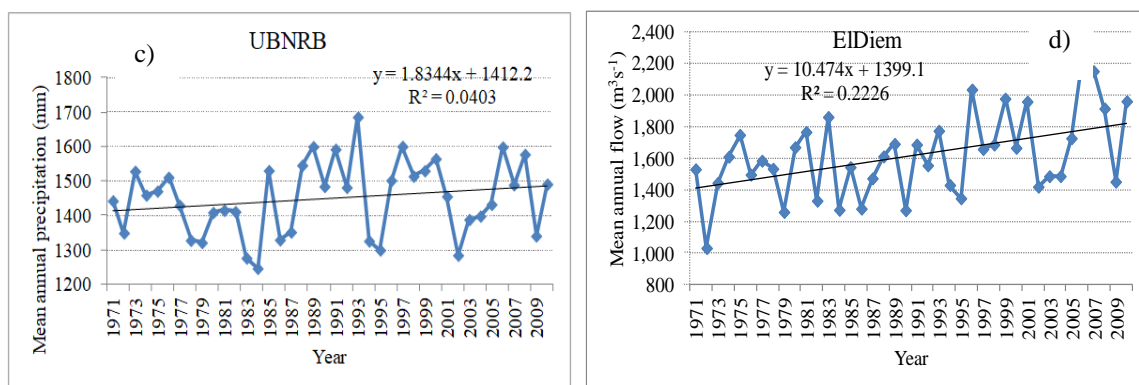
Figure 3: The Pettitt homogeneity test a) annual rainfall, b) annual flow of the UBNRB, C) linear trend of mean annual rainfall and d) linear trend of mean annual streamflow
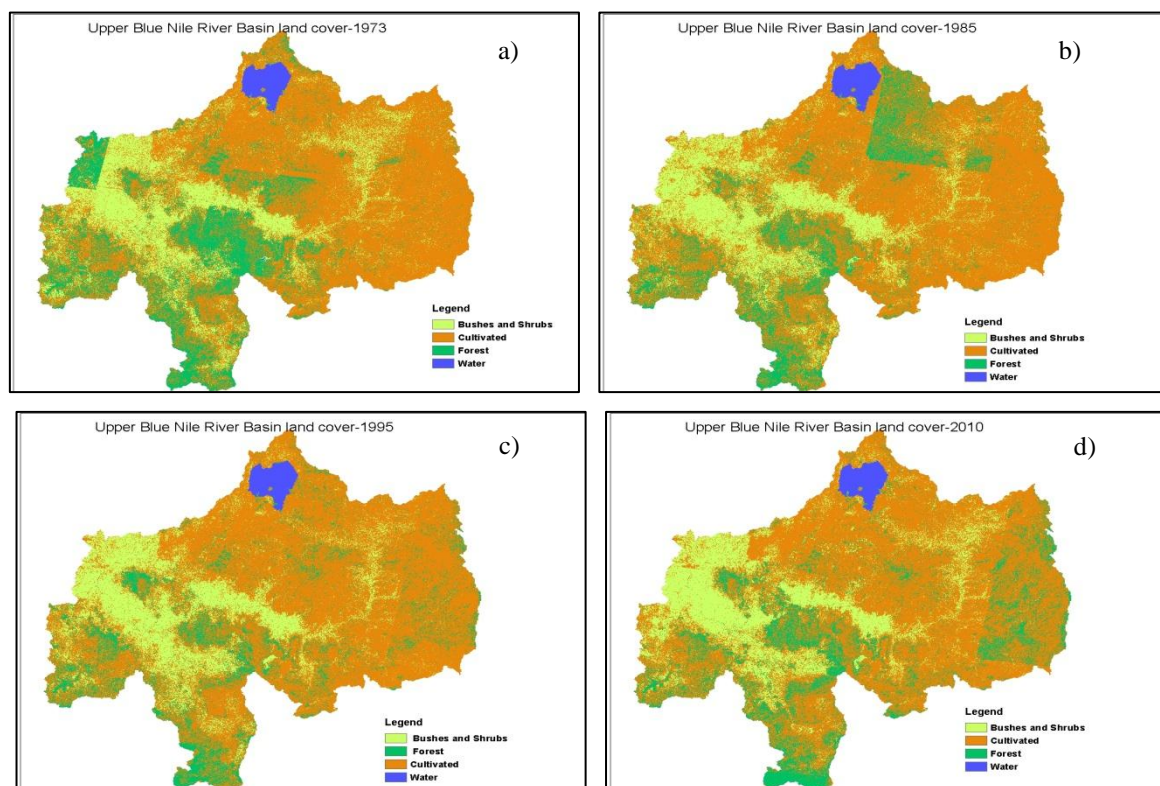
Figure 4: Landcover map of UBNRB derived from Landsat images a) 1973, b) 1985, c) 1995 and d) 2010

a)   b)
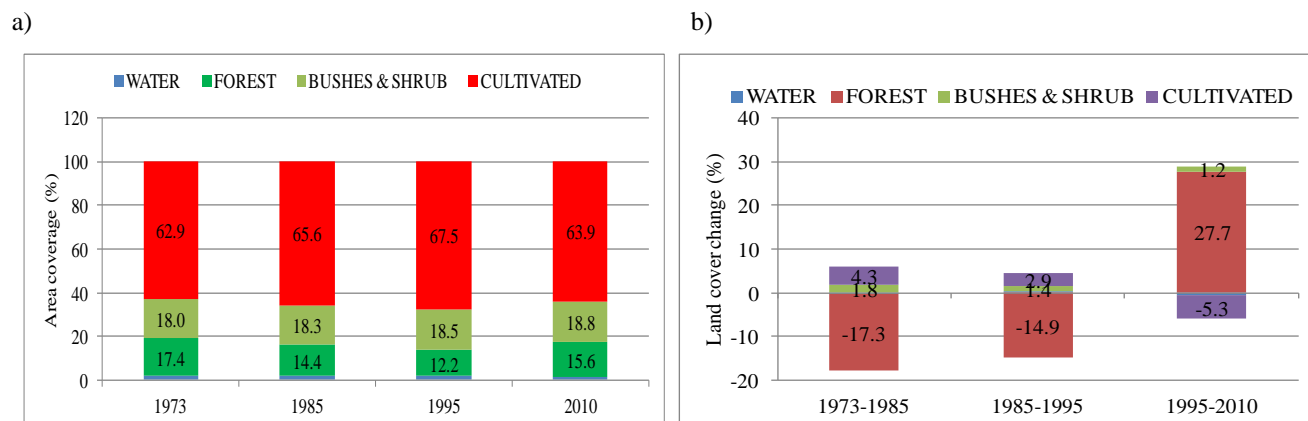
Figure 5: a) LULC composition, b) LULC change in the UBNRB during the period from 1973 to 2010

5   a)

b)

10   c)

d)
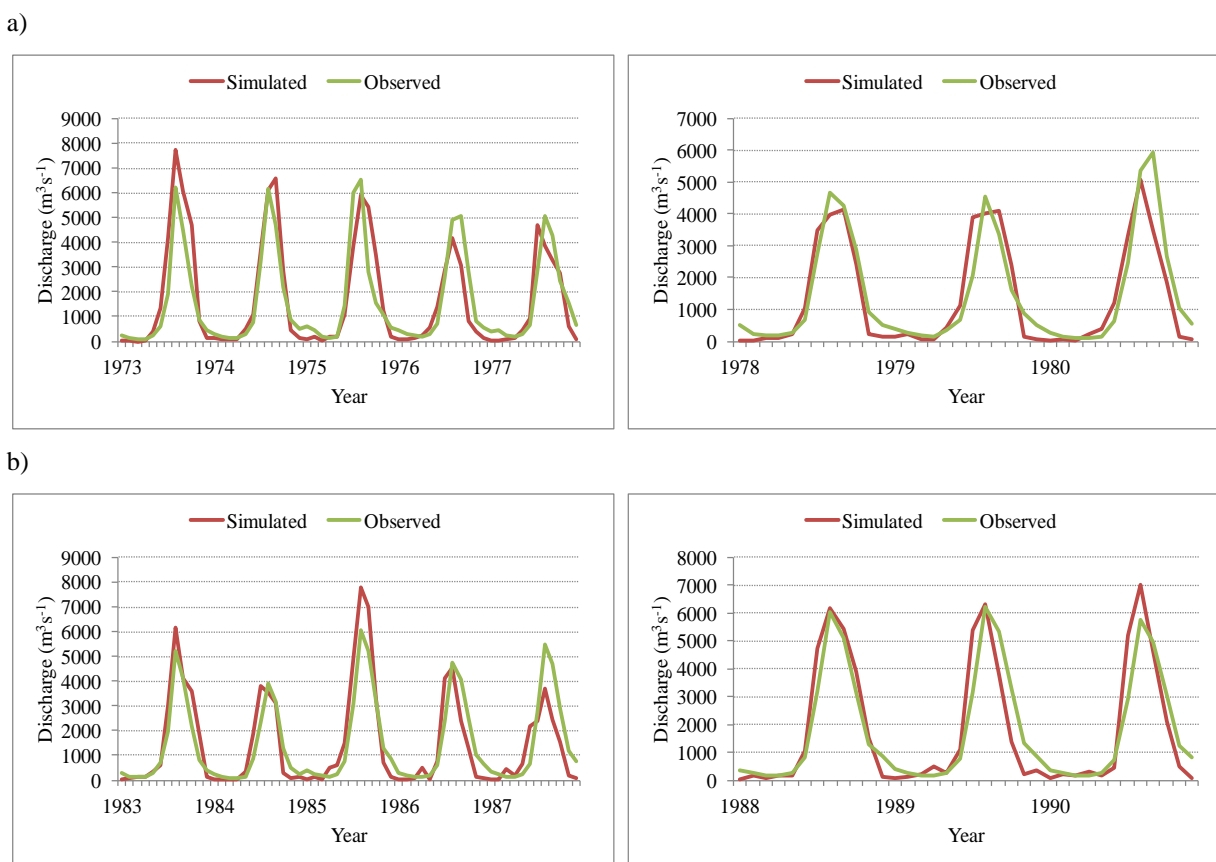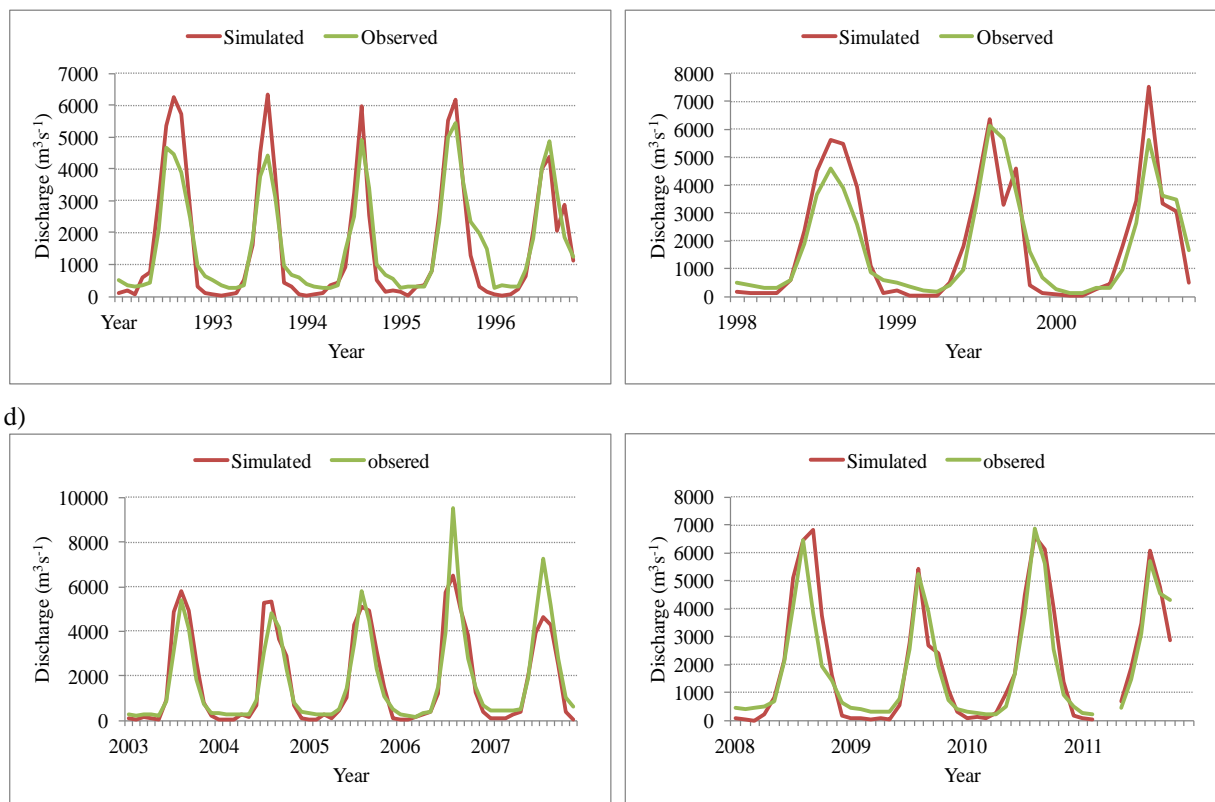
Figure 6: Calibration and validation of the SWAT hydrological model (left and right) respectively
5   a) 1970s, b) 1980s, c) 1990s and d) 2000s monthly time scale
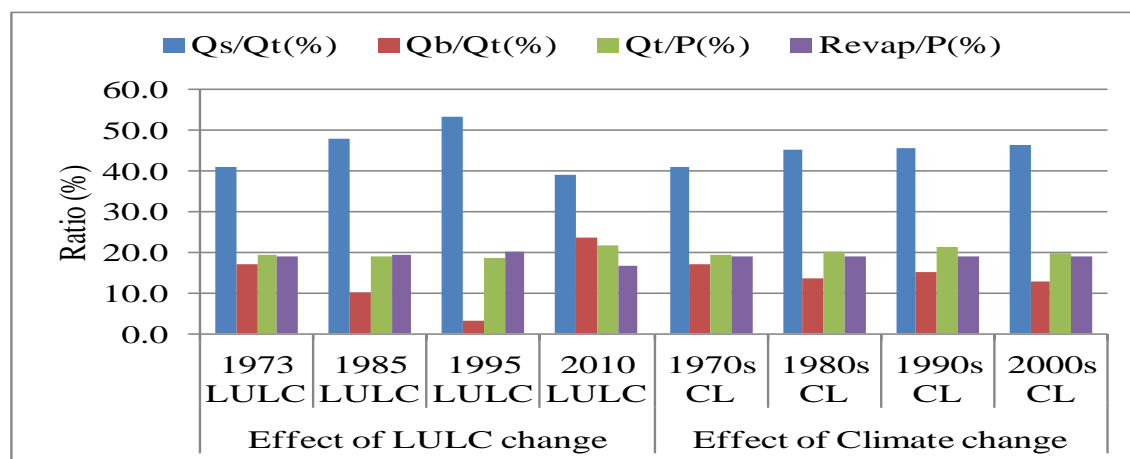
Figure 7: Ratio of water balance component analysis at El Diem station using a single effect (LULC/climate change).