Beyond streamflow
Plausible hydrological modelling for the Upper Blue Nile Basin, Ethiopia

Mohamed, Aseel; Werner, Micha; van der Zaag, Pieter

DOI
10.1016/j.ejrh.2025.102290

Publication date
2025

Document Version
Final published version

Published in
Journal of Hydrology: Regional Studies

Journal of Hydrology: Regional Studies 59 (2025) 102290

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Beyond streamflow: Plausible hydrological modelling for the Upper Blue Nile Basin, Ethiopia

Aseel Mohamed [a,*], Micha Werner [b], Pieter van der Zaag [b,c]

[a] Utrecht University, P.O. Box 80125, 3508 TC, Utrecht, The Netherlands
[b] IHE Delft Institute for Water Education, P.O. Box 3015, 2601 DA, Delft, The Netherlands
[c] Delft University of Technology, P.O. Box 5048, 2600 GA, Delft, The Netherlands

ARTICLE INFO

Dataset link: https://www.gleam.eu/, https://www.chc.ucsb.edu/data/chirps, https://cds.climate.copernicus.eu/, https://land.copernicus.eu/global/products/lc, https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/, https://earthexplorer.usgs.gov/

Keywords:
Hydrological model plausibility
SWAT+
Calibration
Actual evapotranspiration
Upper Blue Nile Basin

ABSTRACT

Study region: Upper Blue Nile Basin (UBNB), Ethiopia.
Study focus: We explored the potential of using the globally available actual evapotranspiration (ETa) dataset in the model calibration processes to enhance hydrological model plausibility for the large UBNB. We compared three calibration strategies: conventional single-point calibration based on streamflow data, spatially explicit ETa-based calibration, and a multi-variable approach incorporating both streamflow and ETa data.
New hydrological insights for the region: Our results underscore the limitations of single-variable calibration in capturing the heterogeneity of the UBNB, particularly in the estimation of ETa. By integrating ETa into the calibration process, multi-variable calibration offers improved performance across both streamflow and ETa simulations, providing valuable insights into basin dynamics and internal processes. This approach, leveraging ETa as a signature of basin heterogeneity in the calibration, demonstrates significant promise for enhancing the plausibility of hydrological models in the complex, and large UBNB while maintaining computational simplicity. We used SWAT+, which is the most recent version of the most used hydrological model in the UBNB, SWAT. Thus, this study provides a benchmark for the employment and calibration of the SWAT+ model.

1. Introduction

The Upper Blue Nile Basin (UBNB) is a crucial water source for Ethiopia, Sudan and Egypt. The basin is facing challenges, in particular, due to rapid population expansion and uncontrolled land use changes (Pricope et al., 2013). Additionally, as reported by Swain (2011) and Hassan et al. (2024), significant tensions have emerged among the countries sharing the Blue Nile River because of uncertainty regarding the availability of water resources. These tensions became more significant during the construction and filling of the Grand Ethiopian Renaissance Dam (GERD) in the Blue Nile just a few kilometres from the border between Sudan and Ethiopia. Furthermore, there is no data-sharing protocol between these countries, making hydrological modelling difficult and exacerbating the uncertainty over water availability in downstream countries, i.e., Sudan and Egypt.

Conventionally, hydrological models in the UBNB are calibrated against in-situ observed streamflow data at the outlet point of the basin (e.g., Lemann et al., 2016; Tegegne et al., 2017; Tigabu et al., 2020). This conventional calibration technique can, however, fail to accurately represent internal watershed processes (Daggupati et al., 2015; Apostel et al., 2021), leading to equifinality. Equifinality is a principle by which different parameter sets can generate equally acceptable results when evaluating model performance (Beven

* Corresponding author.
E-mail address: aseel-a@windowslive.com (A. Mohamed).

https://doi.org/10.1016/j.ejrh.2025.102290
Received 5 September 2024; Received in revised form 7 February 2025; Accepted 27 February 2025
Available online 17 March 2025

and Freer, 2001). Recent studies have aimed to reduce equifinality for various basins by employing a cascade calibration, where the model is calibrated to streamflow at the different gauge sites starting from upstream or by employing a simultaneous multi-site streamflow calibration (e.g., Wi et al., 2015; Xue et al., 2016; Lazin et al., 2020; Song et al., 2021). Although these studies have shown improvement in model performance, they necessitate access to multiple streamflow gauge stations. Additionally, these techniques entail significant computational resources.

Research also suggests the employment of other variables in the calibration along with streamflow, such as actual evapotranspiration, can reduce equifinality (López López et al., 2017; Shah et al., 2021). Dembélé et al. (2020) applied a multi-variable calibration for the mesoscale Hydrological Model (mHM) in a basin in West Africa. They incorporated ground observations of streamflow with remote sensing-based actual evapotranspiration (GLEAM), soil moisture (ESA CCI), and terrestrial water storage change (GRACE). Shah et al. (2021) calibrated a SWAT model to streamflow and remotely sensed actual evapotranspiration (MODIS) in a small basin in China. Rane and Jayaraj (2022) improved the performance of a SWAT model of a basin in India by applying multi-variable calibration. They calibrated the model to in-situ observed streamflow, remotely sensed actual evapotranspiration, and Leaf Area Index (LAI) maps.

Although the UBNB is considered an extensively studied basin Dile et al. (2018), previous hydrological models applied in literature have been calibrated solely to streamflow data, as in Gebremicael et al. (2013), Onyutha (2019), Mengistu et al. (2021), Abebe et al. (2022); and Ali et al. (2023). This limitation underscores a significant research gap in the study of the hydrology of the UBNB, particularly given its extensive area and high seasonality. Considering the constraints posed by limited hydrological data to adequately predict the hydrological dynamics of the Upper Blue Nile, this study aims to develop a plausible hydrological model leveraging remotely sensed evapotranspiration (ETa) data as well as the only available streamflow data due to the absence of a data-sharing protocol.

We assessed and analysed the internal catchment processes, as well as the streamflow at calibrated and non-calibrated points when employing different calibration strategies for the hydrological model. As the water balance in the UBNB hydrological system is mainly influenced by precipitation, streamflow and evapotranspiration (Abiodun et al., 2018), we used actual evapotranspiration and streamflow in the calibration of the model. The selection of the model for this study was based on its performance and its common use in previous studies of the study area (Van Griensven et al., 2012; Dash et al., 2021; Ali et al., 2023). Thus, we selected the SWAT+ model, the newest Soil and Water Assessment Tool (SWAT) version.

This study underscores the importance of the calibration process in hydrological model development and emphasises how calibration decisions directly influence model outcomes. Furthermore, it highlights the societal benefits, including improved water availability predictions and more effective water resource planning and allocation across various sectors, which may lead to less tension between Blue Nile riparian countries.

2. Data and methods

2.1. Study area

The Upper Blue Nile River flows 940 kilometres from Lake Tana to El Diem, the river discharge station on the Ethiopian-Sudanese border. The river basin area is approximately 176,000 km^2, comprising 17% of the area of Ethiopia (Conway, 1997). The basin topography is highly varied, ranging from lowlands to mountain ranges (Tesemma et al., 2010). As shown in Fig. 1, the majority of the basin altitude is above 2000 m.a.s.l. Towards the border with Sudan, the topography is almost flat with a minimum altitude of 494 m.a.s.l. This varied topography has a considerable impact on the spatial rainfall pattern, as most of the precipitation falls in the mountainous areas in the eastern part of the basin (Taye et al., 2015). Mean annual precipitation ranges from 800 to 2200 mm/year, while the mean annual temperature is 18.5 °C, with mean daily maximum and minimum temperatures of 25.5 °C and 10.8 °C, respectively (Legesse Gebre, 2015). The potential evaporation according to Steenhuis et al. (2009) in the UBNB is 3.5 mm/day during the long rainy season (June to September) and 5 mm/day during the rest of the year. Reflecting the seasonality of rainfall, streamflow is highly seasonal, with high flows concentrated between July and October (Taye et al., 2015). The average streamflow at the UBNB outlet in August (≈ 5950 m^3/s), when it is at its maximum, some 22 times the minimum flow in March (≈ 270 m^3/s).

The geology of the basin is primarily composed of basement and basaltic rocks, with the latter found in more than 50% of the basin area located in the Ethiopian highlands. The UBNB lowlands are composed of basement rocks and metamorphic rocks (e.g., gneisses and marbles). In addition, there are small areas of unconsolidated sediments, sandstones, and pyroclastic ashes (ENTRO, 2007). Groundwater recharge is rapid through fractured basalts, resulting in varied aquifer productivity and storage (Kebede et al., 2005). The major soil types in the UBNB are Leptosols, Nitisols, Vertisols, and Luvisols (ENTRO, 2007). Nitisols, which cover almost half of the basin, have good retention and high permeability, while Leptosols, found on steep slopes in the eastern basin, have low water retention and are prone to erosion. Luvisols around Lake Tana have moderate to good water retention and permeability (Getahun and Selassie, 2017). According to the Food and Agriculture Organisation (FAO) classification system, the basin land cover is dominated by croplands, complemented by forests and shrublands.

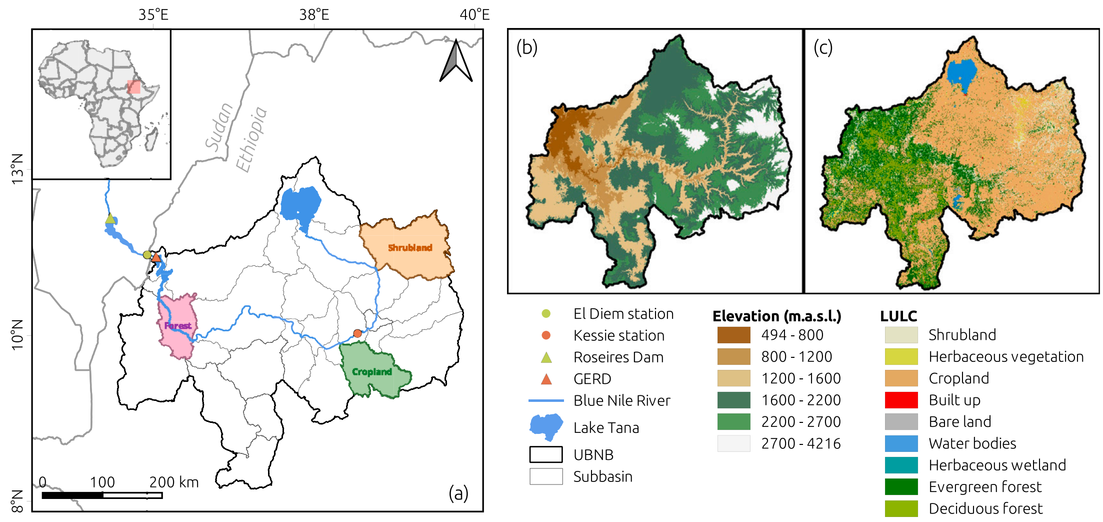


Fig. 1. Upper Blue Nile Basin and its 29 sub-basins, where the three sub-basins used for ETa calibration and analysis are coloured and indicated by pink for Forest, green for Cropland, and orange for Shrubland (a), topography (b) and land use/cover (LULC) classes (c). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. Datasets

2.2.1. Model input data

High-quality and representative rainfall data are essential in hydrological modelling. In this study, we did not have access to in-situ rainfall data. Several studies evaluated the performance of globally available remotely sensed rainfall data over the UBNB, finding that the Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) v2.0 outperformed several other products (Bayissa et al., 2017; Ayehu et al., 2018; Ali et al., 2023). The CHIRPS dataset is an integration of infrared satellite data, which offers a comprehensive overview of cloud activity, with ground-based station observations, which provide precise and localised rainfall measurements. For the UBNB, CHIRPS uses more than 40 gauge station datasets, with varied availability. The CHIRPS dataset covers the period from 1981 to the present, with a spatial resolution of 0.05° (Funk et al., 2014). The fifth-generation reanalysis product of the European Centre for Medium-Range Weather Forecasts (ERA5-Land) was used to derive other climatic data, i.e., temperature, relative humidity, and solar radiation. These datasets have hourly timestep and spatial resolution of 0.1°, spanning from 1950 to the present (Hersbach et al., 2019). The combination of CHIRPS and ERA5-Land data sources was chosen to ensure a robust dataset for hydrological modelling, considering the strengths and characteristics of each. These datasets were aggregated to form one virtual station per sub-basin with daily temporal resolution.

In addition to the dynamic meteorological datasets, three static maps were employed in the model. The Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) with a spatial resolution of 90 m (Farr et al., 2007) was used to calculate the catchment characteristics (e.g., flow direction, slope, and river network). For the soil map, the Harmonised World Soil Database (HWSD), developed by FAO and IIASA (Fischer et al., 2008), was used to extract soil information specific to the study area. For land cover, the dataset for the year 2019 with a resolution of 100 metres developed by the Copernicus Global Land Service was used (Tsendbazar et al., 2021).

2.2.2. Model calibration and validation data

Model calibration and validation were done using daily streamflow and monthly ETa. Streamflow data was obtained from the El Diem station, which is located at the basin outlet on the Ethiopian-Sudanese border. The Sudanese Ministry of Irrigation and Water Resources provided streamflow data at El Diem from 2000 to 2019. Streamflow data for El Diem station from 1997 to 1999 were unavailable, while the filling of the GERD started in 2020, limiting our analysis to the years 2000–2019. For the ETa, there is no in-situ data, but there are several global products available. The extended triple collocation approach (Stoffelen, 1998; McColl et al., 2014) was used to select the product that is best suited to represent the ETa for the basin. GLEAM v3.8a (Martens et al., 2017) dataset was found to be better representing the ETa in comparison with SSEBop and WaPOR (the methodology followed and results are in the supplementary material). GLEAM is an advanced land surface model from the Global Land Evaporation Amsterdam Model with a monthly temporal resolution and 0.25° spatial resolution. It uses satellite-derived measurements such as soil moisture and vegetation optical depth. Reanalysis of solar radiation, rainfall, and temperature datasets are used as forcing data. This product has been evaluated in different areas across the globe and has been previously applied in the Nile basin (Abera et al., 2017; Nooni et al., 2019; Lazin et al., 2020).

2.3. SWAT+ hydrological model

The SWAT model (Arnold et al., 1998) is a continuous time, process-based, semi-distributed hydrological model. Since its development, the model has been continuously updated to improve its accuracy and effectiveness (Neitsch et al., 2011). SWAT+ v2.3 (Dile et al., 2023) was used for this study. SWAT+ is an upgraded version of SWAT, which is more flexible and easier in terms of development and maintenance (Bieger et al., 2017). Inputs to the model include DEM, land cover, soil, and climatic data. Soil characteristics from the FAO database were verified to ensure they were within the appropriate ranges presented by Salter and Williams (1967), Mohr et al. (2000), Saxton and Rawls (2006); and Zeri et al. (2018). In SWAT, the watershed is partitioned into multiple sub-basins, which are further divided into hydrologic response units (HRUs) representing unique combinations of land use, soil, and slope characteristics. The simulation of the hydrological variables is done at the HRU level and subsequently aggregated to the basin scale. As it is possible to generate thousands of these combinations, a filter was implemented to reduce the number of HRUs to typically used values (Chawanda et al., 2020).

The semi-distributed SWAT+ model was selected for this study due to its widespread use in the UBNB as well as its balance between computational efficiency and hydrological detail. Although fully distributed models may be better suited for spatial calibration, these require extensive input data and present parameterisation challenges. In contrast, SWAT+ effectively represents basin heterogeneity while maintaining practical applicability.

In SWAT+ Editor, weather data were imported as one virtual station per sub-basin and the calculation methods were specified. Surface runoff was computed using the Soil Conservation Service Curve Number method, which calculates runoff based on the partitioning of precipitation between soil infiltration and surface runoff (Neitsch et al., 2011). There are three methods to calculate the potential evapotranspiration (ETp) in SWAT, namely the Penman–Monteith, Priestley–Taylor, and Hargreaves methods. Potential evapotranspiration can also be directly provided as input data. The Priestley–Taylor method (Priestley and Taylor, 1972) was used to calculate ETp so that a compatible simulation with GLEAM can be achieved. Actual evapotranspiration was determined as evaporation from canopy interception and soil and transpiration from plants.

SWAT+ Editor provides two calibration routines, soft and hard calibration (SWAT, 2023). Soft calibration requires data as information on specific water balance processes, such as the mean annual runoff coefficient and mean annual depth of the unsaturated zone (Nkwasa et al., 2022), while hard calibration data are time series at specific points, such as streamflow and ETa (SWAT, 2023). Additionally, a number of calibration tools have been developed or adapted for use with SWAT+, including IPEAT+ (Yen et al., 2019) and RSWAT (Nguyen et al., 2022).

2.4. Calibration strategies

Calibration is the process that limits the range of possible parameter values to determine the most suitable values that adequately represent the physical processes in a basin (Rajib et al., 2018). The data splitting for calibration and validation was determined based on data analysis (Dahmen and Hall, 1990; Venneker, 2011), ensuring the two datasets are statistically similar. The calibration period was from 2002 to 2014 and the validation period was from 2016 to 2019. The measurements for the year 2015 seemed to be unreliable, showing unusual flow, and were therefore omitted. The years 2000 and 2001 were used as a warm-up period.

We followed a systematic approach to analyse the model parameters and calibrate them. A local sensitivity analysis for SWAT+ parameters was performed to identify those most sensitive to streamflow and ETa. The advantage of this local analysis, where each parameter was changed individually, is that it allows for an in-depth understanding of the impact a parameter variation has on the model output. Particular attention was paid to the highly sensitive parameters reported by Sultana et al. (2019), Mengistu et al. (2021); and Takele et al. (2021) in their studies of the UBNB, as shown in Table 1. In addition to the local analysis, a global sensitivity analysis was conducted with 300 simulations, following the uniform Latin Hypercube Sampling approach (LHS). The uniform LHS is a global search method. It divides each parameter's range into equal intervals and randomly samples within these intervals, ensuring that the entire space is explored uniformly.

After identifying the most sensitive parameters, their values were fitted to streamflow and ETa using three calibration strategies, detailed below. The first of these strategies (calibration I) was based on the conventional calibration method followed in the literature, calibrating the model to streamflow at the basin outlet at El Diem (Fig. 1). A second strategy (calibration II) followed was by calibration of the model to monthly ETa of three sub-basins, while the third strategy (calibration III) was based on calibrating the model to both daily streamflow and monthly ETa, i.e., multi-variable calibration. These calibration processes were performed sequentially: calibration II was performed on the model already calibrated by calibration I and calibration III was conducted on the model post-calibration II. The calibration of ETa was performed using globally available GLEAM datasets, while the streamflow was based on the in-situ streamflow data of El Diem.

SWAT+ was first calibrated automatically using the RSWAT tool (Nguyen et al., 2022), constraining the model parameters with the prior ranges shown in Table 1. After identifying the best parameter set from RSWAT, the model was critically fine-tuned by manual calibration using the hard calibration routine in SWAT+ Editor to ensure that parameter values were realistic and resulted in acceptable simulation based on the literature and our understanding of the catchment. Additionally, for ETa calibration in the second and third strategies, biophysical parameters, shown in Table 2, were fitted manually as they are not available in the automatic calibration tool. The values of these parameters, e.g., biomass and leaf area index, were adjusted following the ranges identified by Dile et al. (2020).

Two types of change were applied to the model parameters: relative and replace. The relative change, or percentage change, was applied to parameters related to geographical attributes such as soil type, land use class, and/or slope band. In this change type, the value of each parameter is multiplied by a factor. This is essential for maintaining the spatial variability of basin characteristics. In contrast, a replacement change was applied for the parameters that exhibit uniformity across the study area. The parameter value changed to a new value in each simulation. This includes parameters such as channel hydraulic conductivity.

Table 1
Ranges of sensitive parameter in SWAT+ model, the fitted value ranges reported in the literature for the UBNB and the ranges used in this study for calibration.

Parameter	Description	Change type	Parameter range	Reported range	Calibration range
cn2	Initial SCS runoff CN for moisture condition II	Relative	35–95	−0.15–0.11	−0.25–0.25
alpha_bf[a]	Baseflow alpha factor (days)	Replace	0–1	0.19–0.35	0.03[a]
revap_co	Groundwater "revap" coefficient	Replace	0.02–0.2	0.12–0.17	0.02–0.2
esco	Soil evaporation compensation factor	Replace	0–1	0.68–0.95	0.1– 1
revap_min	Threshold depth of water in the shallow aquifer to return to the root zone (m)	Replace	0–50	0.07–0.228	0–1.5
perco	Percolation coefficient; adjusts soil moisture for percolation to occur	Replace	0–1	0.28	0–1
chn	Manning's 'n' value for the main channel	Replace	−0.01–0.3	0.084–0.15	0.01–0.15
awc	Available water capacity (mm_{H_2O}/mm_{soil})	Relative	0.01–1	0–0.5	−0.25–0.25
k	Saturated soil hydraulic conductivity (mm/h)	Relative	0–2000	−0.43–0.15	−0.25–0.25
chk	Channel effective hydraulic conductivity (mm/h)	Replace	−0.01–500	75–134	50–200
canmx	Maximum canopy storage (mm)	Replace	0–100	3–4.9	0–15
epco	Plant uptake compensation factor	Replace	0–1	–	0–1
surlag	Surface runoff lag coefficient (days)	Replace	0.05–24	–	0.05–24
evlai	Leaf area index at which no evaporation occurs from the water surface	Relative	0–10	–	−0.25 –0.25
ovn	Manning's factor 'n' for overland flow	Relative	0.01–0.3	–	−0.25–0.25
bd	Soil bulk density (g/cm^3)	Relative	0.9–2.5	–	−0.25–0.25
alb	Soil albedo when soil is moist	Relative	0–2.5	–	−0.25–0.25

[a] alpha_bf was calculated from the observed hydrograph using the equation $(\ln\left(\frac{Q_N}{Q_0}\right)/N)$, N is the number of time steps after the recession began (Neitsch et al., 2011).

2.5. Model performance evaluation

The streamflow simulated from each calibration strategy was evaluated using Nash–Sutcliffe Efficiency (NSE), Kling Gupta Efficiency (KGE), and Percent Bias (PBias). The use of these metrics provides broad insights into model performance. NSE (Nash and Sutcliffe, 1970) is a commonly used metric, examining both the mean and variability compared to the observed data. However, NSE is sensitive to extreme flows and biases (Gupta et al., 2009). To overcome these drawbacks, KGE was also used. KGE (Gupta et al., 2009) is a comprehensive metric that considers bias, correlation, and variability, providing a more holistic assessment of the model's ability to recreate various aspects of hydrological behaviour. On the other hand, the performance evaluation of simulated ETa was through the Ratio of Root Mean Square Error to Standard Deviation (RSR) and Percent Bias (PBias). These metrics were chosen differently for ETa simulations compared to streamflow based on their suitability for assessing the adequacy of ETa simulations in capturing the variability and bias. KGE and NSE, on the other hand, reflect their effectiveness in assessing flow dynamics and overall performance in streamflow simulations. The performance ratings of these metrics were based on the classification proposed by Moriasi et al. (2007) and Rogelis et al. (2016).

During the calibration process, the monthly ETa was evaluated for three sub-basins (see analysed sub-basins in Fig. 1a), each characterised by its unique land use of forest, cropland, and shrubland. The forest and cropland sub-basins were mostly covered by their respective land uses, with over 80% of the area occupied by forests or crops. However, the shrubland sub-basin had only 35% of its area covered by shrubland. Although it is not dominated by shrubland, it has the highest percentage of shrubland cover among all the 29 sub-basins generated by SWAT+. The selection of these sub-basins was made to ensure that the model could effectively capture the variability in ETa associated with diverse land use patterns.

2.6. Model comparison

The results from the three calibration strategies were compared and analysed at various spatio-temporal scales. In the study, ETa was chosen as the internal catchment process due to its significance in the basin's water balance dynamics accounting for more than 40% of the UBNB water balance (Abera et al., 2017; Abebe et al., 2022). In addition to evaluating overall model performance, the results were analysed in quantile domains (duration curves) to assess the model's ability to simulate low and high values of streamflow and ETa (Tegegne et al., 2017). Furthermore, we investigated the relationship between the simulated ETa performance of the 29 sub-basins and the sub-basin elevation and land cover types to understand how these factors influence ETa performance across the basin.

3. Results and analysis

3.1. Fitting the model parameters

Starting with a sensitivity analysis, our result revealed that the parameter sensitivity is dependent on the calibration strategy. The analysis results indicated that the curve number (cn2) is the parameter that the streamflow is most sensitive to, while ETa

Table 2
Fitted parameters for the three calibration strategies and their performances during the validation period (2016–2019).

Parameter	Change type	Fitted value		
		Calibration I	Calibration II	Calibration III
cn2	Relative	−0.12	−0.12	–
awc	Relative	−0.1	−0.1	−0.18
perco	Replace	0.4	0.4	0.4
epco	Replace	0.95	0.95	0.95
esco	Replace	0.9	0.9	0.9
chk	Replace	150	150	150
chn	Replace	0.011	0.011	0.011
bd	Relative	0.1	0.1	0.1
k	Relative	−0.03	−0.03	−0.03
alpha_bf	Replace	0.03	0.03	0.03
revap_min	Replace	0.6	0.6	0.6
revap_co	Replace	0.18	0.18	0.18
Forest: Initial biomass (kg/ha)	Replace	–	1000	1000
Forest: Initial residue cover (kg/ha)	Replace	–	1000	1000
Forest: Initial leaf area index (m^2/m^2)	Replace	–	5	5
Forest: Biomass energy ratio ((kg/ha)/(MJ/m^2))	Replace	–	10	10
Forest: Maximum leaf area index (m^2/m^2)	Replace	–	7	7
Shrubland: Initial biomass (kg/ha)	Replace	–	1000	1000
Shrubland: Biomass energy ratio ((kg/ha)/(MJ/m^2))	Replace	–	40	40
Performance metrics				
NSE		0.83	0.20	0.61
KGE		0.69	0.44	0.66
PBias		21.6%	62.7%	42.7%

Note: The changes presented in this table show deviations from the default values rather than sequential modifications between calibrations. For example, CN2 was adjusted by 12% relative to its default value in both Calibration I and Calibration II and remained at its default value in Calibration III.

is most sensitive to the soil evaporation compensation factor (esco). The soil bulk density (bd) was found to be sensitive for both calibration variables, i.e., streamflow and ETa. Although these parameters were identified as the most sensitive, adjusting each has implications for the physical representation of the watershed. For instance, the curve number is influenced by soil type and land use, which were incorporated into the model using globally available data and subsequently validated with Google Earth imagery and literature. Adjusting these parameters entails altering the physical characteristics of the study area, highlighting the importance of careful calibration. Given the lack of field data, calibration remains the optimal approach for adequate model performance.

After identifying all the sensitive parameters, they were fitted to streamflow and ETa following calibration I, calibration II, and calibration III strategies. In calibration I, changes were made to the model parameters to improve the daily streamflow simulation at El Diem, the basin outlet point. Generally, the results show that the model has the potential to capture the dynamics and variability of streamflow. However, as depicted by the high positive bias percentage shown in Table 2, the model significantly underestimates streamflow during the validation period from 2016 to 2019. Calibration II was based on considering calibration I as a baseline model, and changes were made to the biophysical parameters only. The biophysical parameters were adjusted following the ranges specified by Dile et al. (2020) for Ethiopia. The focus on ETa performance led to the degradation of the streamflow performance, although the streamflow sensitive parameters were kept constant as reflected in Table 2. This second model was then used as a baseline for the multi-variable calibration III strategy. For this strategy, it was found that keeping all the parameters as they were in calibration II, while only removing the -12% from the curve number (cn2), i.e., keeping it as model default, and reducing the available water capacity (awc) parameter by -18% produced the best simulations of streamflow and ETa (see Table 2). Hydrographs for all simulations are provided in the supplementary material, Fig. S4.

The model performance for simulating daily streamflow in the three strategies during the validation period is shown in Table 2. These metrics indicate that the model performs better in representing the daily streamflow at El Diem when calibrated solely to streamflow, while it exhibits the poorest performance when streamflow is not taken into consideration in calibration, even though the parameters streamflow is sensitive to were not changed. However, balanced performances for streamflow and ETa were achieved when both calibration variables were considered during calibration. The high PBias for the three strategies are due to the significant underestimation of streamflow in the model, resulting from a drawback in the SWAT model, where infiltrated water fails to contribute to the aquifer and instead disappears from the system (Van Griensven et al., 2012). Some improvements have been made in SWAT+ to address this issue, as discussed by Sánchez-Gómez et al. (2024). However, despite the careful calibration conducted in this study, water losses persist, indicating that some limitations remain within SWAT+.

3.2. Exploring the ETa simulation from different land uses

Similar to the changes observed in the daily streamflow simulation depending on the calibration strategy, the simulation results of ETa also vary. The performance of ETa simulation was explored in three sub-basins, which are shown in Fig. 1a, and compared

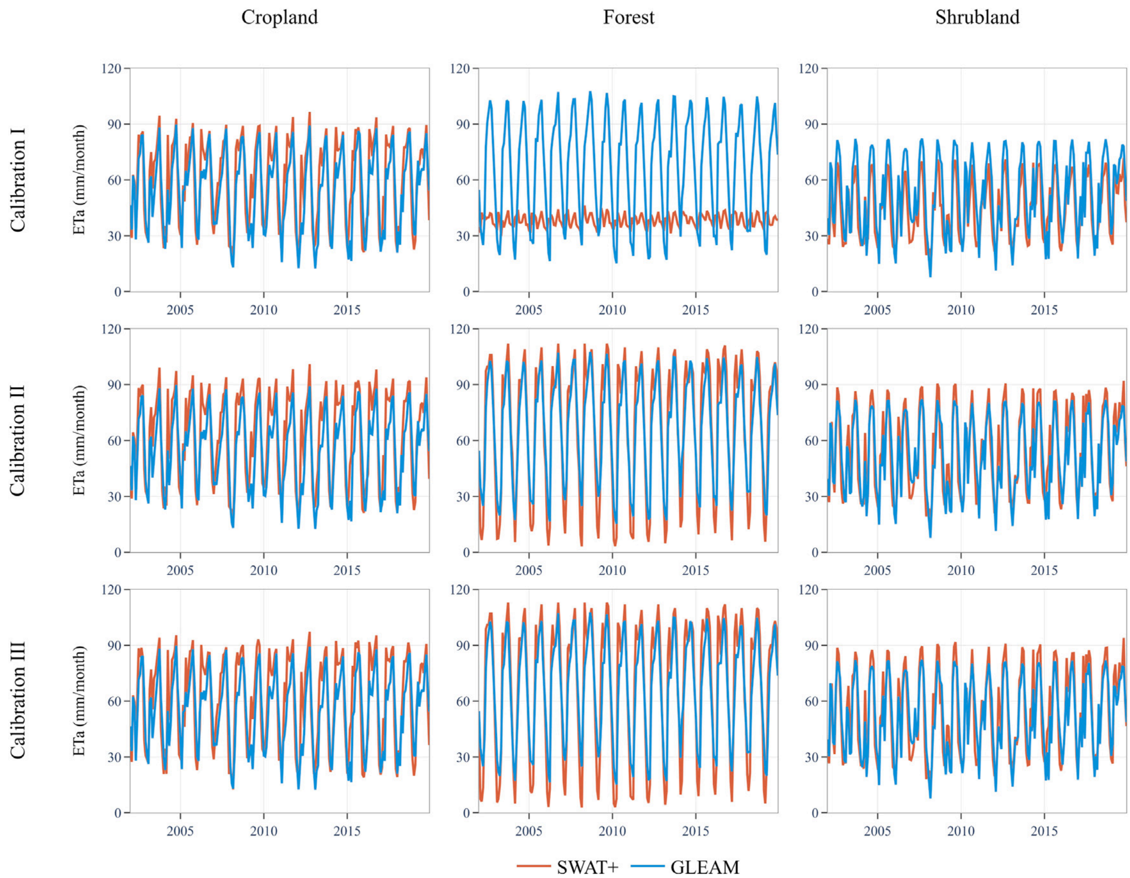


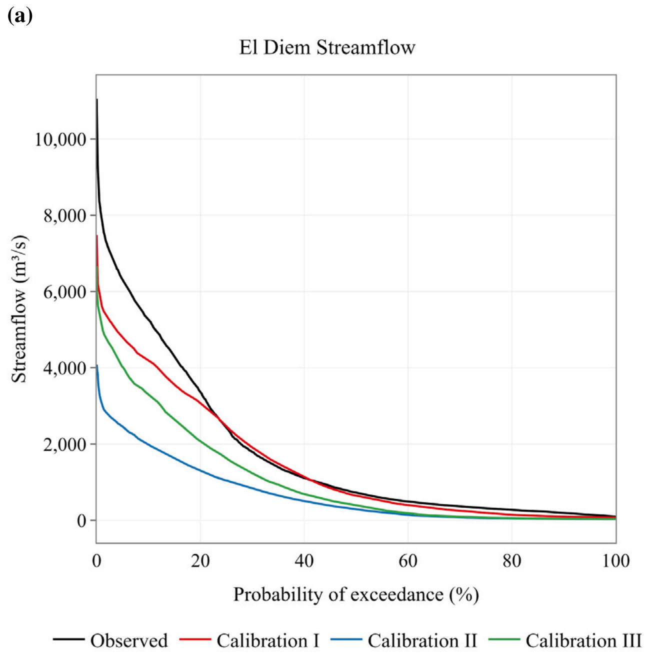
Fig. 2. Monthly ETa of GLEAM (blue lines) and SWAT+ simulations (red lines) calibrated to streamflow (first row), GLEAM ETa (second row), and streamflow and ETa (third row). The simulations of the cropland sub-basin are shown in the first column on the left, the forest sub-basin in the middle, and the shrubland sub-basin in the third column.

to GLEAM monthly ETa. Although the model subjected to calibration I yielded the best performance for streamflow, it provided inaccurate estimations for monthly ETa, especially in forested areas, as illustrated in Fig. 2. The ETa simulation for the forest basin exhibited an unrealistic pattern. However, after adjusting the biophysical parameters in calibration II, the ETa simulation for the forest sub-basin significantly improved. Similarly, the calibration II model produced better simulations for ETa in the shrubland basin. Comparison between calibration II and calibration III indicates a slight decline in ETa simulation performances in the three sub-basins, compensating for the higher performance in streamflow values (see Table 2).

Fig. 3 presents the flow duration curves at El Diem and ETa at the three sub-basins for the three calibration strategies. It is evident from Fig. 3(a) that all calibration strategies resulted in an underestimation of the El Diem flow compared to the observed data (black line) across all flow conditions. The degree of underestimation decreased with decreasing streamflow; in other words, the most significant underestimation occurred during high-flow conditions (0—10%), while the least underestimation occurred during dry conditions (90—100%). On the other hand, a slight overestimation of ETa in the three sub-basins was observed, especially during the high ETa period from August to October, except for calibration I, which showed unusual performance in the forest sub-basin and underestimation in shrubland. Furthermore, the results demonstrate that calibration III notably enhances streamflow simulation at El Diem compared to calibration II, although a very slight deterioration in ETa simulation was observed. This implies that calibration II improved the performance of ETa at the expense of streamflow.

3.3. Spatial variation in ETa simulation performance

Analysing the monthly ETa simulation results across all the 29 sub-basins of the UBNB simultaneously, Fig. 4 highlights that in calibration I, all sub-basins exhibit higher PBias compared to calibrations II and III. Additionally, approximately 45% of the sub-basins show slightly lower RSR. Moreover, it is apparent that calibrations II and III yield similar ETa performance in terms of PBias and RSR. Specifically, Fig. 4(a) indicates that approximately 30% of sub-basins in calibration I overestimate ETa, while around 90% do so in calibrations II and III.



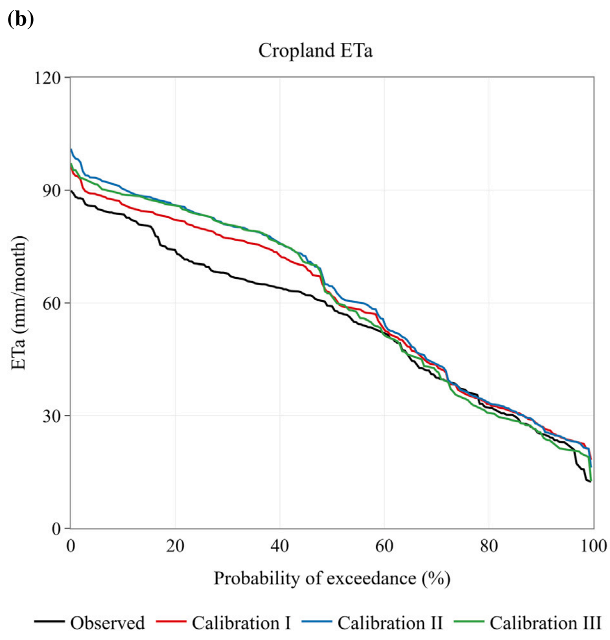


Fig. 3. Duration curves of SWAT+ simulation results and observed data using the three calibration strategies. (c) ETa in the forest sub-basin, and (d) ETa in the shrubland sub-basin. The black line in (a) represents El Diem observed streamflow, while in (b), (c), and (d), it represents GLEAM ETa data.

In analysing the simulations concerning the sub-basin characteristics, Fig. 5(a) indicates that most of the underestimation in calibration I occurs in forest-dominated sub-basins, as well as those dominated by cropland & forest. This tendency may be attributed to the high sensitivity of forest areas to ETa. The analysis included a total of 7 sub-basins dominated by forest, 14 by cropland, 1

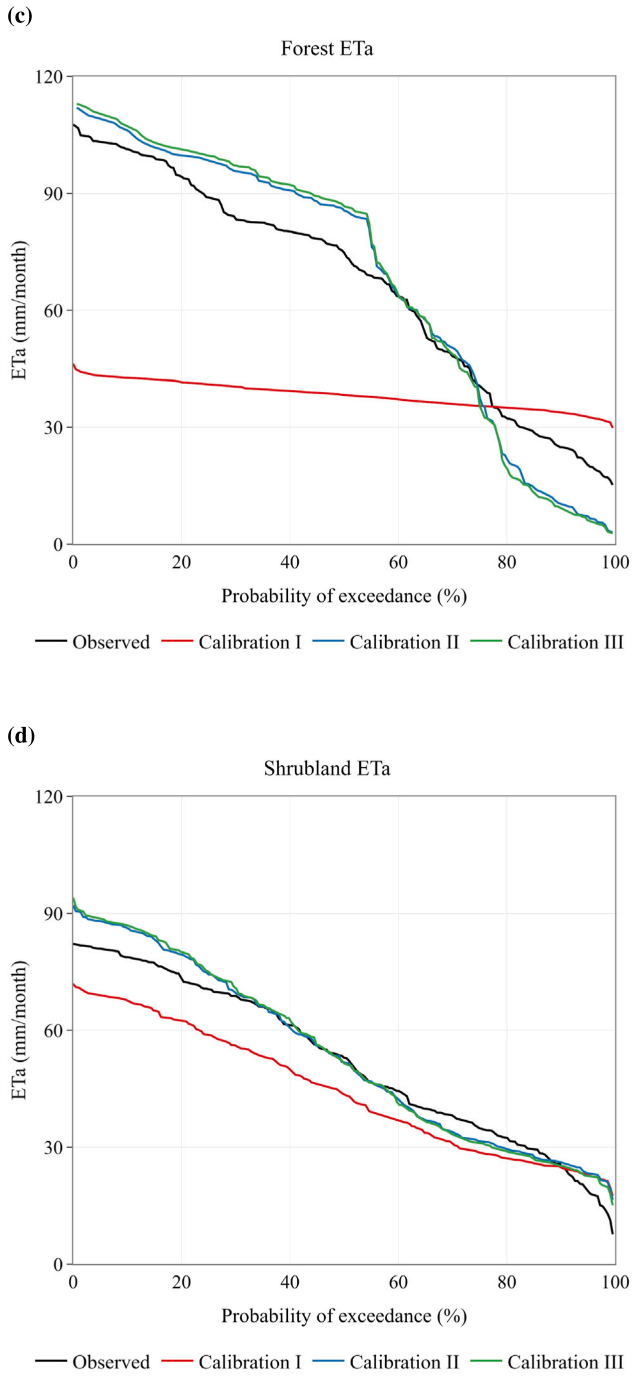


Fig. 3. (*continued*).

by a mix of cropland & shrubland, and 7 by cropland & forest. The best performance in calibration I, as depicted in Figs. 5(a) and 5(b), is observed in cropland sub-basins. The land use management routine in SWAT, where existing crops are integrated into the model, enables better simulation of cropland sub-basins even with little or no calibration. Analysing the correlation between average sub-basin elevation and ETa performance, calibration I exhibits better simulations for higher elevations, as illustrated in Figs. 5(a)

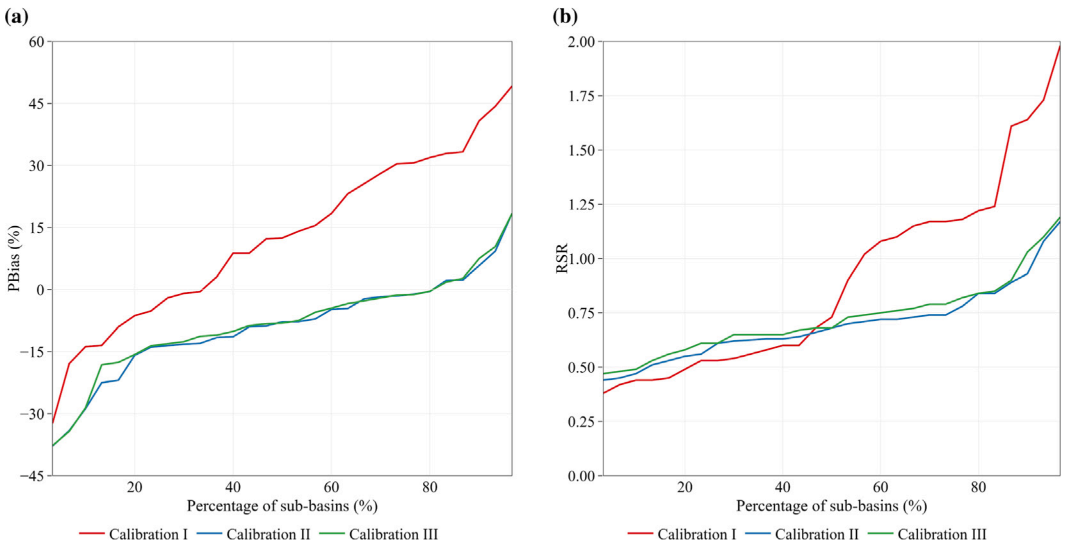


Fig. 4. Distribution of ETa simulation performances for the three calibration strategies across the 29 sub-basins of the UBNB. The x-axis represents the percentage of sub-basins out of all the 29 sub-basins. (a) Performance distribution using PBias, and (b) performance distribution using RSR. Calibration I is illustrated in red in the two plots, calibration II in blue, and calibration III in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and 5(b). However, this link cannot be solely attributed to elevation, given that nearly all forest sub-basins are located in lower elevation bands. In contrast, calibrations II and III show no clear relation or pattern between ETa performance and the elevation band. The spatial distribution map of the model performance concerning elevation and land use type averaged per sub-basin is provided in the supplementary materials, Fig. S6.

4. Discussions

4.1. Application to large and heterogeneous basins

Simulating water flows in large and spatially heterogeneous basins, such as the UBNB, presents significant challenges due to the complex nature of basin processes. Previous studies have demonstrated improvements in hydrological simulations when additional variables are included in the model calibration (Dembélé et al., 2020; Shah et al., 2021; Rane and Jayaraj, 2022). However, these studies were primarily conducted in relatively small basins, whereas the internal watershed processes in large and heterogeneous basins are more significant (Daggupati et al., 2015). In our study, the SWAT+ model successfully captured streamflow during calibration I, but significant disparities were detected in the simulation of ETa, indicating limitations in representing internal catchment processes. This finding aligns with those of Gao et al. (2018) and Ávila et al. (2022), who emphasised the difficulties in accounting for all catchment processes in large heterogeneous basins. For example, Ávila et al. (2022) conducted a review of landscape-based hydrological models, highlighting how landscape heterogeneity affects water balance components and underscoring the importance of considering this complexity when applying hydrological models. Additionally, Srivastava et al. (2020) presented an improved approach to address the challenge of representing land use heterogeneity in hydrological models. Their method involved incorporating both uniform and heterogeneous agricultural land use to analyse its impact on hydrological responses, leading to more realistic simulations that better reflect the impact of land use variations.

In our study, integrating ETa in the calibration process as a signature of land use heterogeneity proved effective in improving model performance. This approach is supported by Shah et al. (2021), who proposed incorporating remotely sensed ETa data into model calibration to capture crucial land surface processes, thereby constraining model parameters and achieving more accurate representations of catchment processes. One advantage of using ETa as a signature is its feasibility within a lumped modelling framework, where land use is simplified. Furthermore, as Mei et al. (2023) discussed, the inclusion of ETa in the calibration reflects the relationship between soil moisture and ETa and enhances soil moisture simulation, thereby improving the overall performance of the model. The Gravity Recovery and Climate Experiment (GRACE) can be utilised to provide key parameters for terrestrial water storage estimation, thereby reducing biases in hydrological models. GRACE data (Rodell and Famiglietti, 2001) is particularly powerful for tracking changes in total water storage, including surface water, soil moisture, and groundwater (Abera et al., 2017). However, while we acknowledge the potential of GRACE, it was beyond the scope of our research, which focused on improving the model performance at the sub-basin level. A study by Chavarría et al. (2022) found that the performance of GRACE diminishes as the catchment area decreases. The average area of the UBNB sub-basins is approximately 10,000 km^2, significantly smaller than the threshold of around 70,000 km^2, below which GRACE's performance notably declines.

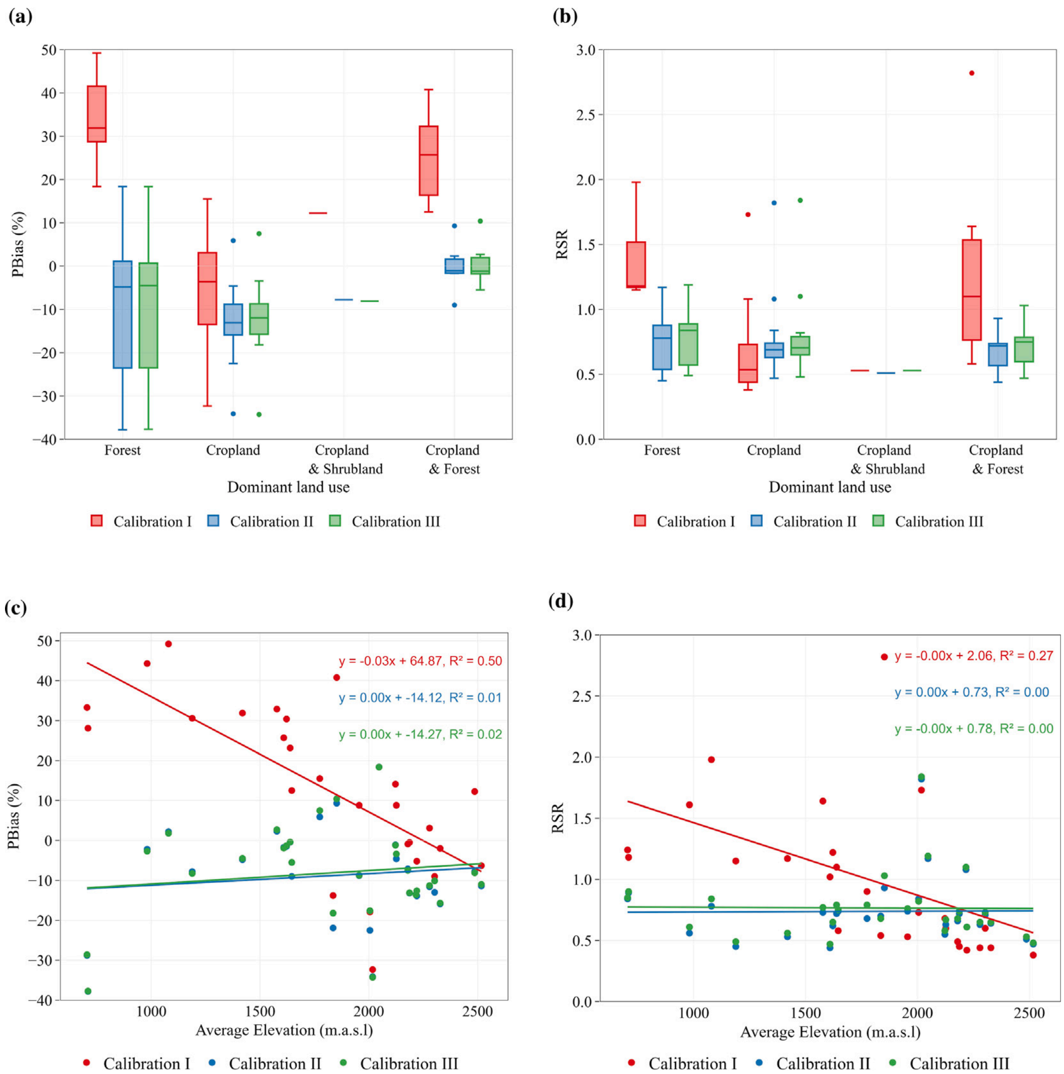


Fig. 5. The correlation between ETa simulation performance and sub-basin characteristics, where (a) is the dominant land use per sub-basin and PBias, (b) is the dominant land use per sub-basin and RSR, (c) is the sub-basin average elevation with PBias, and (d) is the sub-basin average elevation with RSR. The four plots illustrate calibration I in red, calibration II in blue, and calibration III in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.2. Hydrological model plausibility

In further analysis, we assessed the simulation of streamflow at the Kessie Bridge gauge station (see Fig. 1). Due to the unavailability of streamflow data from this station for this study, we referenced observed flow graphs presented by Takele et al. (2021) and Abebe et al. (2022). Our results from the three calibration strategies revealed a significant underestimation of flow at this location by the model (see Fig. S5 in the supplementary materials). The assessment revealed that while the calibration II strategy significantly underestimated streamflow at El Diem compared to calibration I, there was only a small difference between them in simulating the streamflow at Kessie Bridge. This is because parameter adjustments in calibration II primarily focused on the forested area downstream of Kessie Bridge, with minor adjustments to shrubland parameters (upstream area; see Fig. 1c).

Calibration III, however, produced improved monthly streamflow results at Kessie Bridge. This shows that the incorporation of ETa into the calibration process for El Diem had a cascading positive effect on streamflow representation in additional locations. This improvement can be understood by examining the compensatory effects in the other two calibration strategies. In calibration I, the model achieved better streamflow results at the basin outlet by compensating for poor ETa performances, effectively redistributing water in a way that masked underlying inaccuracies. In calibration II, while the focus on ETa improved its representation across the catchment, it did so at the expense of streamflow, as the calibration did not fully resolve the compensatory mechanisms at play. Calibration III, with its multi-variable approach, reduced the model’s reliance on such compensations. By adjusting key parameters of curve number (cn2) and available water capacity (awc), calibration III achieved a more accurate and balanced water distribution across the catchment, leading to plausible simulations of both ETa and streamflow at El Diem and other locations. In agreement with Yang et al. (2018) and Haas et al. (2022, 2024), our findings emphasise that a sole focus on streamflow, as often seen in hydrological modelling practice, can lead to significant misrepresentation of internal catchment processes, particularly regarding ETa from forested areas.

The erroneous simulation of ETa in forest basins is a problem previously noted in the literature. Yang and Zhang (2016) have highlighted that the default settings of SWAT often include unrealistic radiation use efficiency and leaf-to-biomass fraction, leading to inadequate estimations of ETa. Alemayehu et al. (2017) reported that the model's simplistic representation of canopy processes and inadequate parameterisation of the soil-plant-atmosphere continuum further contribute to these inaccuracies. By incorporating spatially distributed ETa data alongside streamflow data into the calibration process, we significantly improved the model's performance. This adjustment allowed for a more realistic representation of forest ETa and better alignment with GLEAM ETa values, effectively addressing the previously noted discrepancies. In future revisions of the SWAT+, the representation of canopy processes, as well as the default settings used could well be improved.

Our comprehensive analysis of basin hydrological processes, compared with real-world hydrology, challenges the efficacy of the widely used calibration strategy focusing solely on streamflow at the UBNB outlet (calibration I) using SWAT+. Despite yielding the best performance in terms of the NSE metric, the calibration I strategy is critiqued for its implausibility. Incorporating spatially distributed ETa data alongside streamflow into the calibration process allowed for a better simulation of vertical and horizontal water fluxes and balanced energy exchange between the canopy and ground. Moreover, highlighting that calibration I achieved the highest performance scores agrees with the notion that "the best is the enemy of the good", as argued also by Beven (2006) and Cinkus et al. (2023).

Although the use of remotely sensed ETa served to improve the plausibility of the calibrated model, the reliance on GLEAM as an observational dataset introduces a layer of uncertainty. While a comparison of three ETa products through triple collocation favoured GLEAM for the UBNB, it is important to recognise that no dataset is entirely error-free (Weerasinghe et al., 2020). Assumptions of error-free data may overlook biases stemming from indirect estimation methods, retrieval algorithm limitations, or inaccuracies in meteorological inputs. Moreover, the spatial resolution of the data and the use of monthly, rather than daily, ETa measurements may restrict the strategy's efficacy. This may overlook rapid fluctuations in evapotranspiration, potentially affecting model accuracy under dynamic environmental conditions (Zhang et al., 2021). Similarly, the use of CHIRPS for rainfall as an observational input also introduces uncertainty. CHIRPS has demonstrated strong performance in the UBNB through correcting satellite rainfall estimation with the available data from the limited number of rainfall stations in the basin Ayehu et al. (2018), Ali et al. (2023) and Lemma et al. (2019). Despite this, its limitations must be acknowledged, as CHIRPS tends to underrepresent high-intensity rainfall events. One reason for this is the data assimilation approach, which is not dynamically nudging-based (Funk et al., 2014), but rather applies more spatially generalised corrections. This may result in extremes being underrepresented when compared to those observed at individual rainfall stations.

4.3. Which calibration strategy to follow?

Calibrating hydrological models is pivotal for adequate water resource management and policy formulation. However, choosing the appropriate calibration strategy remains a subject of debate within the hydrological modelling community. Effective calibration of hydrological models necessitates the integration of both horizontal and vertical flux components, along with an understanding of the interconnectedness of model parameters. In this study, we combined a spatial (GLEAM ETa) with a single-point (El Diem streamflow) calibration. This hybrid approach was identified as the most effective in producing adequate hydrological simulations, as discussed by Rajib et al. (2018). Additionally, researchers such as López López et al. (2017) and Herman et al. (2018) emphasise the importance of incorporating both horizontal (e.g., streamflow) and vertical (e.g., ETa) flux components in the calibration process. Neglecting either component can lead to incomplete representations of watershed processes. This was apparent in calibrations I and II, as well as in attempts to calibrate the model to ETa from scratch, without using calibration I as a baseline, which produced unrealistic streamflow simulations (see Fig. S4 in the supplementary materials). Furthermore, Zhang et al. (2021) caution against relying solely on single variable calibration, as it may result in simulations that are only acceptable for the calibrated variable, failing to capture the complexities of the entire system as outlined in our results. These findings highlight the interconnected nature of hydrological systems and the importance of considering multiple variables in calibration.

In a more in-depth analysis, we applied the sequential calibration concept but started with ETa, followed by streamflow, and then multi-variable calibration. This analysis revealed that calibrating the model to ETa first generated adequate ETa simulations ($NSE_{average}$ = 0.81) but erroneous streamflow simulations (NSE = −8.7) with substantial spikes, as expected. However, subsequent calibration to streamflow significantly improved streamflow simulation (NSE = 0.81) and kept the very good performance of ETa ($NSE_{average}$ = 0.76). This performance can be attributed to the sensitivity of model parameters to these variables. Streamflow is highly sensitive to parameters that also influence ETa. For example, soil bulk density affects both variables differently: the t-stat, a ratio of a parameter value to its error (Abbaspour, 2015) offering a measure of the degree of sensitivity, is 12.64 for streamflow and −1.76 for ETa. This difference in sensitivity significance led to better performance in the second sequential calibration, even though ETa was not (directly) included in the calibration process. These results suggest that starting the calibration with a variable that represents catchment heterogeneity, rather than one that represents a lumped response, can be a more effective approach.

In the UBNB, our review reveals a predominant focus on calibrating hydrological models solely on streamflow data. For example, studies by Gebremicael et al. (2013) and Mekonnen et al. (2018) investigated the impacts of land use and land cover changes on streamflow using the SWAT model, calibrated to El Diem streamflow data. While their models achieved good performance in simulating streamflow at the basin outlet point, it would be important to assess how well the basin heterogeneity is represented. Thus, to ensure more plausible simulations and address the complexity of the UBNB, we suggest the adoption of a multi-variable calibration strategy in such studies. Similarly, Lazin et al. (2020) and Takele et al. (2021) analysed water cycle components in the basin using SWAT and the distributed CREST model, respectively. Although all these studies achieved high-performance ratings for streamflow in terms of NSE, according to the criteria established by Moriasi et al. (2007), relying solely on streamflow performance may not fully capture the intricacies of the hydrological system within the UBNB. Yet, some experts argue that for studies focusing solely on streamflow, multi-variable calibration may not be necessary. The applicability of such models should be considered carefully.

In this study, we primarily focused on time-series evaluation metrics (e.g., NSE). However, some recent studies incorporate pattern performance metrics for calibrating gridded data, such as ETa. These metrics complement temporal-based metrics (e.g., NSE) for streamflow, providing additional insights into model performance. A commonly used measure is the SPAtial EFficiency metric (SPAEF), developed by Koch et al. (2018), which evaluates spatial pattern agreement. Since these metrics focus on spatial variability, they are mostly applied to fully distributed models (e.g., Demirel et al. (2024, 2018)). Future research could explore the impact of calibrating ETa using spatial metrics (Eini et al., 2023) at the HRU and sub-basin levels, comparing their effectiveness to traditional temporal-based approaches. This could enhance SWAT+ model performance by improving spatial pattern representation.

While acknowledging computational costs associated with multi-variable calibration strategies, it is evident from our results that such an approach is required for achieving plausible hydrological simulations. The findings underscore the necessity of investing computational resources to ensure the accuracy of model outputs. Given the common single-variable calibration routines in existing models, we suggest that model developers incorporate multi-variable calibration capabilities in their modelling software, which may simplify the calibration process for researchers and hydrological modellers, thereby enhancing the accuracy and robustness of hydrological predictions.

5. Conclusions

This study investigated the plausibility of the Soil and Water Assessment Tool Plus (SWAT+) model in the large and heterogeneous Upper Blue Nile Basin (UBNB), addressing the challenges posed by the limited presence of ground gauge stations and the lack of data-sharing protocols between the Blue Nile riparian countries. Globally available remote sensing and reanalysis datasets were used as inputs for SWAT+. CHIRPS, which integrates satellite precipitation estimates with ground-based stations, was used as precipitation forcing. The reanalysis ERA5-Land datasets were used for temperature, solar radiation, and relative humidity. Based on statistical data analysis, the model was developed for a period of 20 years: a two-year warmup (2000–2001), thirteen years for calibration (2002–2014), and four years for validation (2016–2019). Three calibration strategies were designed to assess and analyse the plausibility of the hydrological model, with streamflow and actual evapotranspiration (ETa) as calibration variables. Actual evapotranspiration was selected as a signature of internal catchment processes due to its representation of more than 40% of the water budget in the UBNB. The Global Land Evaporation Amsterdam Model (GLEAM) ETa dataset was chosen for this study due to its good performance over the UBNB. The internal catchment process analysis was conducted across various sub-basins with different characteristics to capture the heterogeneity of the region effectively, along with El Diem streamflow at the basin outlet point.

Through sensitivity analysis, we identified the most sensitive parameters to be used for calibrating the model. Subsequent model parameter fitting, employing three calibration strategies, provided an in-depth insight into the internal hydrological dynamic of the basin. The first strategy (Calibration I), focusing solely on the observed streamflow at El Diem, excelled in capturing daily streamflow but failed to adequately simulate monthly ETa. In contrast, Calibration II, using GLEAM ETa, showed improved ETa simulations after adjustment of the biophysical parameters. Calibration III, which included both streamflow and ETa in calibration, proved to be the most viable strategy, producing satisfactory results for both daily streamflow and monthly ETa.

Our findings underscore the limitations of conventional calibration strategies focusing solely on streamflow or ETa. This study aligns with previous research, highlighting the interconnected nature of hydrological processes and advocating for incorporating both point (e.g., streamflow) and spatial (e.g., ETa) variables in calibration. This approach ensures more plausible simulations while effectively resolving the complexities of basin heterogeneity. Furthermore, our analysis reveals critical insights into hydrological model plausibility in the UBNB, highlighting the significance of calibration strategies in enhancing model realism. By adopting a multi-variable calibration approach with SWAT+, our study offers valuable benchmarks for future research endeavours in the UBNB and basins with similar characteristics. From these results, it can also be suggested that the third calibration strategy can be applied in lumped models where land use characteristics are significantly simplified, to improve model plausibility and performance.

Given the findings of this study, it is recommended that model developers incorporate capabilities for multi-variable calibration in hydrological models. This would simplify the calibration process for researchers and hydrological modellers, enhancing the accuracy and robustness of hydrological predictions. Furthermore, enhancements to the model, such as better representation of canopy processes and adjustments to default settings, could address the known issues with ETa simulation, particularly in forest areas. It is suggested that future research focuses on refining calibration strategies, incorporating other variables beyond streamflow and ETa, such as GRACE data, to further investigate model accuracy and address the complexity of calibration of hydrological models in large and heterogeneous basins such as the UBNB, where data is also scarce. and data scarcity. By adopting multi-variable calibration approaches and leveraging advancements in observational datasets and modelling techniques, we can deepen our understanding of basin hydrology and inform more robust water resource management strategies.

CRediT authorship contribution statement

Aseel Mohamed: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Micha Werner:** Writing – review & editing, Supervision, Conceptualization. **Pieter van der Zaag:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study builds upon a thesis submitted to fulfil the requirements for the Master of Science diploma at IHE Delft Institute for Water Education, The Netherlands. The master's program was financially supported by the Sudan Scholarship in the framework of the project: "*Supporting water management in the Sudan through capacity strengthening and applied research*". The authors would also like to thank the Sudanese Ministry of Irrigation and Water Resources for providing us with El Diem streamflow data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ejrh.2025.102290.

Data availability

El Diem streamflow data is available upon request from the Ministry of Irrigation and Water Resources of Sudan. The GLEAM dataset was retrieved from https://www.gleam.eu/. Precipitation data (CHIRPS v2.0) was downloaded from https://www.chc.ucsb.edu/data/chirps. Meteorological data from ERA5 was downloaded from https://cds.climate.copernicus.eu/. The land cover map was retrieved from the Copernicus Global Land Cover at https://land.copernicus.eu/global/products/lc. The harmonised global soil dataset from FAO was downloaded from https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/. The SRTM DEM was downloaded from https://earthexplorer.usgs.gov/.

References

Abbaspour, K.C., 2015. Swat-Cup: swat Calibration and Uncertainty Programs—A User Manual. Eawag, Dbendorf, Switzerland, pp. 16–70.

Abebe, S.A., Qin, T., Zhang, X., Li, C., Yan, D., 2022. Estimating the water budget of the Upper Blue Nile River basin with Water and Energy Processes (WEP) model. Front. Earth Sci. 10, http://dx.doi.org/10.3389/feart.2022.923252.

Abera, W., Formetta, G., Brocca, L., Rigon, R., 2017. Modeling the water budget of the Upper Blue Nile basin using the JGrass-NewAge model system and satellite data. Hydrol. Earth Syst. Sci. 21 (6), 3145–3165. http://dx.doi.org/10.5194/hess-21-3145-2017.

Abiodun, O.O., Guan, H., Post, V.E.A., Batelaan, O., 2018. Comparison of MODIS and SWAT evapotranspiration over a complex terrain at different spatial scales. Hydrol. Earth Syst. Sci. 22 (5), 2775–2794. http://dx.doi.org/10.5194/hess-22-2775-2018.

Alemayehu, T., van Griensven, A., Woldegiorgis, B.T., Bauwens, W., 2017. An improved SWAT vegetation growth module and its evaluation for four tropical ecosystems. Hydrol. Earth Syst. Sci. 21 (9), 4449–4467. http://dx.doi.org/10.5194/hess-21-4449-2017.

Ali, A.M., Melsen, L.A., Teuling, A.J., 2023. Inferring reservoir filling strategies under limited data availability using hydrological modelling and earth observation: the case of the Grand Ethiopian Renaissance Dam (GERD). Hydrol. Earth Syst. Sci. http://dx.doi.org/10.5194/hess-2023-19.

Apostel, A., Kalcic, M., Dagnew, A., Evenson, G., Kast, J., King, K., Martin, J., Muenich, R.L., Scavia, D., 2021. Simulating internal watershed processes using multiple SWAT models. Sci. Total Environ. 759, 143920. http://dx.doi.org/10.1016/j.scitotenv.2020.143920.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: model development 1. JAWRA J. Am. Water Resour. Assoc. 34 (1), 73–89.

Ávila, L., Silveira, R., Campos, A., Rogiski, N., Gonçalves, J., Scortegagna, A., Freita, C., Aver, C., Fan, F., 2022. Comparative evaluation of five hydrological models in a large-scale and Tropical River Basin. Water 14 (19), http://dx.doi.org/10.3390/w14193013.

Ayehu, G.T., Tadesse, T., Gessesse, B., Dinku, T., 2018. Validation of new satellite rainfall products over the Upper Blue Nile Basin, Ethiopia. Atmospheric Meas. Tech. 11 (4), 1921–1936. http://dx.doi.org/10.5194/amt-11-1921-2018.

Bayissa, Y., Tadesse, T., Demisse, G., Shiferaw, A., 2017. Evaluation of satellite-based rainfall estimates and application to monitor meteorological drought for the Upper Blue Nile Basin, Ethiopia. Remote. Sens. 9 (7), http://dx.doi.org/10.3390/rs9070669.

Beven, K., 2006. A manifesto for the equifinality thesis. J. Hydrol. 320 (1–2), 18–36. http://dx.doi.org/10.1016/j.jhydrol.2005.07.007.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249 (1–4), 11–29.

Bieger, K., Arnold, J.G., Rathjens, H., White, M.J., Bosch, D.D., Allen, P.M., Volk, M., Srinivasan, R., 2017. Introduction to SWAT+, A completely restructured version of the soil and water assessment tool. JAWRA J. Am. Water Resour. Assoc. 53 (1), 115–130. http://dx.doi.org/10.1111/1752-1688.12482.

Chavarría, S.B., Werner, M., Salazar, J.F., Vargas, T.B., 2022. Benchmarking global hydrological and land surface models against GRACE in a medium-sized tropical basin. Hydrol. Earth Syst. Sci. 26 (16), 4323–4344. http://dx.doi.org/10.5194/hess-26-4323-2022.

Chawanda, C.J., George, C., Thiery, W., Griensven, A.v., Tech, J., Arnold, J., Srinivasan, R., 2020. User-friendly workflows for catchment modelling: Towards reproducible SWAT+ model studies. Environ. Model. Softw. 134, http://dx.doi.org/10.1016/j.envsoft.2020.104812.

Cinkus, G., Mazzilli, N., Jourde, H., Wunsch, A., Liesch, T., Ravbar, N., Chen, Z., Goldscheider, N., 2023. When best is the enemy of good – critical evaluation of performance criteria in hydrological models. Hydrol. Earth Syst. Sci. 27 (13), 2397–2411. http://dx.doi.org/10.5194/hess-27-2397-2023.

Conway, D., 1997. A water balance model of the Upper Blue Nile in Ethiopia. Hydrol. Sci. J. 42 (2), 265–286. http://dx.doi.org/10.1080/02626669709492024.

Daggupati, P., Yen, H., White, M.J., Srinivasan, R., Arnold, J.G., Keitzer, C.S., Sowa, S.P., 2015. Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. Hydrol. Process. 29 (26), 5307–5320. http://dx.doi.org/10.1002/hyp.10536.

Dahmen, E., Hall, M.J., 1990. Screening of Hydrological Data: Tests for Stationarity and Relative Consistency. ILRI.

Dash, S.S., Sahoo, B., Raghuwanshi, N.S., 2021. How reliable are the evapotranspiration estimates by Soil and Water Assessment Tool (SWAT) and Variable Infiltration Capacity (VIC) models for catchment-scale drought assessment and irrigation planning? J. Hydrol. 592, http://dx.doi.org/10.1016/j.jhydrol.2020.125838.

Dembélé, M., Hrachowitz, M., Savenije, H.H.G., Mariéthoz, G., Schaefli, B., 2020. Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite data sets. Water Resour. Res. 56 (1), http://dx.doi.org/10.1029/2019wr026085.

Demirel, M.C., Koch, J., Rakovec, O., Kumar, R., Mai, J., Müller, S., Thober, S., Samaniego, L., Stisen, S., 2024. Tradeoffs between temporal and spatial pattern calibration and their impacts on robustness and transferability of hydrologic model parameters to ungauged basins. Water Resour. Res. 60 (1), e2022WR034193. http://dx.doi.org/10.1029/2022WR034193, e2022WR034193 2022WR034193.

Demirel, M.C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., Stisen, S., 2018. Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model. Hydrol. Earth Syst. Sci. 22 (2), 1299–1315. http://dx.doi.org/10.5194/hess-22-1299-2018.

Dile, Y.T., Ayana, E.K., Worqlul, A.W., Xie, H., Srinivasan, R., Lefore, N., You, L., Clarke, N., 2020. Evaluating satellite-based evapotranspiration estimates for hydrological applications in data-scarce regions: A case in Ethiopia. Sci. Total Environ. 743, 140702. http://dx.doi.org/10.1016/j.scitotenv.2020.140702.

Dile, Y., Srinivasan, R., George, C., 2023. QGIS Interface for SWAT+: QSWAT+. SWAT.

Dile, Y.T., Tekleab, S., Ayana, E.K., Gebrehiwot, S.G., Worqlul, A.W., Bayabil, H.K., Yimam, Y.T., Tilahun, S.A., Daggupati, P., Karlberg, L., Srinivasan, R., 2018. Advances in water resources research in the Upper Blue Nile basin and the way forward: A review. J. Hydrol. 560, 407–423. http://dx.doi.org/10.1016/j.jhydrol.2018.03.042.

Eini, M.R., Massari, C., Piniewski, M., 2023. Satellite-based soil moisture enhances the reliability of agro-hydrological modeling in large transboundary river basins. Sci. Total Environ. 873, 162396. http://dx.doi.org/10.1016/j.scitotenv.2023.162396.

ENTRO, 2007. Transboundary Analysis: Abay-Blue Nile Sub-basin. Report, Eastern Nile Technical Regional Office.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The shuttle radar topography mission. Rev. Geophys. 45 (2), http://dx.doi.org/10.1029/2005rg000183.

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H., Verelst, L., Wiberg, D., 2008. Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008). Vol. 10, IIASA, Laxenburg, Austria FAO, Rome, Italy.

Funk, C.C., Peterson, P.J., Landsfeld, M.F., Pedreros, D.H., Verdin, J.P., Rowland, J.D., Romero, B.E., Husak, G.J., Michaelsen, J.C., Verdin, A.P., 2014. A quasi-global precipitation time series for drought monitoring. US Geol. Surv. Data Ser. 832 (4), 1–12.

Gao, H., Sabo, J.L., Chen, X., Liu, Z., Yang, Z., Ren, Z., Liu, M., 2018. Landscape heterogeneity and hydrological processes: a review of landscape-based hydrological models. Landsc. Ecol. 33 (9), 1461–1480. http://dx.doi.org/10.1007/s10980-018-0690-4.

Gebremicael, T.G., Mohamed, Y.A., Betrie, G.D., van der Zaag, P., Teferi, E., 2013. Trend analysis of runoff and sediment fluxes in the Upper Blue Nile basin: A combined analysis of statistical tests, physically-based models and landuse maps. J. Hydrol. 482, 57–68. http://dx.doi.org/10.1016/j.jhydrol.2012.12.023.

Getahun, M., Selassie, Y.G., 2017. Characterization, classification and mapping of soils of agricultural landscape in Tana Basin, Amhara National Regional State, Ethiopia. In: Social and Ecological System Dynamics: Characteristics, Trends, and Integration in the Lake Tana Basin, Ethiopia. pp. 93–115.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91. http://dx.doi.org/10.1016/j.jhydrol.2009.08.003.

Haas, H., Kalin, L., Yen, H., 2024. Improved forest canopy evaporation leads to better predictions of ecohydrological processes. Ecol. Model. 489, http://dx.doi.org/10.1016/j.ecolmodel.2024.110620.

Haas, H., Reaver, N.G.F., Karki, R., Kalin, L., Srivastava, P., Kaplan, D.A., Gonzalez-Benecke, C., 2022. Improving the representation of forests in hydrological models. Sci. Total. Env. 812, 151425. http://dx.doi.org/10.1016/j.scitotenv.2021.151425.

Hassan, A.M., Hassan, M., Mohamed, Y., Awad, W., 2024. Dam operation using satellite data and hydrological models: the case of Roseires dam and Grand Ethiopian Renaissance Dam in the Blue Nile River. Water Int. 48, 1–25. http://dx.doi.org/10.1080/02508060.2023.2286412.

Herman, M.R., Nejadhashemi, A.P., Abouali, M., Hernandez-Suarez, J.S., Daneshvar, F., Zhang, Z., Anderson, M.C., Sadeghi, A.M., Hain, C.R., Sharifi, A., 2018. Evaluating the role of evapotranspiration remote sensing data in improving hydrological modeling predictability. J. Hydrol. 556, 39–49. http://dx.doi.org/10.1016/j.jhydrol.2017.11.009.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Sabater, J.M., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., Dee, D., 2019. Global reanalysis: goodbye ERA-Interim, hello ERA5. ECMWF Newsl. 159, http://dx.doi.org/10.21957/vf291hehd7.

Kebede, S., Travi, Y., Alemayehu, T., Ayenew, T., 2005. Groundwater recharge, circulation and geochemical evolution in the source region of the Blue Nile River, Ethiopia. Appl. Geochem. 20 (9), 1658–1676. http://dx.doi.org/10.1016/j.apgeochem.2005.04.016.

Koch, J., Demirel, M.C., Stisen, S., 2018. The SPAtial EFficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models. Geosci. Model. Dev. 11 (5), 1873–1886. http://dx.doi.org/10.5194/gmd-11-1873-2018.

Lazin, R., Shen, X., Koukoula, M., Anagnostou, E., 2020. Evaluation of the hyper-resolution model-derived water cycle components over the Upper Blue Nile Basin. J. Hydrol. 590, http://dx.doi.org/10.1016/j.jhydrol.2020.125231.

Legesse Gebre, S., 2015. Hydrological response to climate change of the Upper Blue Nile River basin: Based on IPCC fifth assessment report (AR5). J. Clim. Weather. Forecast. 03 (01), http://dx.doi.org/10.4172/2332-2594.1000121.

Lemann, T., Zeleke, G., Amsler, C., Giovanoli, L., Suter, H., Roth, V., 2016. Modelling the effect of soil and water conservation on discharge and sediment yield in the upper Blue Nile basin, Ethiopia. Appl. Geogr. 73, 89–101. http://dx.doi.org/10.1016/j.apgeog.2016.06.008.

Lemma, E., Upadhyaya, S., Ramsankaran, R., 2019. Investigating the performance of satellite and reanalysis rainfall products at monthly timescales across different rainfall regimes of Ethiopia. Int. J. Remote Sens. 40 (10), 4019–4042. http://dx.doi.org/10.1080/01431161.2018.1558373.

López López, P., Sutanudjaja, E.H., Schellekens, J., Sterk, G., Bierkens, M.F.P., 2017. Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. Hydrol. Earth Syst. Sci. 21 (6), 3125–3144. http://dx.doi.org/10.5194/hess-21-3125-2017.

Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., Verhoest, N.E.C., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model. Dev. 10 (5), 1903–1925. http://dx.doi.org/10.5194/gmd-10-1903-2017.

McColl, K.A., Vogelzang, J., Konings, A.G., Entekhabi, D., Piles, M., Stoffelen, A., 2014. Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. Geophys. Res. Lett. 41 (17), 6229–6236. http://dx.doi.org/10.1002/2014gl061322.

Mei, Y., Mai, J., Do, H.X., Gronewold, A., Reeves, H., Eberts, S., Niswonger, R., Regan, R.S., Hunt, R.J., 2023. Can hydrological models benefit from using global soil moisture, evapotranspiration, and runoff products as calibration targets? Water Resour. Res. 59 (2), http://dx.doi.org/10.1029/2022wr032064.

Mekonnen, D.F., Duan, Z., Rientjes, T., Disse, M., 2018. Analysis of combined and isolated effects of land-use and land-cover changes and climate change on the upper Blue Nile River basin's streamflow. Hydrol. Earth Syst. Sci. 22 (12), 6187–6207. http://dx.doi.org/10.5194/hess-22-6187-2018.

Mengistu, D., Bewket, W., Dosio, A., Panitz, H.-J., 2021. Climate change impacts on water resources in the Upper Blue Nile (Abay) River Basin, Ethiopia. J. Hydrol. 592, http://dx.doi.org/10.1016/j.jhydrol.2020.125614.

Mohr, K.I., Famiglietti, J.S., Boone, A., Starks, P.J., 2000. Modeling soil moisture and surface flux variability with an untuned land surface scheme: A case study from the Southern Great Plains 1997 hydrology experiment. J. Hydrometeorol. 1 (2), 154–169.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. J. Hydrol. 10 (3), 282–290. http://dx.doi.org/10.1016/0022-1694(70)90255-6.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009. Report, Texas Water Resources Institute.

Nguyen, T.V., Dietrich, J., Dang, T.D., Tran, D.A., Van Doan, B., Sarrazin, F.J., Abbaspour, K., Srinivasan, R., 2022. An interactive graphical interface tool for parameter calibration, sensitivity analysis, uncertainty analysis, and visualization for the Soil and Water Assessment Tool. Environ. Model. Softw. 156, 105497.

Nkwasa, A., Chawanda, C.J., van Griensven, A., 2022. Regionalization of the SWAT+ model for projecting climate change impacts on sediment yield: An application in the Nile basin. J. Hydrol.: Reg. Stud. 42, 101152. http://dx.doi.org/10.1016/j.ejrh.2022.101152.

Nooni, I.K., Wang, G., Hagan, D.F.T., Lu, J., Ullah, W., Li, S., 2019. Evapotranspiration and its components in the Nile River Basin based on long-term satellite assimilation product. Water 11 (7), http://dx.doi.org/10.3390/w11071400.

Onyutha, C., 2019. Hydrological model supported by a step-wise calibration against sub-flows and validation of extreme flow events. Water 11 (2), 244.

Pricope, N.G., Husak, G., Lopez-Carr, D., Funk, C., Michaelsen, J., 2013. The climate-population nexus in the East African Horn: Emerging degradation trends in rangeland and pastoral livelihood zones. Glob. Environ. Chang. 23 (6), 1525–1541. http://dx.doi.org/10.1016/j.gloenvcha.2013.10.002.

Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92.

Rajib, A., Evenson, G.R., Golden, H.E., Lane, C.R., 2018. Hydrologic model predictability improves with spatially explicit calibration using remotely sensed evapotranspiration and biophysical parameters. J. Hydrol. 567, 668–683. http://dx.doi.org/10.1016/j.jhydrol.2018.10.024.

Rane, N.L., Jayaraj, G.K., 2022. Enhancing SWAT model predictivity using multi-objective calibration: effects of integrating remotely sensed evapotranspiration and leaf area index. Int. J. Environ. Sci. Technol. 20 (6), 6449–6468. http://dx.doi.org/10.1007/s13762-022-04293-7.

Rodell, M., Famiglietti, J.S., 2001. An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE). Water Resour. Res. 37 (5), 1327–1339. http://dx.doi.org/10.1029/2000WR900306.

Rogelis, M.C., Werner, M., Obregón, N., Wright, N., 2016. Hydrological model assessment for flood early warning in a tropical high mountain basin. Hydrol. Earth Syst. Sci. 2016, 1–36. http://dx.doi.org/10.5194/hess-2016-30.

Salter, P.J., Williams, J.B., 1967. The influence of texture on the moisture characteristics of soils. J. Soil Sci. 18 (1), 174–181. http://dx.doi.org/10.1111/j.1365-2389.1967.tb01498.x.

Sánchez-Gómez, A., Schürz, C., Molina-Navarro, E., Bieger, K., 2024. Groundwater modelling in SWAT+: Considerations for a realistic baseflow simulation. Groundw. Sustain. Dev. 26, 101275. http://dx.doi.org/10.1016/j.gsd.2024.101275.

Saxton, K.E., Rawls, W.J., 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Sci. Am. J. 70 (5), 1569–1578. http://dx.doi.org/10.2136/sssaj2005.0117.

Shah, S., Duan, Z., Song, X., Li, R., Mao, H., Liu, J., Ma, T., Wang, M., 2021. Evaluating the added value of multi-variable calibration of SWAT with remotely sensed evapotranspiration data for improving hydrological modeling. J. Hydrol. 603, 127046.

Song, Y., Zhang, J., Lai, Y., 2021. Influence of multisite calibration on streamflow estimation based on the hydrological model with CMADS inputs. J. Water Clim. Chang. 12 (7), 3264–3281. http://dx.doi.org/10.2166/wcc.2021.115.

Srivastava, A., Kumari, N., Maza, M., 2020. Hydrological response to agricultural land use heterogeneity using variable infiltration capacity model. Water Resour. Manag. 34 (12), 3779–3794. http://dx.doi.org/10.1007/s11269-020-02630-4.

Steenhuis, T.S., Collick, A.S., Easton, Z.M., Leggesse, E.S., Bayabil, H.K., White, E.D., Awulachew, S.B., Adgo, E., Ahmed, A.A., 2009. Predicting discharge and sediment for the Abay (Blue Nile) with a simple model. Hydrol. Process. 3728–3737. http://dx.doi.org/10.1002/hyp.7513.

Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. J. Geophys. Res.: Ocean. 103 (C4), 7755–7766. http://dx.doi.org/10.1029/97jc03180.

Sultana, R., Dhungana, S., Bhatta, N., 2019. Application of SWAT model for estimating runoff in Upper Nile River basin. Eur. J. Eng. Sci. Technol. 2 (4), 20–35.

Swain, A., 2011. Challenges for water sharing in the Nile basin: changing geo-politics and changing climate. Hydrol. Sci. J. 56 (4), 687–702. http://dx.doi.org/10.1080/02626667.2011.577037.

SWAT, 2023. SWAT+ Documentation. Texas A&M University, URL: https://swatplus.gitbook.io/docs/.

Takele, G.S., Gebre, G.S., Gebremariam, A.G., Engida, A.N., 2021. Hydrological modeling in the Upper Blue Nile basin using soil and water analysis tool (SWAT). Model. Earth Syst. Environ. 8 (1), 277–292. http://dx.doi.org/10.1007/s40808-021-01085-9.

Taye, M.T., Willems, P., Block, P., 2015. Implications of climate change on hydrological extremes in the Blue Nile basin: A review. J. Hydrol.: Reg. Stud. 4, 280–293. http://dx.doi.org/10.1016/j.ejrh.2015.07.001.

Tegegne, G., Park, D.K., Kim, Y.-O., 2017. Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin. J. Hydrol.: Reg. Stud. 14, 49–66. http://dx.doi.org/10.1016/j.ejrh.2017.10.002.

Tesemma, Z.K., Mohamed, Y.A., Steenhuis, T.S., 2010. Trends in rainfall and runoff in the Blue Nile Basin: 1964–2003. Hydrol. Process. 24 (25), 3747–3758. http://dx.doi.org/10.1002/hyp.7893.

Tigabu, T.B., Wagner, P.D., Hörmann, G., Fohrer, N., 2020. Modeling the spatio-temporal flow dynamics of groundwater-surface water interactions of the Lake Tana Basin, Upper Blue Nile, Ethiopia. Hydrol. Res. 51 (6), 1537–1559. http://dx.doi.org/10.2166/nh.2020.046.

Tsendbazar, N.-E., Tarko, A., Li, L., Herold, M., Lesiv, M., Fritz, S., Maus, V., 2021. Copernicus global land service: land cover 100m: Version 3 Globe 2015–2019: validation report.

Van Griensven, A., Ndomba, P., Yalew, S., Kilonzo, F., 2012. Critical review of SWAT applications in the upper Nile basin countries. Hydrol. Earth Syst. Sci. 16 (9), 3371–3381. http://dx.doi.org/10.5194/hess-16-3371-2012.

Venneker, R., 2011. DScreen User Manual. UNESCO-IHE, Delft, The Netherlands.

Weerasinghe, I., Bastiaanssen, W., Mul, M., Jia, L., van Griensven, A., 2020. Can we trust remote sensing evapotranspiration products over Africa? Hydrol. Earth Syst. Sci. 24 (3), 1565–1586. http://dx.doi.org/10.5194/hess-24-1565-2020.

Wi, S., Yang, Y.C.E., Steinschneider, S., Khalil, A., Brown, C.M., 2015. Calibration approaches for distributed hydrologic models in poorly gaged basins: implication for streamflow projections under climate change. Hydrol. Earth Syst. Sci. 19 (2), 857–876. http://dx.doi.org/10.5194/hess-19-857-2015.

Xue, X., Zhang, K., Hong, Y., Gourley, J.J., Kellogg, W., McPherson, R.A., Wan, Z., Austin, B.N., 2016. New multisite cascading calibration approach for hydrological models: Case study in the Red River Basin using the VIC model. J. Hydrol. Eng. 21 (2), http://dx.doi.org/10.1061/(asce)he.1943-5584.0001282.

Yang, Q., Almendinger, J.E., Zhang, X., Huang, M., Chen, X., Leng, G., Zhou, Y., Zhao, K., Asrar, G.R., Srinivasan, R., Li, X., 2018. Enhancing SWAT simulation of forest ecosystems for water resource assessment: A case study in the St. Croix River basin. Ecol. Eng. 120, 422–431. http://dx.doi.org/10.1016/j.ecoleng.2018.06.020.

Yang, Q., Zhang, X., 2016. Improving SWAT for simulating water and carbon fluxes of forest ecosystems. Sci. Total Environ. 569–570, 1478–1488. http://dx.doi.org/10.1016/j.scitotenv.2016.06.238.

Yen, H., Park, S., Arnold, J.G., Srinivasan, R., Chawanda, C.J., Wang, R., Feng, Q., Wu, J., Miao, C., Bieger, K., Daggupati, P., Griensven, A.v., Kalin, L., Lee, S., Sheshukov, A.Y., White, M.J., Yuan, Y., Yeo, I.-Y., Zhang, M., Zhang, X., 2019. IPEAT+: A built-in optimization and automatic calibration tool of SWAT+. Water 11 (8), http://dx.doi.org/10.3390/w11081681.

Zeri, M., S. Alvalá, R., Carneiro, R., Cunha-Zeri, G., Costa, J., Rossato Spatafora, L., Urbano, D., Vall-Llossera, M., Marengo, J., 2018. Tools for communicating agricultural drought over the Brazilian semiarid using the soil moisture index. Water 10 (10), http://dx.doi.org/10.3390/w10101421.

Zhang, L., Zhao, Y., Ma, Q., Wang, P., Ge, Y., Yu, W., 2021. A parallel computing-based and spatially stepwise strategy for constraining a semi-distributed hydrological model with streamflow observations and satellite-based evapotranspiration. J. Hydrol. 599, http://dx.doi.org/10.1016/j.jhydrol.2021.126359.