Inferring reservoir filling strategies under limited data availability using hydrological modelling and Earth observation: the case of the Grand Ethiopian Renaissance Dam (GERD)

Awad M. Ali[1,2], Lieke A. Melsen[1], and Adriaan J. Teuling[1]

[1]Hydrology and Quantitative Water Management Group, Department of Environmental Sciences, Wageningen University & Research, P.O. BOX 47, 6700 AA Wageningen, the Netherlands
[2]Water Research Center, Faculty of Engineering, University of Khartoum, P.O. BOX 321, Khartoum, Sudan
Correspondence: Awad M. Ali (awad.negmeldinawad.mohammedali@wur.nl)

Abstract. The filling of the Grand Ethiopian Renaissance Dam (GERD) started in 2020, posing additional challenges for downstream water management in the Blue Nile River in Sudan, which is already struggling to cope with the effects of climate change. This is also the case for many transboundary rivers that observe a lack of cooperation and transparency during the filling and operation of new dams. Without information about water supply from neighbouring countries, it is risky to manage

5 downstream dams as usual and operation information is needed to apply modifications. This study aims to ~~test the applicability of using lumped hydrological modelling coupled~~ develop a novel approach/framework utilizing hydrological modelling in conjunction with remote sensing data ~~in retrieving~~ to retrieve reservoir filling strategies ~~in regions with~~ under limited data availability. Firstly, five rainfall products (namely; ARC2, CHIRPS, ERA5, GPCC, and PERSIANN-CDR) were evaluated against historical measured rainfall at ten stations. Secondly, to account for input uncertainty, the best three performing rainfall

10 products were forced in the conceptual hydrological model HBV-light with potential evapotranspiration and temperature data from ERA5. The model was calibrated during the period 2006 - 2019 and validated during the period 1991 - 1996. Thirdly, the parameter sets that obtained very good performance (NSE > 0.75) were utilized to predict the inflow of GERD during the operation period (2020 - 2022). Then, from the water balance of GERD, the daily storage was estimated and compared with the storage derived from Landsat ~~observations~~ and Sentinel imageries to evaluate the performance of the selected rainfall products

15 and the reliability of the framework. Finally, three years of GERD filling strategies were retrieved using the best-performing simulation of CHIRPS with RMSE of 1.7 and 1.52 billion cubic meters (BCM) and NSE of 0.77 and 0.86 when compared with ~~Landsat-derived reservoir storage~~ Landsat- and Sentinel-derived reservoir storage, respectively. It was found that GERD stored 14% of the monthly inflow of July 2020, 41% of July 2021, and 37% and 32% of July and August 2022, respectively. Annually, GERD retained 5.2% and 7.4% of the annual inflow in the first two filling phases and between 12.9% and 13.7% in the third

20 phase. The results also revealed that the retrieval of filling strategies is more influenced by input uncertainty than parameter uncertainty. The retrieved daily change in GERD storage with the measured outflow to Sudan allowed further interpretation of the downstream impacts of GERD. The findings of this study provide systematic steps to retrieve filling strategies ~~for data-scarce regions~~, which can serve as a base for future development in the field, especially for data-scarce regions. Locally, the analysis contributes significantly to the future water management of the Roseires and Sennar dams in Sudan.

1 Introduction

The optimal management of water resources is challenging as we increasingly experience the impacts of climate change and increasing water demand due to population growth (Stakhiv, 2011). Transboundary rivers pose even greater challenges to water management, for example, when structures such as dams are built by upstream countries, impacting downstream flow regime (Biswas, 2008) and mean water availability during the filling phase. The absence of collaboration has disastrous impacts in these regions, as collaboration increases water revenues and reduces the risk of extreme events (Basheer et al., 2018; Wheeler et al., 2018). There is, however, the possibility of conflicting dynamics in many transboundary basins due to the different policies and strategies between neighbouring countries (Warner and Zawahri, 2012). As a result, a lack of transparency on the filling/operation of dams and other infrastructures is often witnessed.

The Blue Nile River is an example of a transboundary river where infrastructure influences downstream dynamics. A current challenge faced by the Ministry of Irrigation and Water Resources of Sudan (MoIWR) is the management of the Nile River. Around two-thirds of Nile water is supplied by the Blue Nile River (Dumont, 1986). This river originates in Lake Tana, Ethiopia, and subsequently flows northward to meet the White Nile River (originates in Lake Victoria, Uganda), where the Main Nile River starts in Khartoum, the capital of Sudan, as shown in Fig. 1.A. In April 2011, the Ethiopian government started the construction of the Grand Ethiopian Renaissance Dam (GERD) in the Upper Blue Nile (UBN) basin, around 15 km east of the Ethiopia-Sudan border. The dam, when completed, will be the largest hydropower dam in Africa with a storage capacity of 74 billion cubic meters (BCM) and a power capacity of 5150 Megawatts (MW) (Ezega News, 2019). It was constructed without an agreement with downstream countries (i.e., Sudan and Egypt), which caused hydro-political tension between Sudan, Egypt, and Ethiopia (Gebrehiwet, 2020). Moreover, water management in Sudan is very closely tied to reservoirs. The Roseires dam, for example, plays a significant role in managing drinking water as well as irrigation water for downstream projects and is closely coordinated with the Sennar dam downstream (Alrajoula et al., 2016) (see Fig. 1.B). Around 35% of the Sudan allocation of Nile water is consumed by the Gezira scheme, which is supplied by the Sennar dam (Adam et al., 2003). The Gezira scheme, with a total area of 8800 km^2, is regarded as one of the largest irrigation schemes in the world under a single management structure (Ahmed et al., 2006). It, therefore, plays a considerable role in the socio-economy of Sudan.

The construction of GERD, in combination with a low level of coordination between Ethiopia and Sudan, poses a threat to water management in Sudan. For instance, floods are frequently experienced in Khartoum state, where 16% of the population lives (World Population Review, 2022). In the past 35 years, flood events were observed, on average, every 5 years. The latest flood of 2020 was unprecedented, where water levels exceeded all previous events (NASA Earth Observatory, 2020). In the same year, the filling of GERD started without sharing information with Sudan and Egypt. As shown in Fig. 1.B, the Roseires dam is located only around 120 km downstream of GERD. Without knowledge of GERD operation strategies, management of the Roseires dam is a considerable challenge. Knowledge of operation strategies is thus essential for flow prediction and water management in downstream areas.

It is clear that the operation of the dams along the Blue Nile River has significant implications for food security, flood protection, water availability, and hydropower generation downstream in the Main Nile River. Nonetheless, Ethiopia has yet to
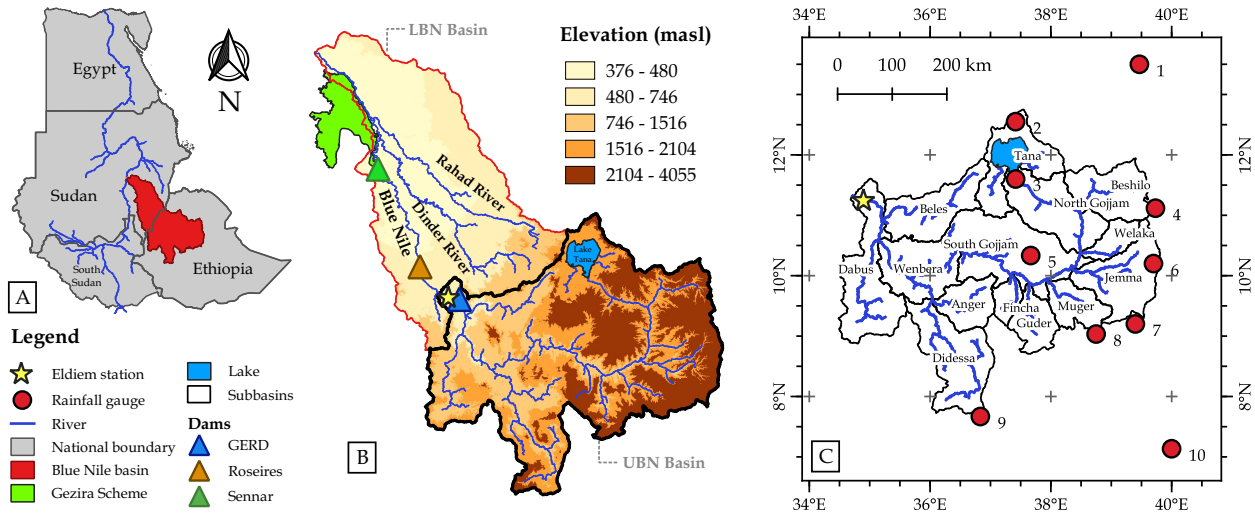
Figure 1. Study area overview: (A) The geographical location of the Blue Nile Basin with the Upper and Lower parts, and (B) a Digital Elevation Model (SRTM 30m) map of the Blue Nile Basin and location of the Gezira Irrigation Scheme, dams and the Eldiem streamflow station, (C) Upper Blue Nile subbasins with the location of the rainfall gauge stations. Basins boundaries and drainage network are obtained from the HydroSHEDS dataset (Lehner et al., 2008) and national boundaries were based on the Database of Global Administrative Areas (GADM, 2012). The name and coordinates of each rainfall gauge station are demonstrated in Table 3.

share information on GERD filling, thus hindering water management in Sudan. Therefore, obtaining filling/operation information using alternative means is urgently needed to support the operation of both; the Roseires and Sennar dams. Different approaches exist to understand dam operation and reservoir water storage. One approach relies on satellite remote sensing and radar altimetry. At a global scale, satellite-derived water heights and extents were used to reconstruct storage dynamics (Hou et al., 2022). Vu et al. (2022) revealed reservoir filling strategies and operating rules of ten dams along the Upper Mekong River using Landsat images & radar altimetry. Another approach is using a process-based hydrological model to simulate water inflow, storage, and release. Eldardiry and Hossain (2019) predicted reservoir operating rules of High Aswan Dam ~~(HAD)~~ in the lower Nile River basin using satellite hydrometeorological observation and a macroscale hydrologic model. Also, Wannasin et al. (2021) simulated the daily storages of two reservoirs using a distributed hydrological model and a reservoir operation module.

In this study, ~~the quantification will be based on the latter approach for three main reasons~~ we aim to develop a novel approach to infer reservoir management strategies under limited data availability. In our approach, two techniques, hydrological modelling and Earth observation, are combined to retrieve reliable reservoir records. Then, we reconstruct management strategies based on further analysis of the retrieved records. The merits of this approach include: (1) ~~Roseires dam is very close to GERD, thus hydrological modelling can support proactive flow prediction for real-life operations~~ facilitating proactive flow forecasting for real-time operations, (2) ~~Finer~~ generating refined and more consistent ~~time steps can be achieved when~~

75 using hydrological models reservoir data, thereby supporting hydrological analyses at various temporal scales, and (3) The impact of rainfall uncertainty on inferring filling strategies can be tested. However, the first approach will be utilized in the validation stage as will be detailed later. permitting the assessment of the influence of different input uncertainties on the inference of the management strategies.

This study aims to test the applicability of lumped hydrological modelling in large basins to retrieve reservoir-filling 80 strategieswith limited data availability. As part of this study, we will analyze the implications of input data and parameter uncertainty in inferring the filling strategies. As a case study, the research will be conducted on the UBN basin to understand the filling stages of GERD and its impact on downstream discharge. Specifically, the study will quantify the volume of water stored behind the GERD based on known outgoing flow to Sudan as well as the daily reservoir storage change. The output of this study intends to provide information to support the management of the Lower Blue Nile River in Sudan. Nonetheless, the 85 work does not attempt to use the findings to optimize current management practices since this would go beyond the goal of this study. Therefore, we will aim to answer the following questions:

1. How can hydrological modelling and remote sensing be leveraged to effectively infer GERD management strategies?

2. What is the implication of rainfall selection and parameter uncertainty on the inference of GERD management strategies?

90 3. What is the impact of GERD reservoir filling on the flow to Sudan during the first three phases (2020 - 2022)?

The outline of the study is as follows: the study area and data are briefly described in Section 2. The methodology, including the evaluation of selected rainfall products, the use of hydrological modelling, and the inference of reservoir storage, are detailed in Section 3. In Section 4, the results of the study are presented and discussed in Section 5. Finally, the conclusion of the work done is summarized in Section 6.

95 Study area overview: (A) The geographical location of the Blue Nile Basin with the Upper and Lower parts, and (B) a Digital Elevation Model (SRTM 30m) map of the Blue Nile Basin and location of the Gezira Irrigation Scheme, dams and the Eldiem streamflow station, (C) Upper Blue Nile subbasins with the location of the rainfall gauge stations. Basins boundaries and drainage network are obtained from the HydroSHEDS dataset (Lehner et al., 2008) and national boundaries were based on the Database of Global Administrative Areas (GADM, 2012). The name and coordinates of each rainfall gauge station are 100 demonstrated in Table 3.

2 Study area and data

2.1 Upper Blue Nile basin

The Blue Nile River basin has a drainage area of 310,000 km^2 that is shared between Ethiopia (64%) and Sudan (36%). From lake Tana, in the north-western Ethiopian Highlands, at 1780 m a.s.l the water is supplied and flows clockwise through the 105 eastern mountainous side of the basin to the western lower altitude side. A few kilometres after the Ethiopia-Sudan border,

the water is measured at the outlet of the ~~Upper Blue Nile (UBN)~~ UBN basin at Eldiem station (481 m a.s.l). Afterward, the water is drained from the Lower Blue Nile (LBN) basin through the Blue Nile River and its tributaries; the Dinder River and the Rahad River, to the Main Nile River as shown in Fig. 1.B.

The UBN basin will be the area of interest for this study which covers an area of around 176,000 km^2. The basin is subdivided into 14 sub-basins as shown in Fig. 1.C. The rainfall of the UBN basin ranges from 1000 mm yr^{-1} near the Ethiopia-Sudan border to 2200 mm yr^{-1} in the Didessa and Dabus subbasins, reflecting the high spatial variability of precipitation associated with topographical features (i.e., so-called water towers) that is typical for much of Eastern Africa (Wamucii et al., 2021). There is a high, but spatially variable, mean potential evapotranspiration, which varies between 1000 and 1800 mm yr^{-1} (Conway, 2000). A typical annual air temperature in the basin ranges from 13°C to 26°C (Tekleab et al., 2013). The discharge regime of the Blue Nile River is highly seasonal, with more than 80% of its annual discharge occurring between July and October and only 4% between January and April (Kim and Kaluarachchi, 2009; Sutcliffe et al., 1999).

Four main land uses dominate in the UBN basin, which are cropland (54.74%), open forest (23.62%), shrubland (8.29%), and closed forest (7.94%) while, the remaining land uses cover only 5.41% of the total area based on the Copernicus Global Land Service ~~(CGLS)~~ (Buchhorn et al., 2020). The large basin is composed of heterogeneous soil types. According to Soil-GRIDS dataset (Hengl et al., 2017), the most common types are Leptosols, Luvisols, Vertisols, and Nitisols, respectively, which collectively occupy 96% of the area.

2.2 Ground measurements

The scarcity of ground measurements in terms of availability and accessibility is a challenge and motivation for this study. Streamflow stations are scattered in the basin and cover many subbasins, but they are not publicly available. Downstream GERD, and after the border, Sudan measures the incoming water at Eldiem station since the 1960s (Fig. 1.B). This data was obtained from the MoIWR for the period 1990-2022 with a daily timestep. Consequently, the calibration and validation of the hydrological model will be based on these measurements as will be detailed later. It is worth mentioning that the discharge at Eldiem station is estimated from measured water levels using rating curves. From 2012 onwards, the Eldiem Station is influenced by the backwater effect of the Roseires dam during filling months (June - October). During these months, the MoIWR estimates the discharge from the water balance of the Roseires dam.

Despite the availability of more than 1200 meteorological stations in Ethiopia, the existing stations within the UBN basin are not sufficient to cover the spatial variability of the rainfall. Additionally, the measured rainfall data is not freely available from the National Meteorological Agency ~~(NMA)~~ of Ethiopia. However, historical monthly rainfall measurements at ten stations were obtained from the ~~NMA~~ National Meteorological Agency as indicated in Fig. 1.C and Table 1. Therefore, rainfall measurements will be used to validate satellite and reanalysis rainfall products available for this region, as remote sensing products cover the entire basin and recent data are available.

Table 1. Main characteristics of the available ground measurements.

Data Type	Temporal Resolution	No. of Stations	Available Years
Discharge	Daily	1	1990 - 2022
Rainfall	Monthly	8	1984 - 2005
Rainfall	Monthly	2 (St. 6&7)*	1993 - 1999

* See Fig. 1.C

Table 2. Summary of the remote sensing products used as inputs in the hydrological modelling over the Upper Blue Nile (UBN) basin.

Input	Dataset	Spatial resolution	Reference
Satellite imagery	Landsat	30 m	(Survey, 2022)
	Sentinel	10 m	(Drusch et al., 2012)
Potential Evapotranspiration	ERA5	0.25°	(Hersbach et al., 2019)
Temperature	ERA5	0.25°	(Hersbach et al., 2019)
Precipitation	ARC 2.0	0.1°	(Novella and Thiaw, 2013)
	CHIRPS	0.05°	(Funk et al., 2014)
	ERA5	0.25°	(Hersbach et al., 2019)
	GPCC	1.0°	(Ziese et al., 2020)
	PERSIANN-CDR	0.25°	(Ashouri et al., 2015)

2.3 Input data

Due to the unavailability of the required observed input data, remote-sensing products will be used in this study. These data include meteorological forcing data (precipitation, potential evapotranspiration, and temperature), and satellite imagery (Landsat and Sentinel). The spatial resolution and references of the selected datasets are summarized in Table 2 and will be utilized for different purposes which are detailed in the following sections.

In this study, the quality of five rainfall products will be tested including the African Rainfall Climatology version 2 (ARC2), the Climate Hazards Group InfraRed Precipitation with Station Version 2 (CHIRPS), the fifth-generation ECMWF atmospheric reanalysis (ERA5), the Global Precipitation Climatology Centre full data reanalysis version 7 (GPCC), and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR). The potential evaporation and temperature data are retrieved from ERA5.
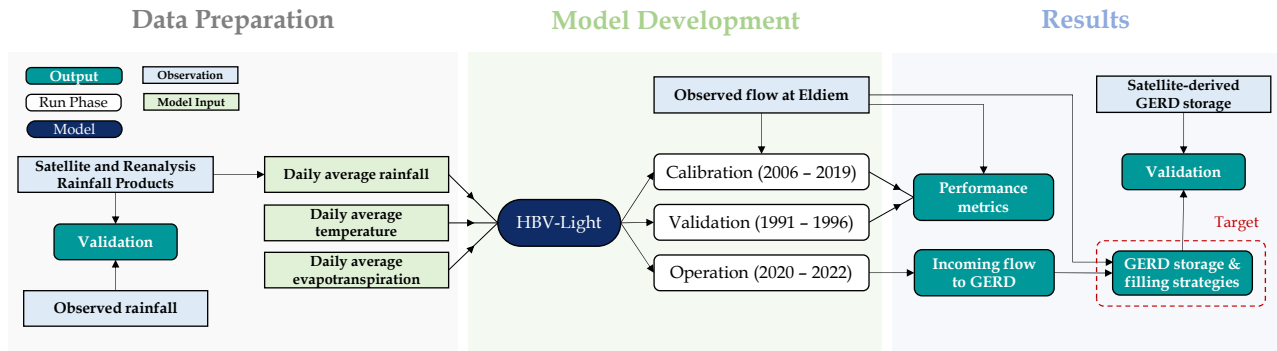
Figure 2. Flowchart representing research methodology. Steps are shown in chronological order, the execution is carried out from left to right. For rainfall product validation five products are selected which are ARC 2.0, CHIRPS 2.0, ERA5, GPCC, and PERSIANN-CDR. For storage validation, observed storage will be computed by converting the water surface area observed by Landsat/Sentinel to water volume using an Elevation-Area-Storage relationship of the GERD reservoir.

3 Methodology

3.1 Evaluation of selected rainfall products

As rainfall products will be used as forcing data, it is important to understand how these data perform in comparison with the available ground measurements. The selection of a representative product is likely to be crucial in minimizing model output uncertainty. In this step, the available monthly rainfall observations in the sites shown in Fig. 1.C were used to investigate the performance of the five mentioned rainfall products using a point-to-pixel approach (see Fig. 2). The approach compares the time series of observed rainfall to the gridded products assuming that grid-cell values are represented by the corresponding rain gauge measurements (e.g., Thiemig et al., 2012; Basheer and Elagib, 2019; McNamara et al., 2021). This approach is useful for data-scarce regions.

A statistical validation using six performance metrics (three linear fit metrics: R^2, Intercept, and Slope; and three error metrics: MAE, MBE, and RMSE) were calculated using the equations in Table B1 in the appendices. Additionally, to draw general conclusions, the products were ranked based on the calculation of the Unified Metric (UM: Eq.(1)) and the Overall Unified Metric (OUM: Eq.(2)) (Elagib and Mansell, 2000). The UM indicates the ranking of each product at each station, while the OUM is the sum of the UMs for each product at all stations. Firstly, each product is ranked per station for each metric. Secondly, the performance rankings are summed up per product with a low value indicating good performance and a high value

indicating poor performance relative to each other.

$$\text{UM}_{p,i} = \sum_{m=1}^{K} R_{p,m,i} \qquad (1)$$

$$\text{OUM}_p = \sum_{i=1}^{L} \text{UM}_{p,i} \qquad (2)$$

where $R_{p,m,i}$ is the rank of the product p based on the performance metric m at station i, and OUM_p is the Overall Unified Metric of the rainfall product p. While K and L are the number of considered performance metrics and rainfall stations, respectively. UM ranges from $K = 6$ to $K \times$ No. of rainfall products = 30. While the minimum OUM value is 60 ($K \times L$) and the maximum value is 300 ($K \times L \times$ No. of rainfall products).

Based on this analysis, the first three ranked rainfall products will be selected to analyse the influence of the uncertainty in the rainfall product on retrieving reservoir storage strategies. This is due to the fact that precipitation changes can dominate reservoir volume changes (Hou et al., 2022). It is important to note that the rank of the rainfall products might differ if the analysis was based on the daily timestep that is used in our hydrological model simulations.

Flowchart representing research methodology. Steps are shown in chronological order, the execution is carried out from left to right. For rainfall product validation five products are selected which are ARC 2.0, CHIRPS 2.0, ERA5, GPCC, and PERSIANN-CDR. For storage validation, observed storage will be computed by converting the water surface area observed by Landsat to water volume using an Elevation-Area-Storage relationship of the GERD reservoir.

3.2 Hydrological modelling

From 57 previous studies in the literature between 2007 and 2022, 26 hydrological models have been implemented in the Upper Blue Nile basin (see Table C1). The Soil and Water Assessment Tool (SWAT) and the Hydrologiska Byråns Vattenbalansavdelning (HBV) are the most commonly used models, utilized in 23 and 13 studies, respectively. Both models performed well in simulating the runoff at the outlet of the UBN basin (Betrie et al., 2009; Teklesadik et al., 2017). Bizuneh et al. (2021) compared their performance in the UBN basin and concluded a better performance of SWAT in the three studied watersheds. However, the HBV conceptual model will be used due to its lower demand for input data compared to the SWAT model and the limited number of parameters. HBV is a widely used conceptual hydrological model developed by Bergström et al. (1995), but with several different implementations being used (Jansen et al., 2021). In this study, a lumped representation of the original HBV model will be used which is HBV-light (Seibert and Vis, 2012). The lumped model simulates daily discharges using daily average precipitation, temperature, and potential evapotranspiration over the basin as input (Seibert, 1996). In addition to its flexibility and computational efficiency, HBV has been successfully applied in the UBN basin (e.g., Uhlenbrook et al., 2010).

HBV-light will be calibrated for a period of 14 years (70%) from 2006 to 2019 and afterward validated for 6 years (30%) from 1991 to 1996 (see Fig. 2). The selection of the years was due to a considerable gap in the measurements from 1997 to 2000. As a Although the climate might show non-stationarity, we still consider the optimized parameters robust enough to draw conclusions for two reasons. First, the calibration and validation periods extend over a long duration. Second, as noted by Melsen et al. (2018), parameter sets are not typically the dominating source of uncertainty when considering long-term

changes simulated with hydrological models such as HBV, which was also employed in this study. Additionally, as a warm-up
year, one year before each period (i.e., 2005 and 1990) will be used to avoid the effect of the initial values. Furthermore, to account for parameter uncertainty, the generalized likelihood uncertainty estimation (GLUE) method will be adopted (Beven and Freer, 2001). The GLUE method was employed as it is simple and commonly used in hydrology. In hydrological models, Nash-Sutcliffe Efficiency (NSE) proposed by Nash and Sutcliffe (1970) is widely used as the objective function (Setegn et al., 2010). Therefore, different parameter sets will be selected based on NSE exceeding a threshold value of 0.75 which is classified as very good performance according to Moriasi et al. (2007). Moreover, to evaluate the performance of the hydrological model, Moriasi et al. (2007) classification will be used.

3.3 Retrieving GERD filling strategies

3.3.1 Reservoir area and storage from satellite observations

As observations of GERD filling are currently not available, the quantification of the reservoir storage using ~~Landsat images~~ satellite imagery will allow to validate the ability of the hydrological modelling in retrieving the filling strategies (see Fig. 2). This can also provide insights into the sensitivity of storage estimation to the selection of rainfall products.

~~To~~ On the one hand, to derive reservoir storage from Landsat images, Landsat Collection 2 Level 2 was obtained. Landsat was selected due to its high spatial resolution (30 m) as well as its reasonable revisit time (16 days). Additionally, to increase the temporal resolution up to 8 days, images from multiple satellites (i.e., Landsat 7/8/9) were used ~~following the steps~~ (steps are shown in Fig. ~~6.A~~A1.A). Different indices allow the estimation of the Water Surface Area (WSA), in this study the Normalized Difference Water Index (NDWI) based on Eq.(3) was selected. Hence, three bands are needed from Landsat which is: the Green (GRN), Near Infrared (NIR), and Quality Assessment bands.

$$\text{NDWI} = \frac{\text{GRN} - \text{NIR}}{\text{GRN} + \text{NIR}} \tag{3}$$

Firstly, the images over the GERD reservoir were obtained from the three satellites. However, the images of Landsat 7 contain missing data stripes caused by the Scan Line Corrector (SLC) failure (Scaramuzza and Barsi, 2005). The missing data were filled using the Fill nodata tool of the Quantum GIS (QGIS, 2009). Then, as the GERD reservoir is covered by two tiles, the tiles were mosaicked and clipped over the reservoir extent. Based on the NDWI maps, the algorithm of Vu et al. (2022) use the data to estimate the WSA. This algorithm is an improvement of the algorithm introduced by Gao et al. (2012) and modified by Zhang et al. (2014) to allow for the use of Landsat Collection 1 data. The algorithm improves the estimation of the images with clouds using the cloudless images by removing the clouds, clouds' shadows, and no-data pixels. The improvement is two-fold; the selection of cloudless images and the identification of additional water zones. From 2022 onwards, Collection 2 structures only will be used for new Landsat data. Therefore, the algorithm was adapted for the processing of all new Landsat acquisitions. ~~Afterward,~~

On the other hand, the computation of WSA from Sentinel-2 imagery was performed using the PyGEE-SWToolbox (Owusu et al., 2022). This open-source toolbox was observed to be a user-friendly, adaptable, and dependable method for quantifying WSA.

The software provides a graphical user interface for obtaining WSA time-series data from diverse satellite imagery datasets, including Sentinel-2. Consequently, by utilizing NDWI calculations and a predetermined threshold value of 0, we can accurately detect and classify pixels as water bodies (i.e., NDWI > 0). Detailed steps can be seen in Fig. A1.B.

Finally, to convert the estimated WSA to ~~convert the area to~~ storage, an Area-Storage relationship is required. For this, the Elevation-Area-Storage relationship of the reservoir was produced from the Shuttle Radar Topography Mission (SRTM) digital elevation model (Farr et al., 2007) using the python code written by Vu et al. (2022). Finally, dates with suspicious values were checked manually and erroneous values were removed. The unrealistic values were found to occur mainly when the images are completely covered with clouds.

3.3.2 Reservoir storage from hydrological modelling

Since July 2020, GERD started filling and the assumption that the observed flow to Sudan (at Eldiem) is equal to the incoming flow to GERD is no longer valid. Accordingly, to predict the unknown inflow of GERD after 2020, the calibrated hydrological models will be utilized. HBV-light will be run for the period 2020-2022 using the optimized parameter sets (see Fig. 2). Thereafter, the observed and simulated discharge together will be used to estimate the water storage volume at the reservoir using the following equations:

$$Q_{out} = Q_{in} - \frac{dS}{dt} - E \tag{4}$$

$$\frac{dS}{dt} = \theta \cdot Q_{in} - E \tag{5}$$

$$\theta = 1 - \frac{Q_{out}}{Q_{in}} \tag{6}$$

where Q_{in} and Q_{out} are the daily inflow to and outflow from the dam, respectively, $\frac{dS}{dt}$ is the daily change in reservoir storage volume, E is the evaporated volume of water from the reservoir, and θ is the fraction of inflow volume retained by the reservoir which ranges between 0 and 1. The last two equations (i.e., Eq. 5 and 6) ~~is~~ are used to estimate the daily storage which when accumulated gives the total stored water in the reservoir. The daily inflow Q_{in} will be estimated from the hydrological modelling, while the daily outflow Q_{out} will be based on discharge measurements at Eldiem station. To calculate E for each timestep, the estimated daily average evaporation per month obtained from Khairy et al. (2019) is multiplied by the reservoir surface area at that timestep. The area is estimated from the storage using the acquired Area-Storage relationship.

4 Results

4.1 Performance of the rainfall products

The performance of the five selected rainfall products based on ten rain gauge stations was estimated and illustrated in a boxplot for each performance metric as shown in Fig. 3. From the boxplots, it can be revealed that CHIRPS has the most accurate estimates as it is the closest product to the optimal values (horizontal dashed red lines) in most of the evaluated metrics. CHIRPS has a high mean R^2 of 0.89, a mean slope of 0.96, and a relatively low mean intercept of 10 mm. CHIRPS
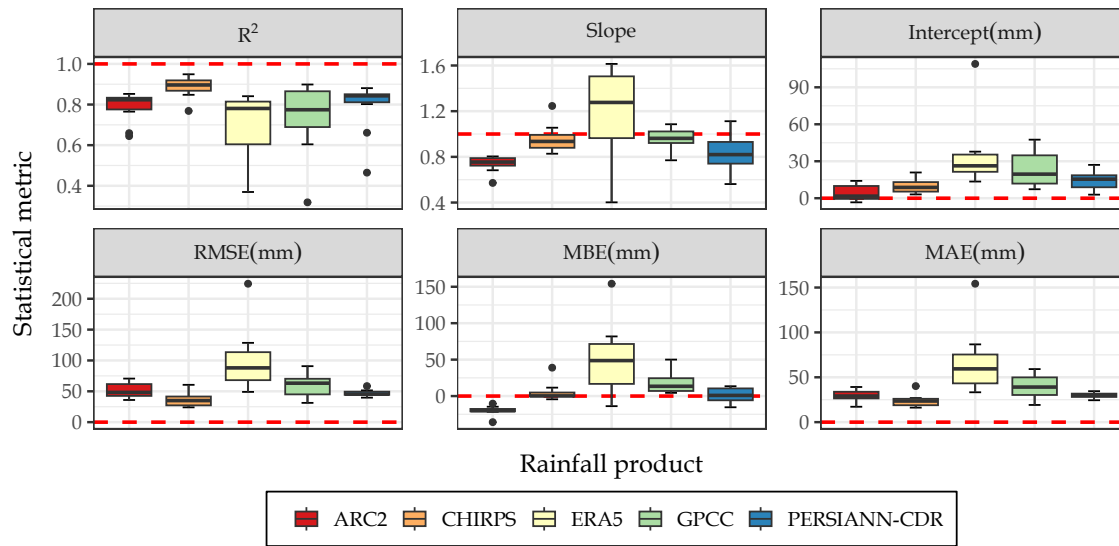
Figure 3. Performance of the five selected rainfall products. Each box reflects the distribution over ten stations (see Fig. 1.C and Table 3). The horizontal dashed red lines indicate the optimal values. Detailed calculations are illustrated in Fig. B2 and the equations are listed in Table B1.

has the lowest values for the error metrics which were found to be 36.66 mm for the RMSE, 5.31 mm for the MBE, and 23.7 mm for the MAE. Moreover, GPCC and ARC2 exhibited the best mean slope and intercept, respectively. On the other hand, ERA5 shows the highest overestimation (mean slope of 1.28) and the highest error values. Additionally, it is also clear that ARC2 consistently underestimates the rainfall in all locations (negative MBE values). The performance of the rainfall products also varied spatially at the different stations. The spatial distribution of the calculated metrics (see Fig. B1) demonstrated consistent performance of CHIRPS over the spatial extent and high spatial variation of ERA5. For further demonstration of the metrics' values and spatial distribution see Fig. B1 and B2 in the appendices.

Furthermore, the UM in Table 3 emphasizes the accuracy of CHIRPS (i.e., has the lowest values) almost in all stations, unlike ERA5 which was the least accurate product. Additionally, the calculation of the OUM concluded that the overall best-performing product is CHIRPS followed by PERSIANN-CDR, ARC2, GPCC, and ERA5. Therefore, CHIRPS is expected to perform better in simulating the discharge in the river, estimating the water stored behind the dam, and retrieving the filling strategies. However, the first three rainfall products were selected as forcing data for the hydrological modelling to account for the input uncertainty in retrieving the filling strategies.

Table 3. Ranking of the Satellite and Reanalysis Rainfall Products for each station based on ten selected rain gauge stations. UM is the Unified Metric and OUM is the Overall Unified Metric.

No.	Station	Longitude	Latitude	Unified Metric (UM)				
				CHIRPS	GPCC	ERA5	PERSIANN-CDR	ARC 2.0
1	Mekelle	39.5	13.5	7	16	25	25	17
2	Gonder Airport	37.4	12.6	11	8	27	20	24
3	B-dar Synoptic	37.4	11.6	8	25	18	15	24
4	Combolcha	39.7	11.1	9	19	28	16	18
5	Debre Markos	37.7	10.3	7	23	29	15	16
6	Mehal Meda	39.7	10.2	14	19	30	13	14
7	Shola Gebeya	39.4	9.2	8	15	29	16	22
8	Addis Ababa Bole	38.8	9.0	9	23	29	15	14
9	Jimma	36.8	7.7	8	27	20	13	22
10	Robe	40.0	7.1	11	24	26	15	14
	OUM			92	199	261	163	185
	Overall rank			1	4	5	2	3

4.2 Hydrological modelling using HBV-light

4.2.1 Model calibration and validation

The HBV-light model was forced by ERA5 temperature and potential evapotranspiration estimates while CHIRPS, PERSIANN-
275 CDR, and ARC2 rainfall data were provided to account for input uncertainty. Initially, the parameter ranges tested in the Gilgel Abay catchment in the UBN basin by Uhlenbrook et al. (2010) were adopted. However, the model performance was limited by the initial range of the maximum soil moisture (FC) which was 200 to 600 mm. Therefore, a new range was selected from 200 to 1000 mm which was sufficient to tackle this issue. HBV-light routines, parameters symbol and description, and parameter ranges are listed in Table D1.

280 To account for parameter uncertainty, 10,000 random parameter sets were generated by the Monte Carlo runs of HBV-light based on the parameter ranges and assuming a uniform distribution. After running the model during the calibration period, 1756 model simulations achieved NSE > 0.75 when using CHIRPS while only 269 and 244 simulations in the case of PERSIANN-CDR and ARC2, respectively. In further analysis, we rely on the parameter set that resulted in the best model performance (best simulation). However, we acknowledge that other parameter sets with a high performance may be equally valid. Fig. D1
285 shows the range of each parameter for the very good simulations in comparison to the applied parameter range (with narrower ranges indicating sensitivity). As a result, the simulated discharge was found to be sensitive to FC, BETA, K1, and K2 when using the three rainfall products.
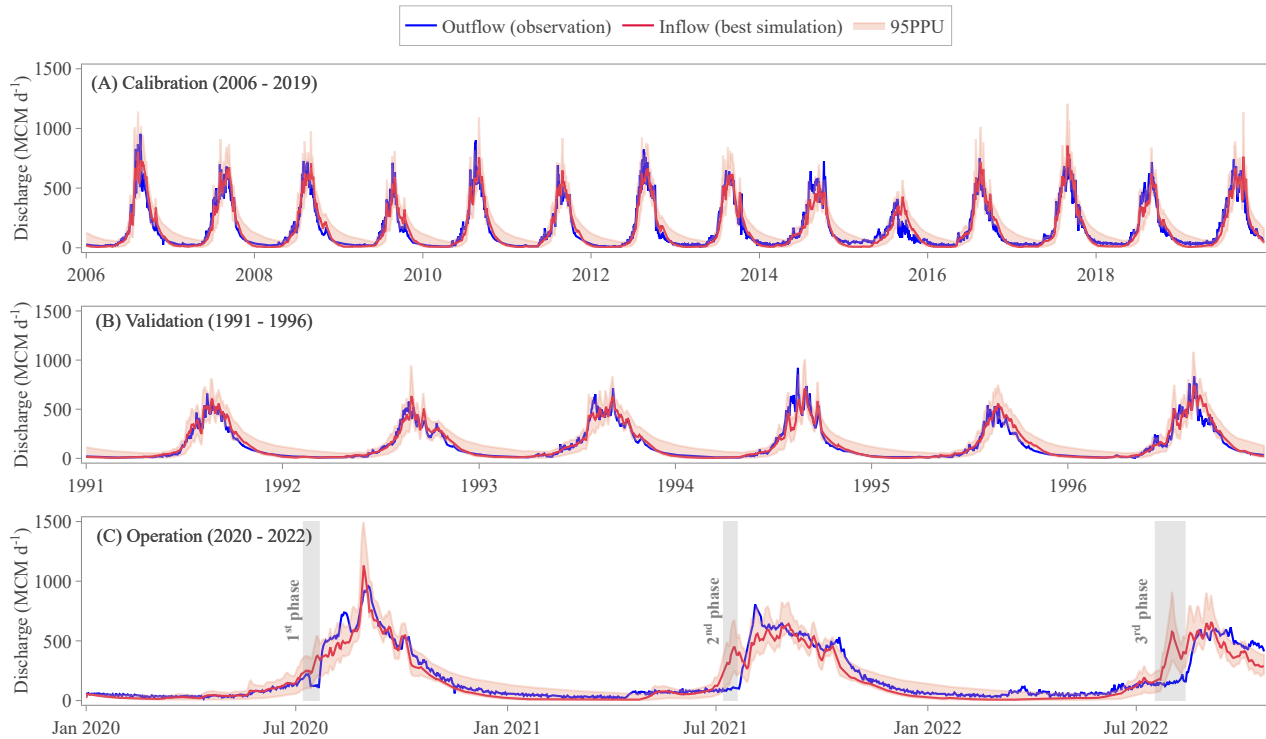
Figure 4. Temporal dynamics of daily simulated and observed discharge at the Upper Blue Nile basin outlet. Simulations forced with CHIRPS rainfall are shown during (A) the calibration period, (B) the validation period, and (C) the operation period. The best simulation was based on the parameter sets that achieved the highest NSE value during calibration. The 95% Prediction Uncertainty (95PPU) represents the 95th percentile of all very good simulations (NSE > 0.75) that was obtained by random generation of the parameter sets. The vertical grey shaded areas in the operation period roughly indicate the days in which the dam is filling. MCM stands for Million Cubic Meters.

Fig. 4.A and 4.B ~~shows~~ show the daily time series of simulated, using CHIRPS, and observed discharges over the calibration and validation periods, respectively (see Fig. D2.A and D2.B for both PERSIANN-CDR and ARC2). The parameter set that obtained the highest performance for each rainfall product is indicated in Table D1. Looking at the best simulation, it can be revealed that HBV-light was able to capture the intra-annual seasonality in the three rainfall cases and in both periods but hardly capture the daily variation in discharge, especially in high flows which were also remarked by Uhlenbrook et al. (2010). Accordingly, the models resulted in very good NSE values (as shown in Table 4) and dropped after removing the seasonality to satisfactory (0.51) and unsatisfactory (-0.32 and -0.25) for CHIRPS, PERSIANN-CDR and ARC2 during calibration years, respectively.

Table 4 also allows comparing the behaviour of the model when utilizing the different rainfall products. Firstly, the performance of the best simulations was represented by the first three metrics (NSE, PBIAS, and RSR) with their equations in Table B1. For all rainfall products, NSE and RSR was found to be very good in the calibration and validation runs except for

Table 4. The statistic summary of HBV-light simulations.

Product	NSE	RSR	PBIAS (%)	P-Factor (%)	R-Factor
Calibration (2006 - 2019)					
CHIRPS	0.90 (Very good)	0.32 (Very good)	−3.61 (Very good)	77.84	0.73
PERCIANN-CDR	0.80 (Very good)	0.45 (Very good)	−5.06 (Very good)	54.98	0.40
ARC2	0.79 (Very good)	0.46 (Very good)	1.53 (Very good)	37.65	0.35
Validation (1991 - 1996)					
CHIRPS	0.91 (Very good)	0.3 (Very good)	4.64 (Very good)	85.13	0.78
PERSIANN-CDR	0.74 (Good)	0.52 (Good)	32.20 (Unsatisfactory)	49.18	0.50
ARC2	0.81 (Very good)	0.44 (Very good)	22.88 (Satisfactory)	58.02	0.39

The classification according to Moriasi et al. (2007) is as follows:

Very good:	$0.75 < NSE \leq 1.0$;	$0 \leq RSR \leq 0.5$;	$PBIAS < \pm10$
Good:	$0.65 < NSE \leq 0.75$;	$0.5 < RSR \leq 0.6$;	$\pm10 \leq PBIAS < \pm15$
Satisfactory:	$0.50 < NSE \leq 0.65$;	$0.6 < RSR \leq 0.7$;	$\pm15 \leq PBIAS < \pm25$
Unsatisfactory:	$NSE \leq 0.5$;	$RSR > 0.7$;	$PBIAS \geq \pm25$

PERSIANN-CDR during validation which were found to be good, confirming the skilful predictability of the HBV-light model. On the other hand, PBIAS was generally very good but ARC2 and PERSIANN-CDR overestimated the discharge when applied from 1991 to 1996, resulting in satisfactory and unsatisfactory performance, respectively. Secondly, the P-Factor and R-Factor were used to measure the uncertainty. The highest percentage of the observations bracketed by the 95PPU was in the case of CHIRPS (78% and 85%) whereas the lowest percentages were ARC2 during calibration (38%) and PERSIANN-CDR during validation (49%). Additionally, the uncertainty range of CHIRPS was found to be the widest (0.73 and 0.78) and ARC2 the narrowest (0.35 and 0.39). Overall, from the analysis of model performance and uncertainty, it can be concluded that CHIRPS is the best product for simulating the discharge at the outlet of the UBN basin followed by ARC2 and then PERSIANN-CDR.

4.2.2 Predicted inflow to GERD

Since 2020, the inflow to the GERD dam remains unknown to Sudan due to data sharing challenges. Therefore, the main purpose of the HBV-light model in this study was the predictability of the inflow to GERD. The selected parameter sets in the last step were run from 2020 to 2022. Fig. 4.C illustrates the daily inflow (from HBV-light using CHIRPS) and outflow (from Eldiem station) of GERD (see Fig. D2.C for both PERSIANN-CDR and ARC2). It is important to note that PERSIANN-CDR data were available up to June 2022 (before the third filling phase).

Firstly, for the three rainfall products, it is obvious that during filling dates (vertical grey shaded areas) the inflows were greater than the outflow indicating the occurrence of reservoir filling. However, the filling periods were better captured by

CHIRPS followed by PERSIANN-CDR and ARC2 which gave longer filling periods (longer periods where inflow > outflow). Moreover, PERSIANN-CDR demonstrates two filling phases in July and August of the first year.

Secondly, during no-fill dates, water passes through the dam so it is expected that the inflow is equal to the outflow. However, due to model, parameter, and data uncertainties, this is difficult thus outflow is supposed to be within the uncertainty range. Only for CHIRPS runs, the 95PPU range was able to bracket the discharge during no-fill dates. In the case of PERSIANN-CDR, the inflow was significantly lower than the outflow, falsely indicating a release from the reservoir (no emptying of the reservoir occurred during this period). Lastly, ARC2 showed very high inflows from June to October suggesting a large storage amount.

Overall, based on the analysis during the filling period, CHIRPS was efficient in reproducing the hydrological properties of the system after 2020, while PERSIANN-CDR and ARC2 missed important features (i.e., filling dates).

Temporal dynamics of daily simulated and observed discharge at the Upper Blue Nile basin outlet. Simulations forced with CHIRPS rainfall are shown during (A) the calibration period, (B) the validation period, and (C) the operation period. The best simulation was based on the parameter sets that achieved the highest NSE value during calibration. The 95% Prediction Uncertainty (95PPU) represents the 95th percentile of all very good simulations (NSE > 0.75) that was obtained by random generation of the parameter sets. The vertical grey shaded areas in the operation period roughly indicate the days in which the dam is filling. MCM stands for Million Cubic Meters.

The statistic summary of HBV-light simulations. Product NSE RSR PBIAS (%) P-Factor (%) R-Factor CHIRPS 0.90 (Very good) 0.32 (Very good) −3.61 (Very good) 77.84 0.73 PERCIANN-CDR 0.80 (Very good) 0.45 (Very good) −5.06 (Very good) 54.98 0.40 ARC2 0.79 (Very good) 0.46 (Very good) 1.53 (Very good) 37.65 0.35 CHIRPS 0.91 (Very good) 0.3 (Very good) 4.64 (Very good) 85.13 0.78 PERSIANN-CDR 0.74 (Good) 0.52 (Good) 32.20 (Unsatisfactory) 49.18 0.50 ARC2 0.81 (Very good) 0.44 (Very good) 22.88 (Satisfactory) 58.02 0.39

4.3   Reservoir storage estimation

4.3.1   Elevation-Area-Storage (E-A-S) relationship

The Elevation-Area-Storage (E-A-S) relationship is essential in this work for two reasons; (1) converting the WSAs observed by ~~Landsat~~ satellite imageries to storages and (2) estimating the WSA corresponding to daily modelled GERD storage hence calculating the total daily evaporation from the reservoir. Therefore, the E-A-S relationship was derived from the SRTM starting with estimating the area corresponding to each water level (E-A relationship) at 1 m increments. Next, the reservoir volume corresponding to each water level was calculated using the trapezoidal approximation (Gao et al., 2012; Bonnema and Hossain, 2019). The obtained E-A, A-S, and E-S relationships were fitted with a fifth-degree polynomial as illustrated in Fig. 5. Further detailed information about the approximation can be found in the study by Vu et al. (2022).

On behalf of ten Nile Basin countries, the Nile Basin Initiative (NBI) has been commissioned to conduct studies in the region. The Eastern Nile Technical Regional Office ~~(ENTRO)~~ of the NBI produced the E-A-S relationship of GERD reservoir and was published in Wheeler et al. (2016). When comparing the relationships derived from the DEM with the NBI, there is
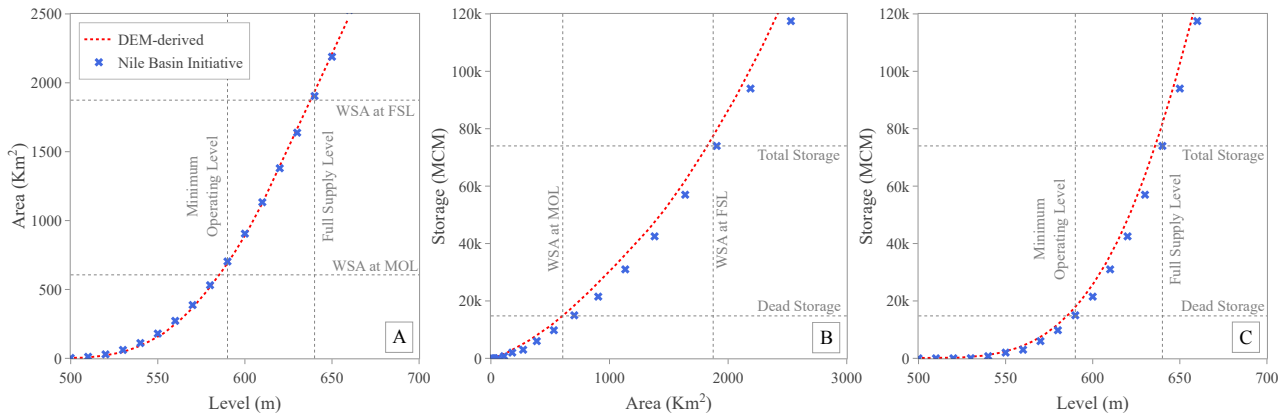
Figure 5. Elevation-Area-Storage (E-A-S) relationship of GERD reservoir obtained from SRTM digital elevation model in comparison with Nile Basin Initiative relationship. (A) Elevation-Area relationship, (B) Area-Storage relationship, and (C) Elevation-Storage relationship. The DEM-derived lines are the results of a fifth-degree polynomial fitting to the data points obtained at 1 m intervals of water levels.

an agreement in the Elevation-Area relationship (see Fig. 5.A). However, the latter underestimates the storage corresponding to the water levels and surface areas (see Fig. 5.B and 5.C).

350   To validate the relationship, the intersected dotted grey lines in Fig. 5 demonstrate the reported characteristics of GERD by the International Panel of Experts (IPOE), (IPoE, 2013). The most important for this study is the A-S relationship and especially for the areas less than 1000 km^2 in size, as GERD recently started to fill. For this range, the DEM-derived relationship intersects with the values reported by IPOE indicating better estimation than the NBI (see Fig. 5.B). Moreover, SRTM was commonly applied to obtain the E-A-S relationship of GERD (Kansara et al., 2021; Chen et al., 2021; Salama et al., 2022). For this reason,

355   the following analysis will be based on the DEM-derived relationship.

~~Elevation-Area-Storage (E-A-S) relationship of GERD reservoir obtained from SRTM digital elevation model in comparison with Nile Basin Initiative relationship. (A) Elevation-Area relationship, (B) Area-Storage relationship, and (C) Elevation-Storage relationship. The DEM-derived lines are the results of a fifth-degree polynomial fitting to the data points obtained at 1 m intervals of water levels.~~

360   **4.3.2   Satellite-derived reservoir area and storage**

4.3.3   ~~Landsat-derived reservoir area and storage~~

The Water Surface Area (WSA) of the GERD reservoir was estimated using Landsat ~~images using the algorithm modified by Vu et al. (2022) following the steps~~ and Sentinel images (steps are shown in Fig. ~~6. A.~~ A1). The results (i.e., reservoir area and storage) were more or less similar. Therefore, we will focus on discussing Landsat results and refer to Sentinel when necessary.

365
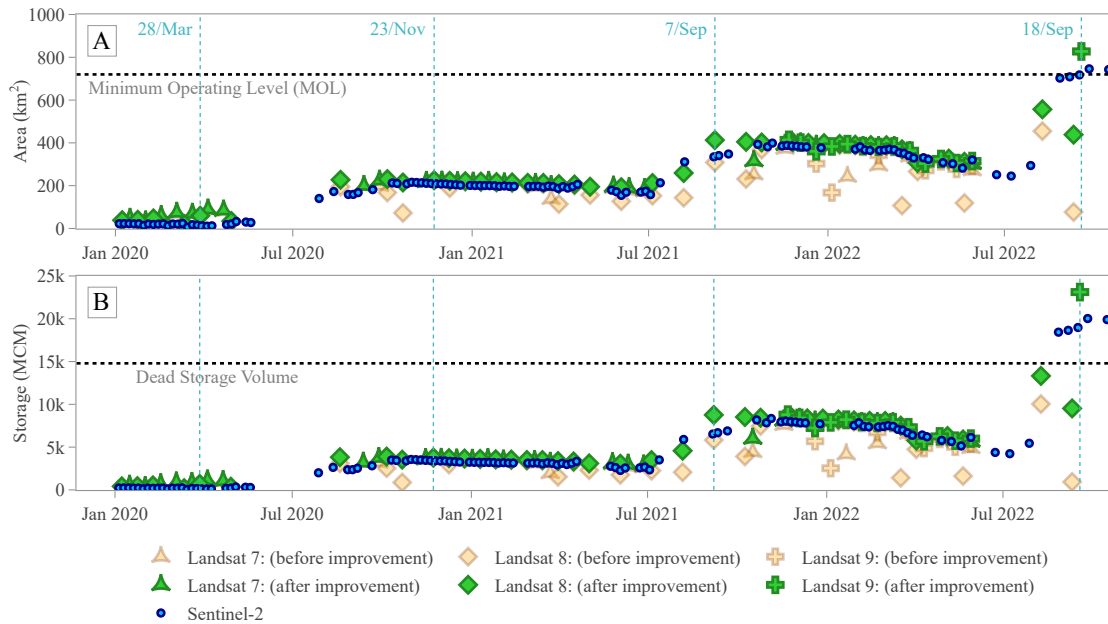
Figure 6. Satellite-derived GERD reservoir information during the filling period 2020 - 2022. Panel (A) shows the derived water surface area, while panel (B) shows the derived storage volume of GERD using Landsat (Green: before improvement; Light orange: after improvement) and Sentinel (blue). The WSA maps of the dates specified in cyan dotted lines are illustrated in Fig. 7. Note that not filling the gaps induced by the Scan Line failure of Landsat 7 produces NDWI = 0 in these gaps when using the algorithm of Vu et al. (2022). Therefore, in case of negative NDWI thresholds (which were the case here), Landsat 7 overestimates the areas and storages.

WSA estimation from Landsat images was based on the algorithm modified by Vu et al. (2022). The algorithm uses the cloudless images (i.e., images with less than 20% clouds) to create an expanded mask and zone mask to correct the disturbed images. 133 Landsat images are available from January 2020 to September 2022, in which 70 images (53%) were classified as cloudless images. Fig. 6.~~B~~ A presents the WSA time series and revealed a significant increase in the WSA after improvement
370 (green points) when compared with the WSA before improvement (light orange points). Additionally, the abrupt changes in the reservoir area are less observed after improvement, showing an important enhancement made by the algorithm. However, the WSA was underestimated in some images due to the high cloud coverage.

As can be seen in Fig. 6.~~B~~ A and 6.~~C~~B, there are missing data during the wetting season (between June and September) of the three years which is due to the high cloud coverage. The first observed storage after the filling was on 19 August 2020 with
375 a volume of 3.82 BCM and continued to decrease to 2.97 BCM on 11 June 2021. After the second phase, the storage jumped to 8.76 BCM (increased by 5.8 BCM) and gradually decrease till 2 March 2022 when a sudden decrease occurred. The water decreased from 8.06 BCM to reach 5.77 BCM by 3 April 2022. This decrease can be attributed to the announced 375 ~~MW~~ Megawatts electricity generation that was started on 20 February 2022 according to Africa News (2022). In the third phase, the storage increased by 17 BCM reaching 23.1 BCM.
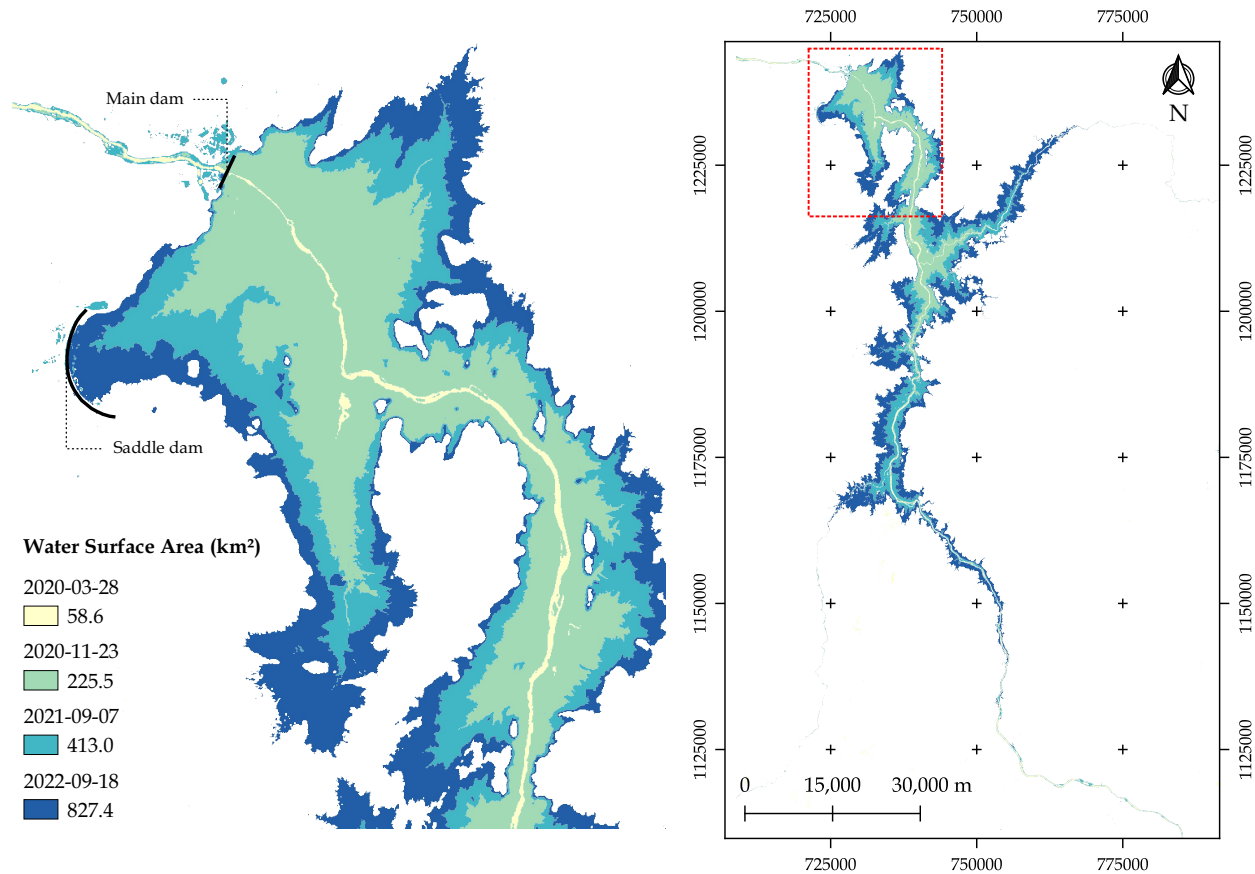
Figure 7. Landsat-derived water surface areas of Grand Ethiopian Renaissance Dam, using the modified algorithm by Vu et al. (2022), showing reservoir development through the first three filling phases from 2020 to 2022. The location of the zoomed area in the left figure is indicated by the rectangle (in red dotted line) in the right figure.

Furthermore, the increase in the WSA due to the filling was well captured by Landsat (and Sentinel) as illustrated in Fig. 7. The area was estimated to be 3.1% of the WSA at the Full Supply Level (FSL; 1874 km^2) before filling started and dramatically increased to reach 12%, 22%, and 44% during the first three filling phases, respectively. This indicates the considerable amount of water stored thus the importance of understanding the filling process for downstream water management.

The utilized algorithm was not always accurate in determining the threshold value of NDWI which in turn influenced the quantification of the area. Therefore, a visual inspection of the threshold was done to improve the estimation. Additionally, in fully cloud images the algorithm still provided additional water which is not true and needs to be checked manually.

Procedure to derive GERD reservoir characteristics from Landsat. (A) The steps followed to derive the Water Surface Area (B) and Storage volume (C) of GERD reservoir from Landsat 7/8/9 Collection 2 Version 2 using the modified algorithm by Vu et al. (2022). The WSA maps of the dates specified in cyan dotted lines are illustrated in Fig. 7. Note that, not filling the

Table 5. The performance of HBV-Light best simulation in retrieving the storage volume of GERD reservoir taking Satellite-derived storage volume as reference.

Satellite Rainfall Product	Landsat		Sentinel	
	NSE	RMSE (BCM)	NSE	RMSE (BCM)
CHIRPS	0.77	1.70	0.86	1.52
PERSIANN-CDR	−4.47	6.26	−2.27	−7.37
ARC2	−307.14	62.40	−206.52	−58.71

390 ~~gaps induced by the Scan Line failure of Landsat 7 produces NDWI = 0 in the gaps when using the algorithm. Therefore, in case of negative NDWI thresholds (which were the case here), Landsat 7 overestimates the areas and storages.~~ It is important to point out that open-access radar altimetry datasets provide valuable information for this validation step. Therefore, for further validation, we also considered radar altimetry data from G-REALM (Birkett et al., 2011) using the Elevation-Storage relationship shown in Fig. 5.C. However, the storage volume derived from G-REALM showed unrealistic behaviour in the

395 recession and only added a few data points during the filling periods, which made us decide to exclude this source from the validation.

~~Landsat-derived water surface areas of Grand Ethiopian Renaissance Dam, using the modified algorithm by Vu et al. (2022), showing reservoir development through the first three filling phases from 2020 to 2022. The location of the zoomed area in the left figure is indicated by the rectangle (in red dotted line) in the right figure.~~

400 ### 4.3.3 Model-derived reservoir storage

The performance of storage volume quantification from HBV-light and the uncertainty due to model parameters and rainfall product selection are highlighted in Fig. 8. The very high modelled storage due to the use of PERSIANN-CDR and ARC2 resulted only in negative NSE values while CHIRPS was the only product that provided reliable reservoir storage as suggested by Fig. 8.A and 8.B. Around 629 runs of CHIRPS (out of 1756) have positive NSE values of which 407 simulations have NSE

405 > 0.5. Additionally, the long filling periods suggested by ARC2 produced the highest RMSE values and range (between 44 and 78 BCM). Comparing the other two products, Fig. 8.C displays a higher and narrower interquartile range in the case of PERSIANN-CDR compared to CHIRPS. The performance of the best simulation of each rainfall product is exhibited in Table 5. Ultimately, the large difference between the rainfall products compared to the deviation caused by parameter uncertainty seen in Fig. 8, concludes that filling strategies are more influenced by input uncertainty than parameter uncertainty.

410 The modelled daily fraction of retained inflow, reservoir storage and area, and evaporated volume of water from HBV-light using CHIRPS are shown in Fig. 9 (for PERSIANN-CDR and ARC2 see Fig. D4). From the first row, it is clear that the duration of filling increases with time suggesting around 13, 20, and 25 days for the first three phases, respectively. Looking at the storage retrieved from HBV-light shown in the second row, one can see the increase in the uncertainty through time which is due to the accumulation of discharge uncertainty with time. It is also important to note that the decrease witnessed by Landsat/Sentinel
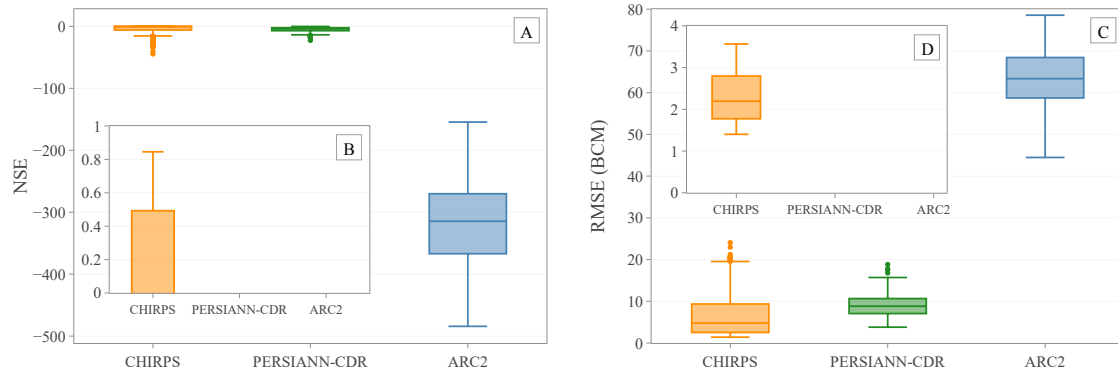
Figure 8. HBV-light performance in retrieving the storage volume of GERD from 2020 to 2022. (A, B) The Nash-Sutcliffe Efficiency (NSE) and (C, D) the Root Mean Square Error (RMSE) between the satellite observed storage (using Landsat) and the simulated GERD reservoir storage (using HBV-light). Panel (B) is a zoom-in of (A) showing NSE between 0 and 1 while (D) presents the RMSE of all simulations with NSE > 0. The x-axis is the three tested rainfall products for input uncertainty while the whiskers demonstrate HBV-light parameter uncertainty. The performance against Sentinel is shown in Fig. D3.

in 2022, due to the hydropower generation, was not captured by the model which does not include release/operation terms thus resulting in additional errors (i.e., lower RMSE can be achieved if this release was included). Additionally, the best simulation suggests that the dead storage (14.79 BCM) was exceeded after the filling of 2022. The last two rows of Fig. 9 exhibit the daily area of the reservoir and the evaporation from the reservoir surface. Obviously, the evaporation increases with time due to the surface area increase which causes a slight gradual decrease in the storage.

~~The performance of HBV-Light best simulation in retrieving the storage volume of GERD reservoir taking Landsat-derived storage volume as reference. Satellite Rainfall Product NSE RMSE (BCM) CHIRPS 0.77 1.70 PERSIANN-CDR −4.47 6.26 ARC2 −307.14 62.40~~

~~HBV-light performance in retrieving the storage volume of GERD from 2020 to 2022. (A, B) The Nash-Sutcliffe Efficiency (NSE) and (C, D) the Root Mean Square Error (RMSE) between the satellite observed storage (using Landsat) and the simulated GERD reservoir storage (using HBV-light). Panel (B) is a zoom-in of (A) showing NSE between 0 and 1 while (D) presents the RMSE of all simulations with NSE > 0. The x-axis is the three tested rainfall products for input uncertainty while the whiskers demonstrate HBV-light parameter uncertainty.~~

4.4 Filling strategies and downstream discharge

4.4.1 GERD filling strategies

Fig. 10 exemplifies GERD filling strategies retrieved from CHIRPS since the other two rainfall products provided insufficient results. The figure shows a time series of monthly θ values and the percentage of reservoir filling. The θ values of May and June 2020 are more than zero due to the low flows in these months and the uncertainty in the model (see also θ in Fig. 9).
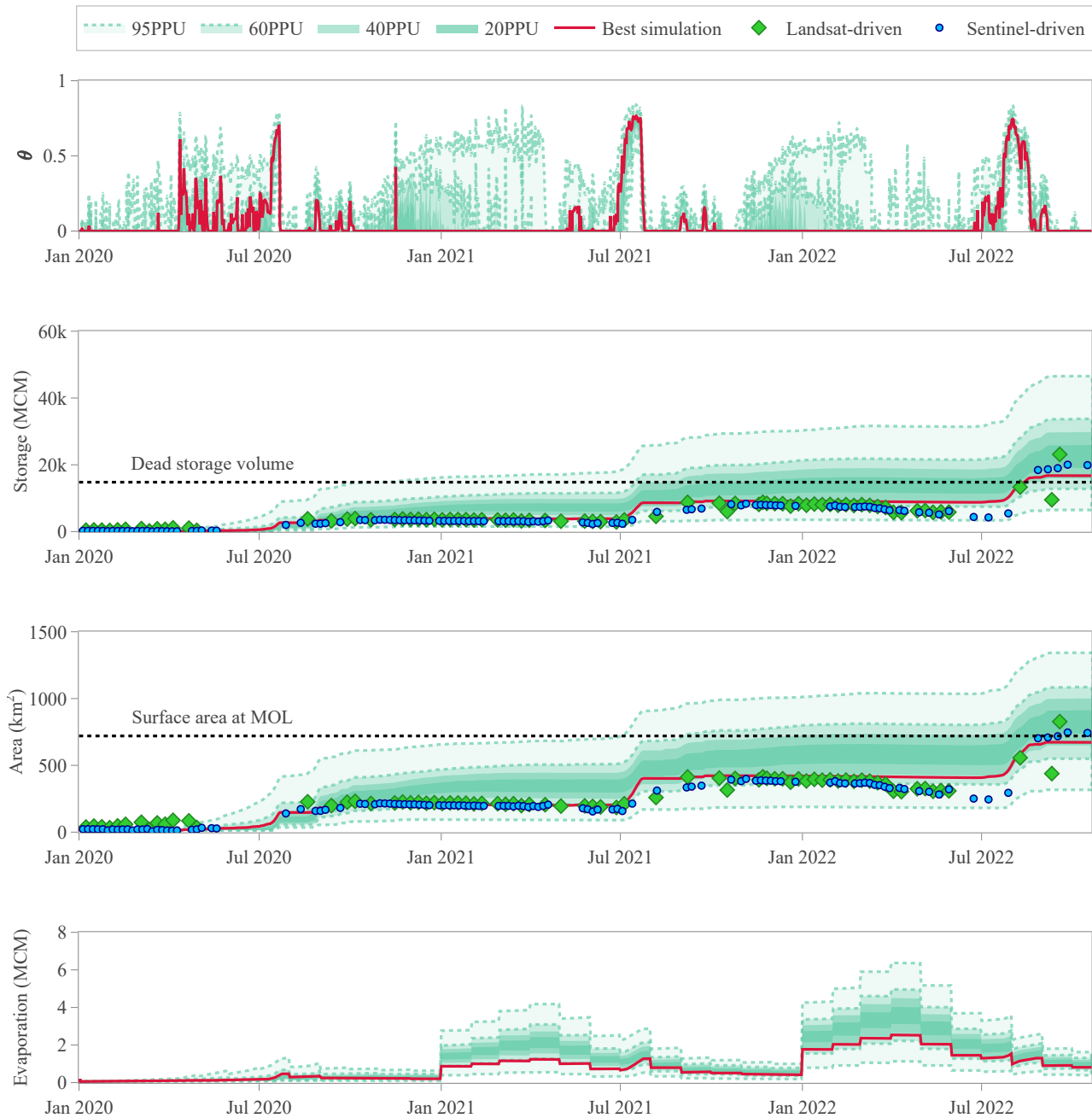
Figure 9. Temporal dynamics of inferred GERD filling strategies. The first three years of the Grand Ethiopian Renaissance Dam (GERD) filling strategies were retrieved from HBV-light lumped model using CHIRPS. The first row is the daily fraction of inflow volume retained by GERD. The second and third rows are the daily water storage and the corresponding water surface area of the GERD reservoir, respectively. The last row is the daily evaporated volume of water from the reservoir surface. The prediction uncertainty ranges indicate the 95, 60, 40, and 20 percentiles of the results obtained from the simulations classified as very good which were 1756.
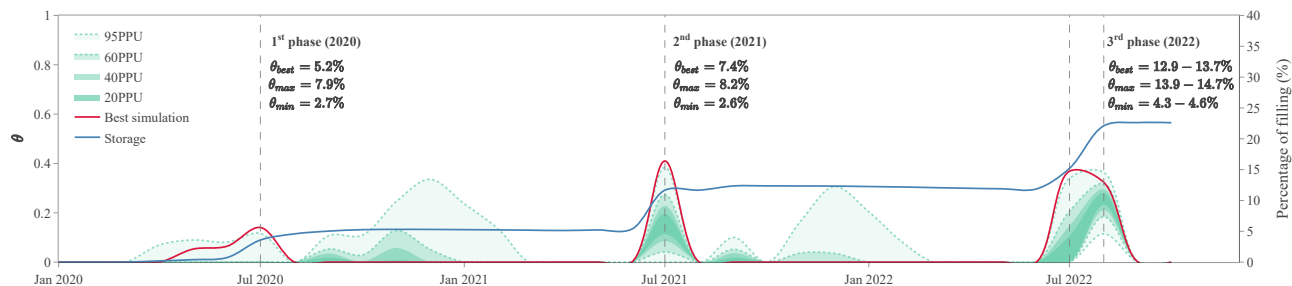
Figure 10. GERD filling strategies. θ expresses the fraction of the monthly inflow volume retained by the dam. The three values θ_{max}, θ_{min}, and θ_{best} (i.e., maximum, minimum, and best simulation) are the fraction of the annual inflow volume retained. The range in the values of the year 2022 is due to the unknown volume of November and December. The percentage of filling was calculated based on the best simulation and the GERD full storage capacity of 74 BCM. The uncertainty ranges (95 - 20 PPU) are based on the 407 runs with NSE > 0.5.

According to the red line, the filling of the three years started in July but continued till August in the year 2022. The percentage of the retained flow were found to be 14% in July 2020, 41% in July 2021, and 37% and 32% in July and August 2022, respectively. Besides the best simulation, the parameter sets that reached NSE > 0.8 in retrieving the storage (Fig. 8.B) were able to well capture the filling dates. For other parameter sets, they suggest filling during other months (see green areas in Fig. 10). Moreover, in terms of annual volume, 5.2% and 7.4% of the total annual flow were retained in the first two years and between 12.9 - 13.7% in the following year. Retaining more than 10% of yearly flow in weeks is significant in terms of volume, and impacts downstream management. Due to parameter uncertainty, the minimum and maximum retained frictions based on the 407 parameter sets (i.e., with NSE > 0.5) were written in Fig. 10. As far as the percentage of filling is concerned, HBV-light suggests that the filling reached 5.3% of the full storage capacity in the first year and considerably rose to reach 12.3% and 22.6% in 2021 and 2022, respectively.

4.4.2 Effect of GERD filling on the inflow to Sudan

With the filling strategies retrieved from the hydrological modelling, the observed daily discharge flowing to Sudan can be better interpreted. Fig. 11 shows further analysis of the filling phases of GERD. At first, the discharge in the last three years showed unnatural behaviour when compared with the discharge before 2020 (see blue lines in Fig. 4). To investigate the role of GERD on the observed hydrograph, the difference in the inflow to Sudan after the construction of GERD and in case of no GERD was estimated and demonstrated in Fig. 11.A. The filling of GERD altered the hydrograph of the flowing water to Sudan. The monthly discharge to Sudan after 2020 showed a notable increase throughout the year except in the filling months (i.e., July 2020-2022, and August 2022) where the flow substantially decreased. This increase is partly due to heavy rains in 2020 and partly due to the operation of GERD. In terms of volume, the filling of the first two years did not influence the annual volume flowing to Sudan unlike in 2022. However, the decrease in the filling months influences the management of
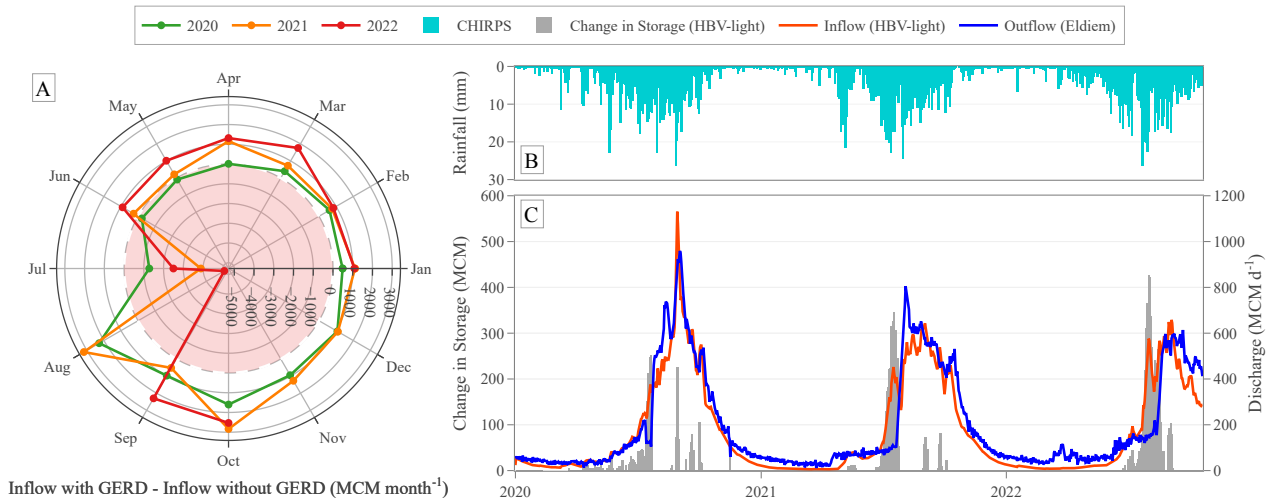
Figure 11. The impact of GERD filling on flow to Sudan: (A) the monthly impact of GERD on the incoming flow to Sudan: calculated as the difference between the inflow to Sudan with GERD and the inflow in case of no GERD, (B) The rainfall time series retrieved from CHIRPS, and (C) The daily change in storage of GERD reservoir and its implication on downstream daily discharge measured in Sudan. The change in storage is defined as the difference between two consecutive daily storage volumes.

the ~~Roseries~~ Roseires and Sennar dams to store and divert enough water to the Gezira scheme and to meet other downstream demands.

455   Moreover, there are abnormal high flows caused by the highly intense rainfalls such as in August to October 2020 and May 2021 (see Fig. 11.B and 11.C). However, the considerable rise in the flowing water observed in March and April 2022 occurred as a result of the hydropower generation that started in February 2022 (see GERD release in Fig. 6.~~C~~B). Paying attention to the storage change in Fig. 11.C, the relatively low outflows during filling days are followed by a sudden and steady increase in the outflow causing a linear hydrograph. Additionally, Table 6 illustrates the start and end dates of these changes and the

460   magnitude of daily increase. The results revealed that the outflow increments are significant and decreased from around 77 MCM per day in 2020 to 40 MCM per day in 2022.

~~The impact of GERD filling on flow to Sudan: (A) the monthly impact of GERD on the incoming flow to Sudan: calculated as the difference between the inflow to Sudan with GERD and the inflow in case of no GERD, (B) The rainfall time series retrieved from CHIRPS, and (C) The daily change in storage of GERD reservoir and its implication on downstream daily discharge~~

465   ~~measured in Sudan. The change in storage is defined as the difference between two consecutive daily storage volumes.~~

5   Discussion

The current work might be of interest to studies on ungauged basins. One of the remaining challenges in ungauged basins is the need for more easily accessible data (Hrachowitz et al., 2013). The deficiencies of meteorological stations were replaced

Table 6. Dates and volumes of the sudden changes in the downstream discharge after the completion of each GERD filling phase.

Filling phase	Start		End		Average daily
	Date	Discharge (MCM)	Date	Discharge (MCM)	increase (MCM)
1st phase (2020)	21 July	104.82	26 July	489.18	76.87
2nd phase (2021)	21 July	100.40	4 August	805.36	50.35
3rd phase (2022)	13 August	159.91	22 August	522.90	40.33

by a selection of a representative rainfall product based on the performance rankings adopted from Elagib and Mansell (2000)
470 . The ranking was found to give a first impression of the hydrological model performance to quantify the discharge as well as retrieve the filling strategies. Additionally, the calibration and validation of the hydrological model based on pre-operation years allowed for successful prediction of the inflow during post-operation years which was proved by the utilization of Earth observation. Moreover, coupling the retrieved change in storage with the measured outflow of the dam allowed a further understanding of the downstream impact of the filling process (Vu et al., 2022).

475 The retrieved GERD storage results are dependent on HBV-light performance, which is subject to various sources of uncertainty, including observed discharge, meteorological products, and model parameters and structures. Rainfall products were found to be a main source of uncertainty with a high influence on the HBV-light results, followed by model parameters (as discussed in section 4.3.3). Although uncertainty arising from potential evapotranspiration products was not investigated, previous studies (Nonki et al., 2021; Wang et al., 2022) have concluded that it is unlikely to have a significant effect on HBV-light
480 performance and simulations. The selected model structure, HBV-light, is also likely to impact the results. However, the limited data availability limits opportunities to explore more complex structures.

The unforeseeable construction progress of GERD complicates the predictability of the years required to completely fill the dam. Nevertheless, some trends can be expected based on the results of the current research. Retaining 20% of the average annual flow at the dam location (around 50 BCM) is equivalent to storing 10 BCM a year. Therefore, given the current stored
485 water (Landsat: 23.1 BCMof water; Sentinel: 20 BCM) and the retrieved friction of retained annual flow, GERD will need at least 5 more years (8 years in total) to complete filling under average flow conditions. However, a minimum of additional 3 years will be needed under wet conditions (6 years in total). During these years, it is important to reduce the risk of water diversions in Sudan by the adaptive management of reservoir operations (Wheeler et al., 2016). For the adaptive management of the Sudanese reservoirs, the current model can be extended to include the Lower Blue Nile basin and used to test operation
490 scenarios for Roseires and Sennar dams. Other cooperative options for mitigating the effect of GERD on downstream flow, are coupling the operation of GERD with solar and wind power (Sterl et al., 2021) and applying water sharing policy during drought conditions (Wheeler et al., 2016; Yang and Block, 2021).

The current work might be of interest to studies on ungauged basins. One of the remaining challenges in ungauged basins is the need for more easily accessible data (Hrachowitz et al., 2013). The deficiencies of meteorological stations were replaced

495 ~~by a selection of a representative rainfall product based on the performance rankings adopted from Elagib and Mansell (2000). The ranking was found to give a first impression of~~ operation of downstream dams (i.e., Roseires and Sennar) depends on the management of the GERD dam. Thus, the outcomes of this approach provide information that can support downstream dam operations. The findings of this study indicate that lower discharge levels were observed during July and August, while higher discharge rates were observed during other periods. It suggests that additional water should be stored during GERD-non-filling

500 periods to achieve the targeted monthly elevation of the ~~hydrological model performance to quantify the discharge as well as retrieve the filling strategies. Additionally, the calibration and validation of~~ Roseires reservoir. A substantial increase in discharge is noted during March-June and October-December. We propose that effective management of Roseires and Sennar should involve storing additional water during these months to compensate for shortages during filling periods. Additionally, for adaptive management of Sudanese reservoirs, the existing model should be expanded to include the Lower Blue Nile basin

505 and used to evaluate alternative operation scenarios for Roseires and Sennar dams. This approach will enable a comprehensive understanding of the implications of various management decisions and facilitate the development of optimized reservoir management plans.

It is worth mentioning the main advantages of the proposed approach. Besides the approach discussed here, the downstream analysis can also be based on satellite images using, for example, the approach as proposed by Vu et al. (2022). However,

510 relying on the latter approach for real-time operation presents certain challenges. Firstly, given the current availability of free satellite data (such as Sentinel and Landsat), it is not feasible to achieve daily time steps, unlike in the ~~hydrological model based on pre-operation years allowed for successful prediction of the inflow during post-operation years which was proved by the utilization of Earth observation. Moreover, coupling the retrieved change in storage with the measured outflow of the dam allowed a further understanding of the downstream impact of the filling process (Vu et al., 2022).~~ case of relying on

515 hydrological modelling. Secondly, waiting for a few days to acquire satellite data can be problematic, particularly during flood events, such as those experienced in Sudan in 2020. Moreover, hydrological modelling allows to investigate future scenarios under different management strategies. It is important to note that our proposed approach relies on outflow observations, which may not always be available/accessible. As such, both approaches have their respective advantages and limitations."

Worldwide, there are ~~268~~ 286 transboundary rivers covering 50% of the global land surface ~~(Draper, 2007)~~(De Stefano et al., 2017; Drape

520 . Going a step forward, the novel methodology described in this paper can be adopted on a global scale, especially for data-scarce regions and transboundary rivers. However, discharge observations and the selection of representative precipitation products are crucial. The global implementation will raise the practical question: "How ~~does~~ is the precipitation product ~~will be~~ selected? the best product globally? or a precipitation product per climate zone? or per reservoir?". Another important aspect is the reliability of the current methodology. With the current error of around 1.7 BCM (or 1.52 BCM as suggested by Sentinel),

525 the reliability increases with the increase of reservoir storage capacity. Thus, it is preferable to adopt the methodology for dams with large storage capacities. Nonetheless, further improvements in the methodology are needed to increase model efficiency and thus cover small reservoirs. The improvements are (but are not limited to):

- Including more observations (e.g. soil moisture, groundwater measurements): with this additional data, independent validation can be done which can help in eliminating many of the high-performed discharges due to their low performance on the other component (i.e., the additional used observation).

- Validating both the precipitation and evapotranspiration when using remote sensing data: to minimize the uncertainty.

- Utilizing actual evaporation measurements/validated product: to improve capturing the day-to-day variation in reservoir storage.

- Including seepage term in the quantification of reservoir storage from reservoir water balance: new water levels are reached during filling, requiring more water initially than the obtained from the Elevation-Storage relationship to reach a certain level.

- Using a more spatially representative hydrological model: daily variation in discharge can be better captured, but it requires extra information which may not be available.

6 Conclusions

The work done in this paper demonstrated ~~the ability to~~ a novel approach that utilize hydrological modelling with limited ground measurements to infer GERD filling strategies ~~. The~~ using three-steps validation. These steps are: (1) precipitation products with rain gauges, (2) simulated discharge with streamflow gauge, and (3) inferred storage with satellite-driven storage. Therefore, the lack of information was compensated by the use of available Earth observation. ~~From~~ First, from the evaluation of five selected rainfall products, CHIRPS was found to be the best-performing product in time (monthly) and space when compared with ten rain gauge stations, followed by PERSIANN-CDR, ARC2, GPCC, and ERA5, respectively. ~~Considering~~ Second, considering the top three products, HBV-light was calibrated and validated which resulted in high performance when the model was forced by CHIRPS, ARC2, and PERSIANN-CDR, respectively. ~~During~~ Third, during the operation period (2020 onwards), CHIRPS outperformed PERSIANN-CDR and ARC2 in capturing the filling properties (dates and magnitudes). The retrieved three years of GERD storage were compared with Landsat and Sentinel imagery, CHIRPS was the only product that produced positive NSE values thus the lowest RMSE values. Subsequently, the selection of the rainfall product was found to have more influence on inferring GERD filling strategies than parameter uncertainty. Additionally, the best simulation of CHIRPS (i.e., simulation with the highest NSE value during calibration), was found to have an NSE of 0.77 and 0.86, and RMSE of 1.7 ~~BCM~~ and 1.52 BCM when compared with Landsat and Sentinel, respectively. It inferred that the friction of the monthly retained inflow was 14% for July 2020, 41% for July 2021, and 37% and 32% for July and August 2022, respectively. Moreover, around 5.2% and 7.4% of the annual inflow to Sudan were retained by the dam in 2020 and 2021, respectively. While after the third phase, in 2022, the reservoir retained between 12.9% and 13.7% of the annual inflow reaching around 22.6% of the full storage capacity. Furthermore, it was revealed that the first three years started to impact the flow downstream. During filling months, the inflow to Sudan significantly decreased below average (2002 - 2019), suggesting the need for adapting reservoir management of the Roseires and Sennar dams to meet downstream demands.

560 *Data availability.* The publicly available data used in this study are the forcing data of the hydrological modelling (i.e., rainfall, temperature, and potential evapotranspiration) as well as the satellite imagery. Landsat Collection 2 Level 2 was retrieved from https://earthexplorer.usgs.gov/. The datasets of CHIRPS, PERSIANN-CDR and ERA5 were downloaded from the Climate Engine research app (Huntington et al., 2017). The data of ARC2 is available in https://ftp.cpc.ncep.noaa.gov. Additionally, the GPCC dataset can be downloaded from https://opendata.dwd.de. For the rain gauge data, it is not publicaly available but can be obtained from the Ethiopian National Meteorological Agency website

565 http://www.ethiomet.gov.et/ (navigate to "Data Service" and then select "Station Information" a request form will appear once a station has been selected). The hydrological data within the Nile system is not publicly available due to governmental restrictions. However, the flow data collected and maintained by the Nile Waters Directorate of the Ministry of Irrigation and Water Resources, Sudan, was made available for this study. Because we cannot share this data directly, we provide the successful HBV-light parameter sets and runs (NSE > 0.75) as a best approximation of the observed flow at https://doi.org/10.4211/hs.ed4530307dda435e9d3dcdb74da86a30 in addition to the E-A-S

570 relationship derived from SRTM DEM (Ali, 2022)

Appendix A: Satellite-based reservoir information
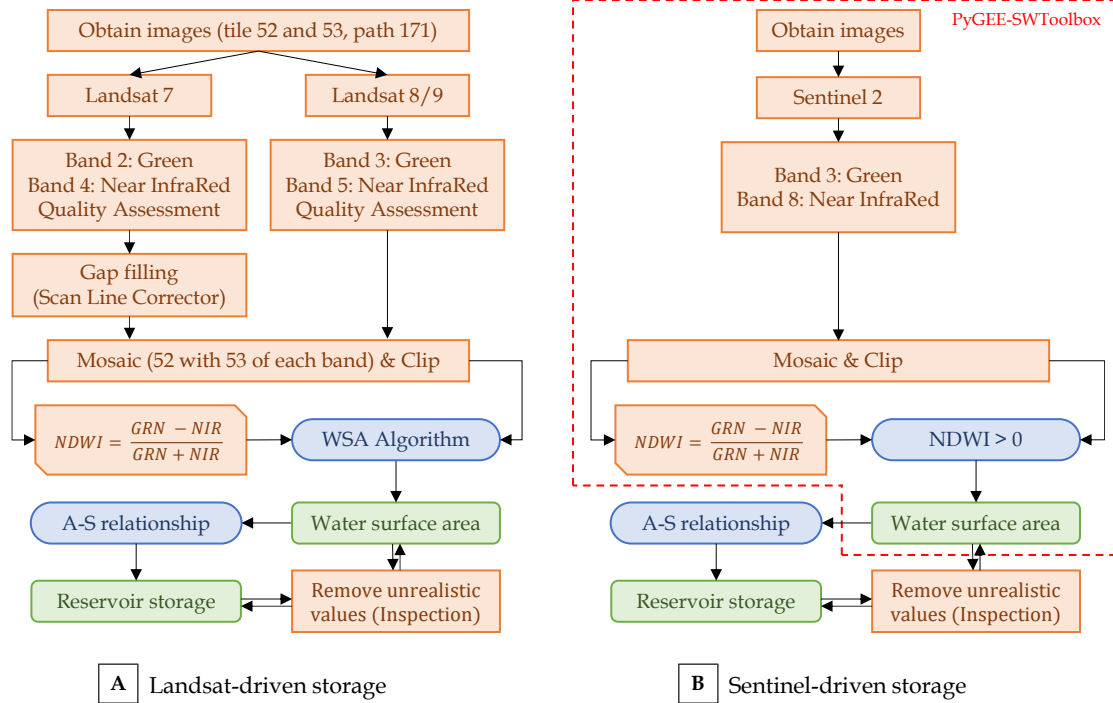
Figure A1. Procedure to derive GERD reservoir characteristics using Landsat (A) and Sentinel (B). Landsat water surface area (WSA) was estimated based on the modified algorithm by Vu et al. (2022). Fixed NDWI threshold was used to estimate WSA for Sentinel images using PyGEE-SWToolbox (Owusu et al., 2022).

Appendix B: Additional data on performance metrics

Table B1. Statistical performance metrics applied in the evaluation of the satellite rainfall products and/or the simulated discharge and reservoir storage.

Name	Symbol/Formula	Optimal value
Root Mean Square Error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - O_i)^2}$	0
RMSE-observations Standard deviation Ratio	$RSR = \frac{RMSE}{\sigma_o}$	0
Mean Bias Error	$MBE = \frac{1}{N}\sum_{i=1}^{N}(S_i - O_i)$	0
Mean Absolute Error	$MAE = \frac{1}{N}\sum_{i=1}^{N}|S_i - O_i|$	0
Percentage Bias	$PBIAS = \frac{\sum_{i=1}^{N}(S_i - O_i)}{\sum_{i=1}^{N}O_i} \times 100$	100
Nash-Sutcliffe Efficiency	$NSE = 1 - \frac{\sum_{i=1}^{N}(S_i - O_i)^2}{\sum_{i=1}^{N}(O_i - \mu_o)^2}$	1
Coefficient of Determination	$R^2 = \left(\frac{\sum_{i=1}^{N}(O_i - \mu_o)(S_i - \mu_s)}{\sqrt{\sum_{i=1}^{N}(O_i - \mu_o)^2(S_i - \mu_s)^2}}\right)^2$	0
Slope	$Slope = \frac{(\sum_{i=1}^{N} O_i \cdot S_i) - \frac{(\sum_{i=1}^{N} O_i)(\sum_{i=1}^{N} S_i)}{N}}{(\sum_{i=1}^{N} O_i^2) - \frac{(\sum_{i=1}^{N} O_i)^2}{N}}$	1
Intercept	$Intercept = \mu_s - Slope \times \mu_o$	0

O = Observed value, S = Gridded product or simulated value, N = Number of samples, μ_o = Mean value of O, μ_s = Mean value of S, σ_o = Standard deviation value of O, σ_s = Standard deviation value of S. The coefficient of determination, intercept and slope are computed using linear fit.
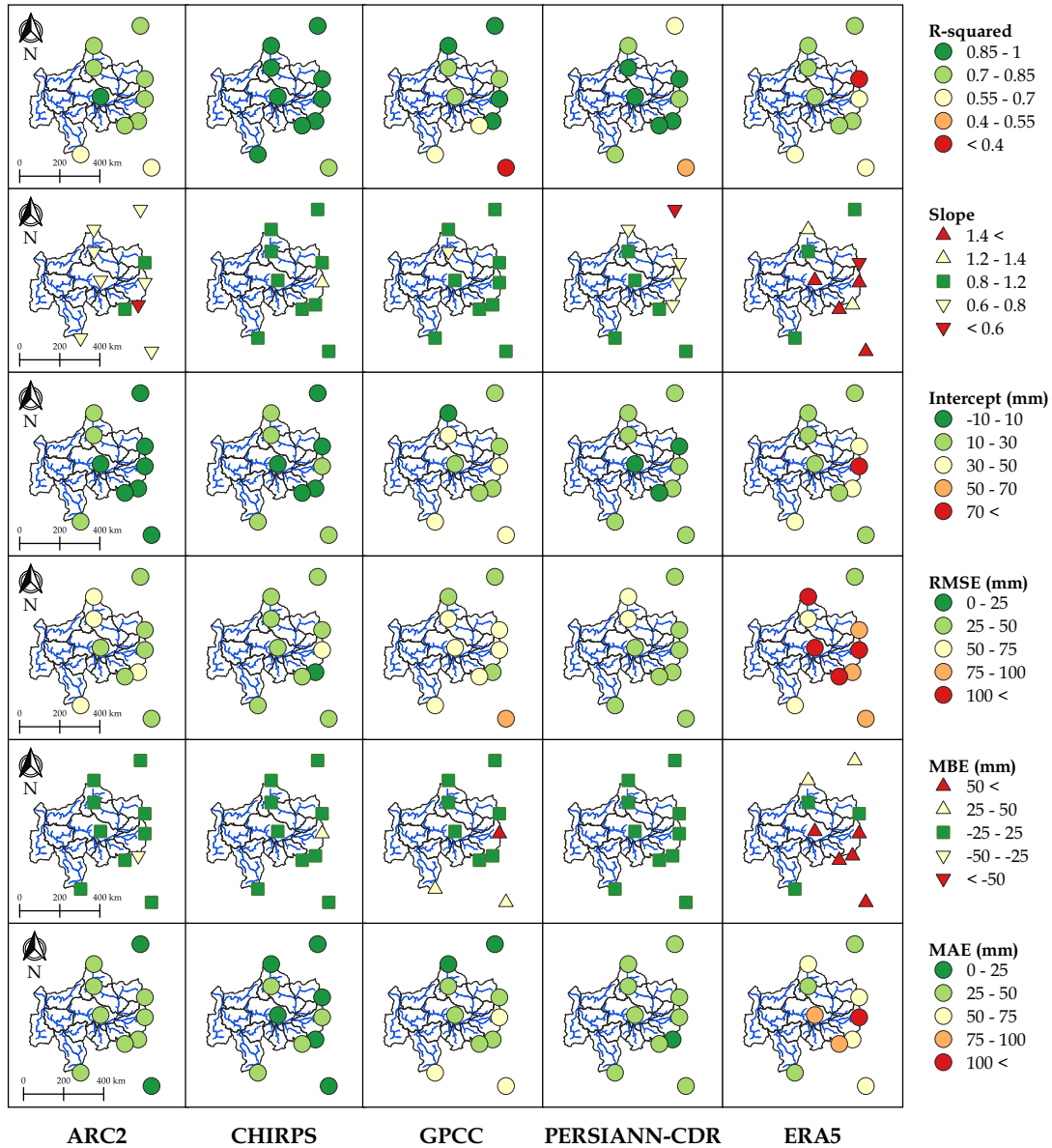
Figure B1. Spatial distribution of the statistical performance metrics over the Upper Blue Nile basin based on the point-to-pixel approach at a monthly scale from 1984 to 2005. Six out of the ten stations are outside the UBN boundaries. However, all are included in the analysis to strengthen the evaluation of the rainfall products over the region. Basins boundaries and drainage network are obtained from the Hy-droSHEDS dataset (Lehner et al., 2008).
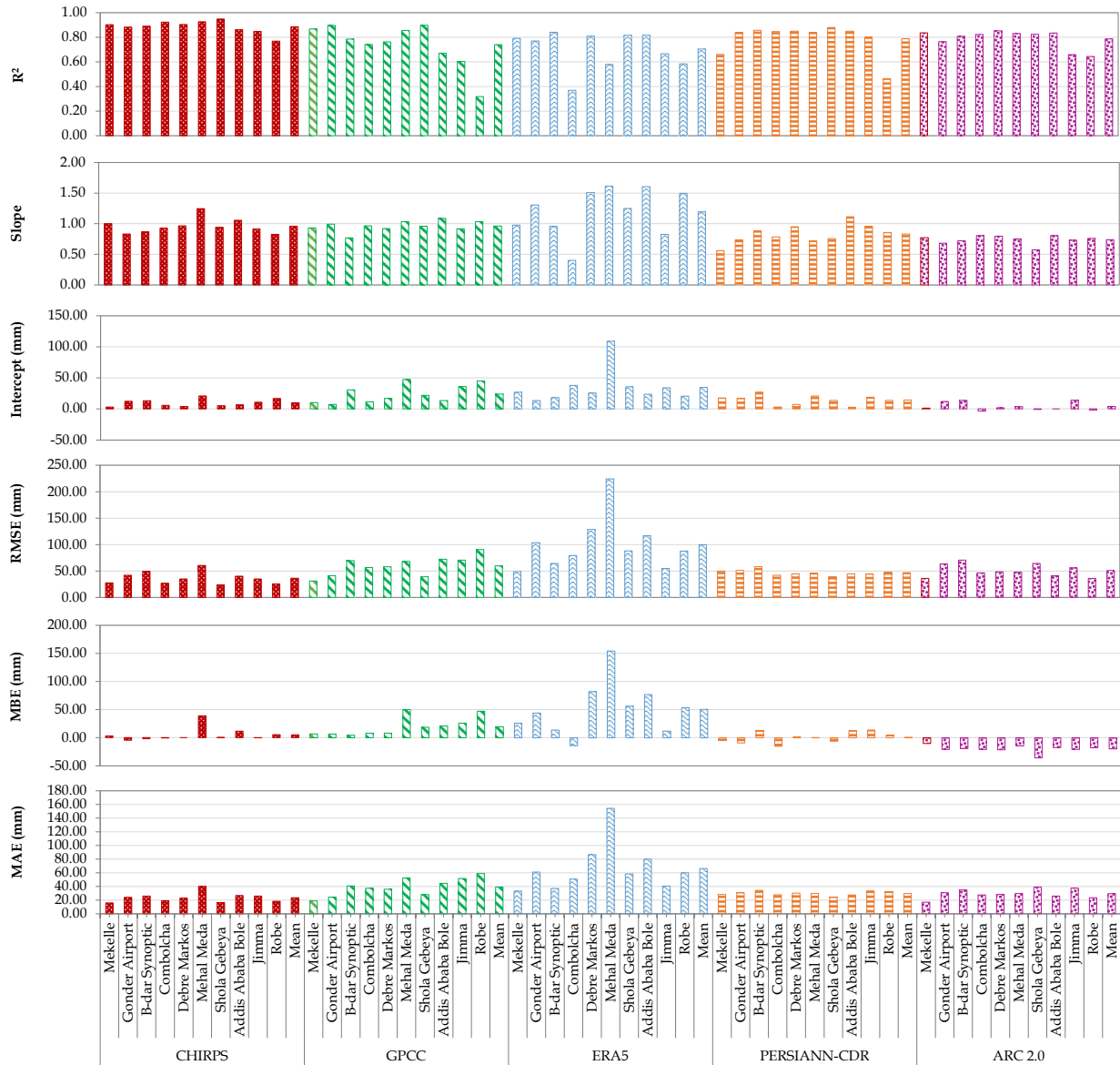
Figure B2. Statistical performance metrics of five remote sensing rainfall products evaluated at ten stations (coordinates in Table 3).

Appendix C: Hydrological models applied in the Upper Blue Nile basin

Table C1. Hydrological models applied over the Upper Blue Nile basin from 57 studies in the literature.

No.	Model Name	No. of studies	References
1	SWAT	23	Betrie et al. (2009); Schmidt and Zemadim (2015); Wosenie (2015); Lemann et al. (2016); Roth and Lemann (2016); Chakilu and Moges (2017); Lemann et al. (2017); Polanco et al. (2017); Tegegne et al. (2017); Teklesadik et al. (2017); Woldesenbet et al. (2017); Lemann et al. (2018); Woldesenbet et al. (2018); Boru et al. (2019); Kessete et al. (2019); Nigussie et al. (2019); Sultana et al. (2019); Teshome et al. (2019); Adem et al. (2020); Sinshaw et al. (2020); Bizuneh et al. (2021); Getachew et al. (2021); Mengistu et al. (2021)
2	HBV	13	Ymeti (2007); Gragne et al. (2008); Wale et al. (2008); Uhlenbrook et al. (2010); Asnik (2015); Gebre et al. (2015); Meresa and Gatachew (2015); Teklesadik et al. (2017); Worqlul et al. (2017); Bihonegn et al. (2020); Bizuneh et al. (2021); Wubneh et al. (2022a, b)
3	HEC-HMS	5	Gebre (2015); Gebre and Ludwig (2015); Agegn (2016); Zelelew and Melesse (2018); Bihonegn et al. (2020)
4	PED	5	Enku et al. (2014); Zimale et al. (2016); Worqlul et al. (2017); Akale et al. (2019); Bihonegn et al. (2020)
5	CREST	3	Lakew et al. (2017); Lakew (2020); Lakew and Moges (2021)
6	SWIM	2	Aich et al. (2014); Teklesadik et al. (2017)
7	GR4J	2	Meresa and Gatachew (2015); Tegegne et al. (2017)
8	Thornthwaite and Mather (1955)	2	Collick et al. (2009); Legesse (2009)
9	FLexB	1	Wosenie (2015)
10	LAPSUS_D	1	Getahun (2016)
11	mHM	1	Teklesadik et al. (2017)
12	VIC	1	Teklesadik et al. (2017)
13	WaterGAP3	1	Teklesadik et al. (2017)
14	CREST- SVAS	1	Lazin et al. (2020)
15	IHACRES	1	Tegegne et al. (2017)
16	JGrass-NewAge	1	Abera et al. (2017)
17	PCRaster	1	Tekleab et al. (2015)
18	Hydro-BEAM	1	Abd-El Moneim et al. (2017)
19	HBV-light-WEAP21	1	Asitatikie and Gebeyehu (2021)
20	GBHM	1	Abdel-Aziz (2014)
21	MCDE	1	Nigussie et al. (2019)
22	SAC-SMA	1	Ymeti (2007)
23	RIBASIM	1	Ghorab et al. (2013)
24	HMETS	1	Meresa and Gatachew (2015)
25	TOPMODEL	1	Deginet (2008)
26	HBV-RIBASIM	1	Booij et al. (2011)

Appendix D: HBV-light simulaitons

Table D1. HBV-Light parameters, selected parameter ranges for Monte Carlo Simulations and optimized values of each selected satellite rainfall product.

Parameter	Description	Unit	Range		Optimize values		
			Min.	Max.	CHIRPS	PERSIANN-CDR	ARC2
Soil moisture routine							
FC	Maximum soil moisture	mm	200	1000	992.26	938.10	974.28
LP	Soil moisture threshold for evaporation reduction	-	0.5	0.7	0.528	0.510	0.520
BETA	Shape coefficient	-	1	4	1.78	1.43	1.06
Groundwater and response routine							
PERC	Maximal flow from upper to lower GW-box	mm d^{-1}	1.4	2.8	1.71	1.85	1.48
UZL	Threshold for K0-outflow	mm	10.2	25.6	22.13	14.78	22.09
K0	Recession coefficient	d^{-1}	0.05	0.2	0.188	0.073	0.119
K1	Recession coefficient	d^{-1}	0.01	0.2	0.130	0.119	0.086
K2	Recession coefficient	d^{-1}	0.006	0.05	0.041	0.049	0.049
Routing routine							
MAXBAS	Length of weighting function	d	1.5	2.9	2.14	1.73	2.76
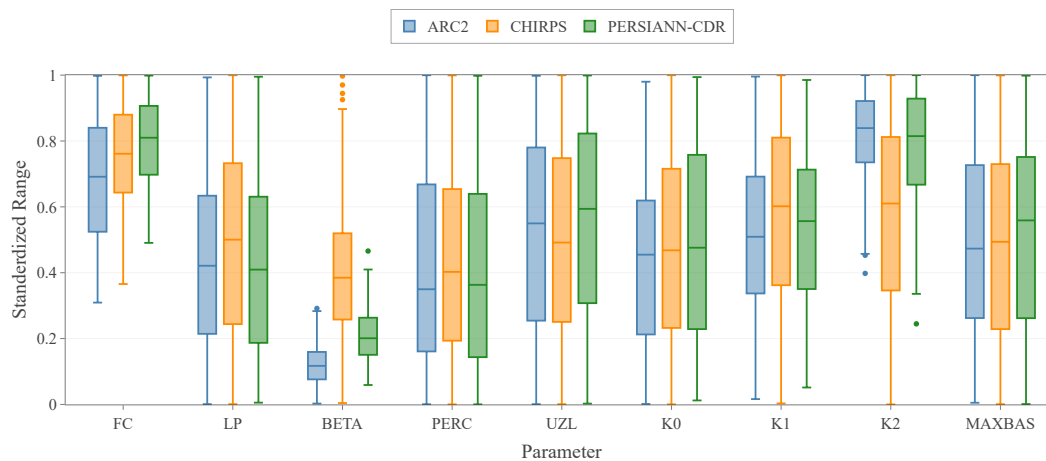
Figure D1. The standardized sensitive value ranges of HBV-light parameters for the simulations that classified as very good (i.e., NSE > 0.75) when using ARC2, CHIRPS, and PERSIANN-CDR as forcing data.
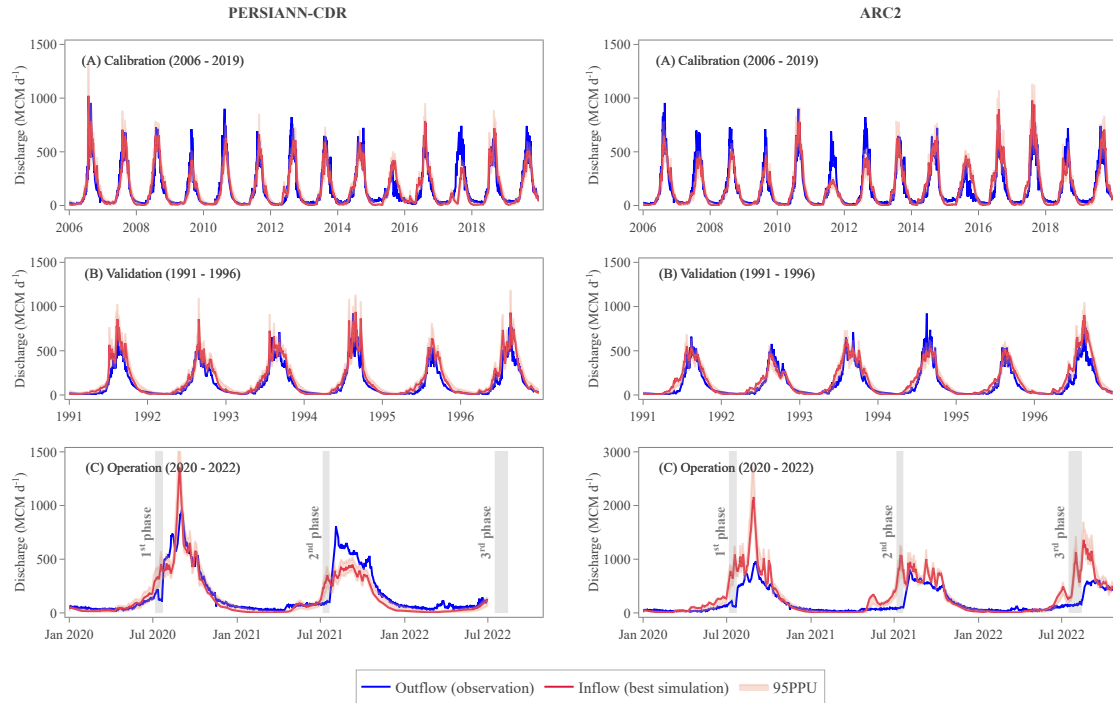
Figure D2. Temporal dynamics of daily simulated and observed discharge at the Upper Blue Nile basin outlet. Simulations forced with PERSIANN-CDR (left column) and ARC2 (right column) are shown during (A) Calibration period, (B) Validation period, and (C) Operation period. The best simulation was based on the parameter sets that achieved the highest NSE value during calibration. The 95% Prediction Uncertainty (95PPU) represents the 95th percentile of all very good simulations (NSE > 0.75) that was obtained by random generation of the parameter sets. The vertical grey shaded areas in the operation period roughly indicate the days in which the dam is filling.
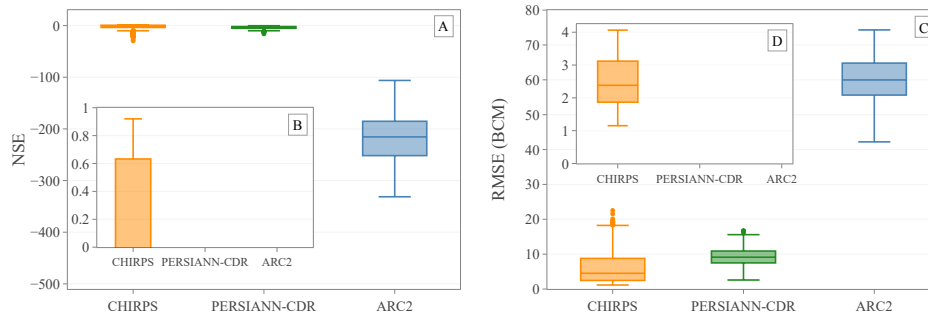
Figure D3. HBV-light performance in retrieving the storage volume of GERD from 2020 to 2022. (A, B) The Nash-Sutcliffe Efficiency (NSE) and (C, D) the Root Mean Square Error (RMSE) between the satellite observed storage (using Sentinel) and the simulated GERD reservoir storage (using HBV-light). Panel (B) is a zoom-in of (A) showing NSE between 0 and 1 while (D) presents the RMSE of all simulations with NSE > 0. The x-axis is the three tested rainfall products for input uncertainty while the whiskers demonstrate HBV-light parameter uncertainty.
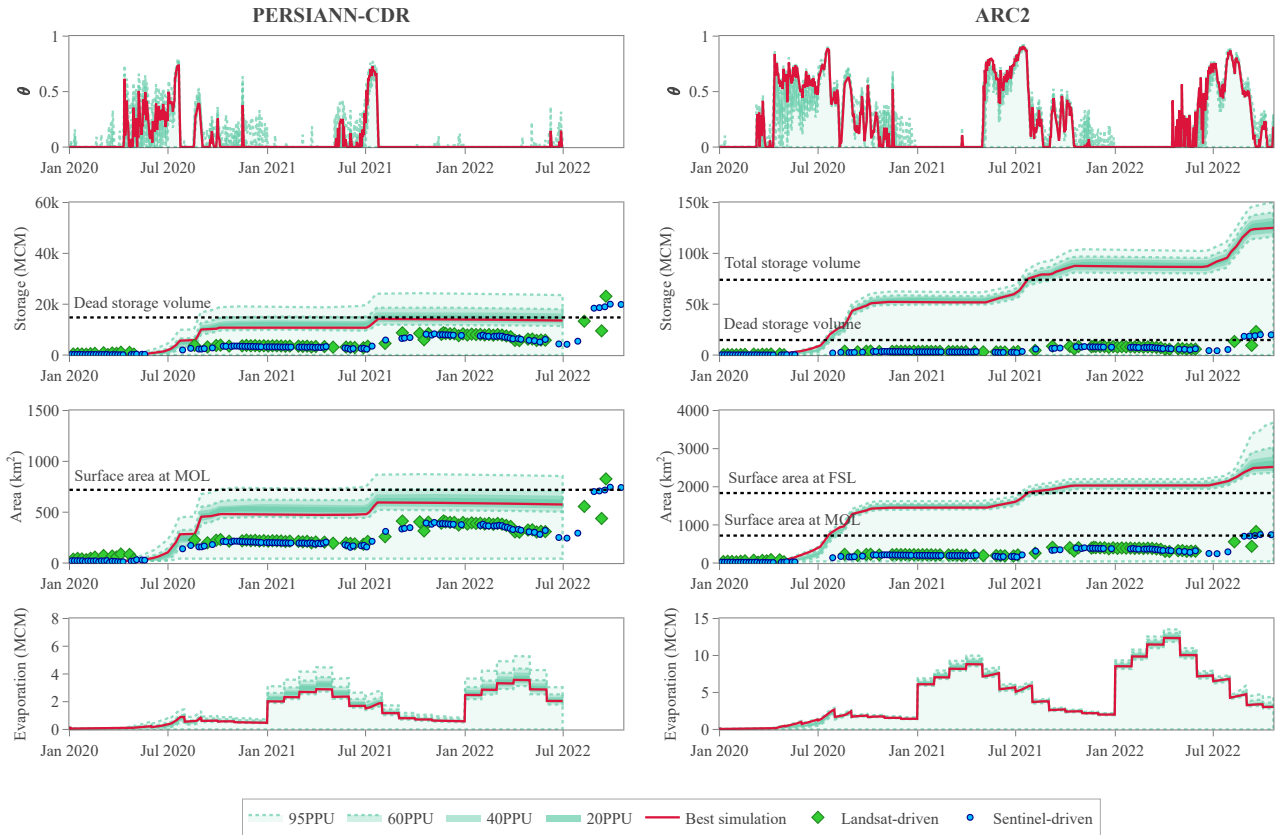
Figure D4. Temporal dynamics of inferred GERD filling strategies. The first three years of the Grand Ethiopian Renaissance Dam (GERD) filling strategies were retrieved from HBV-light lumped model using PERSIANN-CDR (left column) and ARC2 (right column). The first row is the daily fraction of inflow volume retained by GERD. The second and third rows are the daily water storage and the corresponding water surface area of the GERD reservoir, respectively. The last row is the daily evaporated volume of water from the reservoir surface. The prediction uncertainty ranges indicate the 95, 60, 40, and 20 percentiles of the results obtained from the simulations classified as very good which were 269 for PERSIANN-CDR and 244 for ARC2.

Author contributions. AMA was responsible for data curation, formal analysis, investigation, methodology development, acquiring resources, software, validation of data collection and analyses, visualization of results and writing of the original manuscript. AJT and LAM supervised the study which started as AMA's thesis towards a Master of Science degree. AJT acted as the primary supervisor and LAM as the secondary. All the authors contributed to the conceptualization, reviewing, and editing of the paper.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Hydrology and Earth System Sciences, and the authors have also no other competing interests to declare.

Acknowledgements. This study was carried out as a master thesis by Awad M. Ali who is supported by the African Scholarship Programme (ASP) offered by the Wageningen University Fellowship Programme (WUFP). We also thank the Ministry of Irrigation and Water Resources of Sudan for providing the discharge data at Eldiem.

575

580

References

Abd-El Moneim, H., Soliman, M. R., and Moghazy, H. M.: Numerical simulation of Blue Nile Basin using distributed hydrological model, in: 11th international conference on the role of engineering towards a better environment (RETBE'17), 2017.

Abdel-Aziz, O. R.: Flood forecasting in Blue Nile basin using a process-based hydrological model, International Journal of Environment, 3, 10–21, https://doi.org/10.3126/ije.v3i1.9938, 2014.

Abera, W., Formetta, G., Brocca, L., and Rigon, R.: Modeling the water budget of the Upper Blue Nile basin using the JGrass-NewAge model system and satellite data, Hydrology and Earth System Sciences, 21, 3145–3165, https://doi.org/10.5194/hess-21-3145-2017, 2017.

Adam, H., AW, A., and Hata, T.: The Future of Participatory Water Management in Gezira Scheme, Sudan, in: Proceedings of the International Workshop on Participatory Management of Irrigation Systems, Cairo, Egypt, pp. 2–4, 2003.

Adem, A. A., Dile, Y. T., Worqlul, A. W., Ayana, E. K., Tilahun, S. A., and Steenhuis, T. S.: Assessing digital soil inventories for predicting streamflow in the headwaters of the Blue Nile, Hydrology, 7, 8, https://doi.org/10.3390/hydrology7010008, 2020.

Africa News: Ethiopia dismisses rumours Nile mega-dam will starve Egypt and Sudan, https://www.africanews.com/2022/02/21/ethiopia-dismisses-rumours-nile-mega-dam-will-starve-egypt-and-sudan/, accessed December 2, 2022, 2022.

Agegn, G. K.: Climate change impact on stream flow of upper blue nile basin: based on CMIP5 climate model outputs, 2016.

Ahmed, A., Osman, I., and Babiker, A.: Sediment and aquatic weeds management challenges in Gezira irrigated scheme, in: International Sediment Initiative Conference, pp. 12–15, 2006.

Aich, V., Liersch, S., Vetter, T., Huang, S., Tecklenburg, J., Hoffmann, P., Koch, H., Fournet, S., Krysanova, V., Müller, E., et al.: Comparing impacts of climate change on streamflow in four large African river basins, Hydrology and Earth System Sciences, 18, 1305–1321, https://doi.org/10.5194/hess-18-1305-2014, 2014.

Akale, A. T., Dagnew, D. C., Moges, M. A., Tilahun, S. A., and Steenhuis, T. S.: The effect of landscape interventions on groundwater flow and surface runoff in a watershed in the upper reaches of the Blue Nile, Water, 11, 2188, https://doi.org/10.3390/w11102188, 2019.

Ali, A. M.: Data underlying the paper "Inferring reservoir filling strategies under limited data availability using hydrological modelling and Earth observation: the case of the Grand Ethiopian Renaissance Dam (GERD)", https://doi.org/10.4211/hs.ed4530307dda435e9d3dcdb74da86a30, 2022.

Alrajoula, M. T., Al Zayed, I. S., Elagib, N. A., and Hamdi, M. R.: Hydrological, socio-economic and reservoir alterations of Er Roseires Dam in Sudan, Science of the Total Environment, 566, 938–948, https://doi.org/10.1016/j.scitotenv.2016.05.029, 2016.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies, Bulletin of the American Meteorological Society, 96, 69–83, https://doi.org/10.1175/BAMS-D-13-00068.1, 2015.

Asitatikie, A. N. and Gebeyehu, W. Z.: Assessment of hydrology and optimal water allocation under changing climate conditions: The case of Megech river sub basin reservoir, Upper Blue Nile Basin, Ethiopia, Modeling Earth Systems and Environment, 7, 2629–2642, https://doi.org/10.1007/s40808-020-01024-0, 2021.

Asnik, C. T.: Estimation of Discharge for Ungauged Catchments Using Rainfall-Runoff Model in Didessa Sub-Basin: The Case of Blue Nile River Basin, Ethiopia, Ph.D. thesis, Arba Minch University, 2015.

Basheer, M. and Elagib, N. A.: Performance of satellite-based and GPCC 7.0 rainfall products in an extremely data-scarce country in the Nile Basin, Atmospheric Research, 215, 128–140, https://doi.org/10.1016/j.atmosres.2018.08.028, 2019.

620   Basheer, M., Wheeler, K. G., Ribbe, L., Majdalawi, M., Abdo, G., and Zagona, E. A.: Quantifying and evaluating the impacts of cooperation in transboundary river basins on the Water-Energy-Food nexus: The Blue Nile Basin, Science of the Total Environment, 630, 1309–1323, https://doi.org/10.1016/j.scitotenv.2018.02.249, 2018.

Bergström, S. et al.: The HBV model., Computer models of watershed hydrology., pp. 443–476, 1995.

Betrie, G. D., Mohamed, Y. A., Ann van Griensven, I., and Mynett, A.: Modeling of soil erosion and sediment transport in the Blue Nile Basin
625   using the Open Model Interface Approach, Improved Water and Land Management in the Ethiopian Highlands: Its Impact on Downstream Stakeholders Dependent on the Blue Nile, p. 132, 2009.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of hydrology, 249, 11–29, https://doi.org/10.1016/S0022-1694(01)00421-8, 2001.

Bihonegn, B. G., Moges, M. A., Mulu, G. F., and Sinshaw, B. G.: Evaluation of Stream Flow Prediction Capability of Hydrological Models
630   in the Upper Blue Nile Basin, Ethiopia, in: International Conference on Advances of Science and Technology, pp. 210–227, Springer, https://doi.org/10.1007/978-3-030-43690-2_14, 2020.

Birkett, C., Reynolds, C., Beckley, B., and Doorn, B.: From research to operations: The USDA global reservoir and lake monitor, in: Coastal altimetry, pp. 19–50, Springer, https://doi.org/10.1007/978-3-642-12796-0_2, 2011.

Biswas, A. K.: Management of transboundary waters: an overview, Management of transboundary rivers and lakes, pp. 1–20,
635   https://doi.org/10.1007/978-3-540-74928-8_1, 2008.

Bizuneh, B. B., Moges, M. A., Sinshaw, B. G., and Kerebih, M. S.: SWAT and HBV models' response to streamflow estimation in the upper Blue Nile Basin, Ethiopia, Water-Energy Nexus, 4, 41–53, https://doi.org/10.1016/j.wen.2021.03.001, 2021.

Bonnema, M. and Hossain, F.: Assessing the potential of the surface water and ocean topography mission for reservoir monitoring in the Mekong River Basin, Water Resources Research, 55, 444–461, https://doi.org/10.1029/2018WR023743, 2019.

640   Booij, M. J., Tollenaar, D., van Beek, E., and Kwadijk, J. C.: Simulating impacts of climate change on river discharges in the Nile basin, Physics and Chemistry of the Earth, Parts A/B/C, 36, 696–709, https://doi.org/10.1016/j.pce.2011.07.042, 2011.

Boru, G. F., Gonfa, Z. B., and Diga, G. M.: Impacts of climate change on stream flow and water availability in Anger sub-basin, Nile Basin of Ethiopia, Sustainable Water Resources Management, 5, 1755–1764, https://doi.org/10.1007/s40899-019-00327-0, 2019.

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N.-E., Herold, M., and Fritz, S.: Copernicus global land service:
645   Land cover 100m: collection 3: epoch 2019: Globe, Version V3. 0.1)[Data set], https://doi.org/10.5281/zenodo.3518038, 2020.

Chakilu, G. and Moges, M.: Assessing the land use/cover dynamics and its impact on the low flow of Gumara Watershed, Upper Blue Nile Basin, Ethiopia, Hydrol Current Res, 7, https://doi.org/10.4172/2157-7587.1000268, 2017.

Chen, Y., Ahmed, M., Tangdamrongsub, N., and Murgulet, D.: Reservoir-Induced Land Deformation: Case Study from the Grand Ethiopian Renaissance Dam, Remote Sensing, 13, 874, https://doi.org/10.3390/rs13050874, 2021.

650   Collick, A. S., Easton, Z. M., Ashagrie, T., Biruk, B., Tilahun, S., Adgo, E., Awulachew, S. B., Zeleke, G., and Steenhuis, T. S.: A simple semi-distributed water balance model for the Ethiopian highlands, Hydrological Processes: An International Journal, 23, 3718–3727, https://doi.org/10.1002/hyp.7517, 2009.

Conway, D.: The climate and hydrology of the Upper Blue Nile River, Geographical Journal, 166, 49–62, https://doi.org/10.1111/j.1475-4959.2000.tb00006.x, 2000.

655   De Stefano, L., Petersen-Perlman, J. D., Sproles, E. A., Eynard, J., and Wolf, A. T.: Assessment of transboundary river basins for potential hydro-political tensions, Global Environmental Change, 45, 35–46, https://doi.org/10.1016/j.gloenvcha.2017.04.008, 2017.

Deginet, M. D.: Land surface representation for regional rainfall-runoff modelling, upper Blue Nile basin, Ethiopia, 2008.

Draper, S. E.: Introduction to transboundary water sharing, Journal of water resources planning and management, 133, 377–381, https://doi.org/10.1061/(ASCE)0733-9496(2007)133:5(377), 2007.

660 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., et al.: Sentinel-2: ESA's optical high-resolution mission for GMES operational services, Remote sensing of Environment, 120, 25–36, https://doi.org/10.1016/j.rse.2011.11.026, 2012.

Dumont, H. J.: The Nile river system, in: The Ecology of River Systems, pp. 61–88, Springer, https://doi.org/10.1007/978-94-017-3290-1_3, 1986.

665 Elagib, N. A. and Mansell, M. G.: New approaches for estimating global solar radiation across Sudan, Energy Conversion and Management, 41, 419–434, https://doi.org/10.1016/S0196-8904(99)00123-5, 2000.

Eldardiry, H. and Hossain, F.: Understanding reservoir operating rules in the transboundary nile river basin using macroscale hydrologic modeling with satellite measurements, Journal of Hydrometeorology, 20, 2253–2269, https://doi.org/10.1175/JHM-D-19-0058.1, 2019.

Enku, T., Tadesse, A., Yilak, D., Gessesse, A., Addisie, M., Abate, M., Zimale, F., Moges, M., Tilahun, S., and Steenhuis, T.: Biohydrology
670 of low flows in the humid Ethiopian highlands: The Gilgel Abay catchment, Biologia, 69, 1502–1509, https://doi.org/10.2478/s11756-014-0462-9, 2014.

Ezega News: Power Generation Capacity of GERD Slashed to 5150MW – Ethiopian Minister, https://www.ezega.com/News/NewsDetails/7331/Power-Generation-Capacity-of-GERD-Slashed-to-5150MW-Ethiopian-Minister, accessed July 4, 2022, 2019.

675 Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., et al.: The shuttle radar topography mission, Reviews of geophysics, 45, https://doi.org/10.1029/2005RG000183, 2007.

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., et al.: A quasi-global precipitation time series for drought monitoring, US Geological Survey data series, 832, 1–12, https://doi.org/10.3133/ds832, 2014.

680 GADM, G. A. A.: GADM database of global administrative areas, 2012.

Gao, H., Birkett, C., and Lettenmaier, D. P.: Global monitoring of large reservoir storage from satellite remote sensing, Water Resources Research, 48, https://doi.org/10.1029/2012WR012063, 2012.

Gebre, S., Tadele, K., and Mariam, B.: Potential impacts of climate change on the hydrology and water resources availability of Didessa Catchment, Blue Nile River Basin, Ethiopia, J. Geol. Geosci, 4, 193, https://doi.org/10.4172/2329-6755.1000193, 2015.

685 Gebre, S. L.: Application of the HEC-HMS model for runoff simulation of Upper Blue Nile River Basin, Hydrology: Current Research, 6, 1, https://doi.org/10.4172/2157-7587.1000199, 2015.

Gebre, S. L. and Ludwig, F.: Hydrological response to climate change of the upper blue Nile River Basin: based on IPCC fifth assessment report (AR5), Journal of Climatology & Weather Forecasting, 3, 1–15, https://doi.org/10.4172/2332-2594.1000121, 2015.

Gebrehiwet, K.: Hydro-hegemony, an antiquated notion, in the contemporary Nile River basin: the rise of water utilization in up-stream
690 riparian countries, Heliyon, 6, e04 877, https://doi.org/10.1016/j.heliyon.2020.e04877, 2020.

Getachew, B., Manjunatha, B., and Bhat, H. G.: Modeling projected impacts of climate and land use/land cover changes on hydrological responses in the Lake Tana Basin, upper Blue Nile River Basin, Ethiopia, Journal of Hydrology, 595, 125 974, https://doi.org/10.1016/j.jhydrol.2021.125974, 2021.

Getahun, M. M.: Adapting LAPSUS_D model to simulate runoff and sediment yield in Minizr catchment, NW Ethiopia, Sustaining reservoir
695 use through sediment trapping in NW Ethiopia, p. 87, 2016.

Ghorab, E., Helal, Y. E., Mahgoub, S., Sobeih, M., and Shalash, A.: Numerical investigations to the effects of the proposed dams on the blue nile basin discharge, ERJ. Engineering Research Journal, 36, 69–77, 2013.

Gragne, A., Uhlenbrook, S., Mohammed, Y., and Kebede, S.: Catchment modeling and model transferability in upper Blue Nile basin, Lake Tana, Ethiopia, Hydrology and Earth System Sciences Discussions, 5, 811–842, 2008.

700 Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., et al.: SoilGrids250m: Global gridded soil information based on machine learning, PLoS one, 12, e0169 748, https://doi.org/10.1371/journal.pone.0169748, 2017.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Sabater, J. M., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., et al.: Global reanalysis: goodbye ERA-Interim, hello ERA5, ECMWF newsletter, 159, 17–24, https://doi.org/10.21957/vf291hehd7, 2019.

705 Hou, J., van Dijk, A. I., Beck, H. E., Renzullo, L. J., and Wada, Y.: Remotely sensed reservoir water storage dynamics (1984–2015) and the influence of climate variability and management at a global scale, Hydrology and Earth System Sciences, 26, 3785–3803, https://doi.org/10.5194/hess-26-3785-2022, 2022.

Hrachowitz, M., Savenije, H., Blöschl, G., McDonnell, J., Sivapalan, M., Pomeroy, J., Arheimer, B., Blume, T., Clark, M., Ehret, U., et al.: A decade of Predictions in Ungauged Basins (PUB)—a review, Hydrological sciences journal, 58, 1198–1255,
710 https://doi.org/10.1080/02626667.2013.803183, 2013.

Huntington, J. L., Hegewisch, K. C., Daudert, B., Morton, C. G., Abatzoglou, J. T., McEvoy, D. J., and Erickson, T.: Climate engine: Cloud computing and visualization of climate and remote sensing data for advanced natural resource monitoring and process understanding, Bulletin of the American Meteorological Society, 98, 2397–2410, https://doi.org/10.1175/BAMS-D-15-00324.1, 2017.

IPoE: International Panel of Experts on the Grand Ethiopian Renaissance Dam Project Final Report, 2013.

715 Jansen, K. F., Teuling, A. J., Craig, J. R., Dal Molin, M., Knoben, W. J., Parajka, J., Vis, M., and Melsen, L. A.: Mimicry of a Conceptual Hydrological Model (HBV): What's in a Name?, Water Resources Research, 57, e2020WR029 143, https://doi.org/10.1029/2020WR029143, 2021.

Kansara, P., Li, W., El-Askary, H., Lakshmi, V., Piechota, T., Struppa, D., and Abdelaty Sayed, M.: An assessment of the filling process of the Grand Ethiopian Renaissance Dam and its impact on the downstream countries, Remote Sensing, 13, 711,
720 https://doi.org/10.3390/rs13040711, 2021.

Kessete, N., Moges, M. A., and Steenhuis, T. S.: Evaluating the applicability and scalability of bias corrected CFSR climate data for hydrological modeling in upper Blue Nile basin, Ethiopia, in: Extreme Hydrology and Climate Variability, pp. 11–22, Elsevier, https://doi.org/10.1016/B978-0-12-815998-9.00002-6, 2019.

Khairy, W., El-Motasem, M., Mehanna, A., and Hefny, K.: Estimation of evaporation losses from water bodies in the Sudan and Ethiopia,
725 International Journal of Energy and Water Resources, 3, 233–246, https://doi.org/10.1007/s42108-019-00031-x, 2019.

Kim, U. and Kaluarachchi, J. J.: Climate change impacts on water resources in the upper blue Nile River Basin, Ethiopia 1, JAWRA Journal of the American Water Resources Association, 45, 1361–1378, https://doi.org/10.1111/j.1752-1688.2009.00369.x, 2009.

Lakew, H. B.: Investigating the effectiveness of bias correction and merging MSWEP with gauged rainfall for the hydrological simulation of the upper Blue Nile basin, Journal of Hydrology: Regional Studies, 32, 100 741, https://doi.org/10.1016/j.ejrh.2020.100741, 2020.

730 Lakew, H. B. and Moges, S. A.: Comparison of Spatial Interpolation Techniques of Rainfall for Hydrological Applications in a Complex Mountainous Region of the Upper Blue Nile Basin, in: Nile and Grand Ethiopian Renaissance Dam, pp. 461–474, Springer, https://doi.org/10.1007/978-3-030-76437-1_24, 2021.

Lakew, H. B., Moges, S. A., and Asfaw, D. H.: Hydrological evaluation of satellite and reanalysis precipitation products in the Upper Blue Nile Basin: A case study of Gilgel Abbay, Hydrology, 4, 39, https://doi.org/10.3390/hydrology4030039, 2017.

735 Lazin, R., Shen, X., Koukoula, M., and Anagnostou, E.: Evaluation of the hyper-resolution model-derived water cycle components over the upper Blue Nile Basin, Journal of Hydrology, 590, 125 231, https://doi.org/10.1016/j.jhydrol.2020.125231, 2020.

Legesse, E. S.: Modeling rainfall-runoff relationships for the Anjeni watershed in the Blue Nile basin, Ph.D. thesis, Citeseer, 2009.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, Transactions American Geophysical Union, 89, 93–94, https://doi.org/10.1029/2008EO100001, 2008.

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

Lemann, T., Roth, V., and Zeleke, G.: Impact of precipitation and temperature changes on hydrological responses of small-scale catchments in the Ethiopian Highlands, Hydrological sciences journal, 62, 270–282, https://doi.org/10.1080/02626667.2016.1217415, 2017.

745 Lemann, T., Roth, V., Zeleke, G., Subhatu, A., Kassawmar, T., and Hurni, H.: Spatial and temporal variability in hydrological responses of the Upper Blue Nile basin, Ethiopia, Water, 11, 21, https://doi.org/10.3390/w11010021, 2018.

McNamara, I., Baez-Villanueva, O. M., Zomorodian, A., Ayyad, S., Zambrano-Bigiarini, M., Zaroug, M., Mersha, A., Nauditt, A., Mbuliro, M., Wamala, S., et al.: How well do gridded precipitation and actual evapotranspiration products represent the key water balance components in the Nile Basin?, Journal of Hydrology: Regional Studies, 37, 100 884, https://doi.org/10.1016/j.ejrh.2021.100884, 2021.

750 Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J., Clark, M. P., Uijlenhoet, R., and Teuling, A. J.: Mapping (dis) agreement in hydrologic projections, Hydrology and Earth System Sciences, 22, 1775–1791, https://doi.org/10.5194/hess-22-1775-2018, 2018.

Mengistu, D., Bewket, W., Dosio, A., and Panitz, H.-J.: Climate change impacts on water resources in the upper blue nile (Abay) river basin, ethiopia, Journal of Hydrology, 592, 125 614, https://doi.org/10.1016/j.jhydrol.2020.125614, 2021.

Meresa, H. and Gatachew, M.: Modeling of hydrological extremes under climate change scenarios in the Upper Blue Nile River Basin, 755 Ethiopia, Journal of Civil & Environmental Engineering, 6, 252, https://doi.org/10.4172/2165-784X.1000252, 2015.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Transactions of the ASABE, 50, 885–900, https://doi.org/10.13031/2013.23153, 2007.

NASA Earth Observatory: Record Flooding in Sudan, https://earthobservatory.nasa.gov/images/147288/record-flooding-in-sudan, accessed 760 July 22, 2022, 2020.

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

Nigussie, G., Moges, M. A., Moges, M. M., and Steenhuis, T. S.: Assessment of suitable land for surface irrigation in ungauged catchments: Blue Nile basin, Ethiopia, Water, 11, 1465, https://doi.org/10.3390/w11071465, 2019.

765 Nonki, R. M., Lenouo, A., Lennard, C. J., Tshimanga, R. M., and Tchawoua, C.: Comparison between dynamic and static sensitivity analysis approaches for impact assessment of different potential evapotranspiration methods on hydrological models' performance, Journal of Hydrometeorology, 22, 2713–2730, https://doi.org/10.1175/JHM-D-20-0192.1, 2021.

Novella, N. S. and Thiaw, W. M.: African rainfall climatology version 2 for famine early warning systems, Journal of Applied meteorology and Climatology, 52, 588–606, https://doi.org/10.1175/JAMC-D-11-0238.1, 2013.

770 Owusu, C., Snigdha, N. J., Martin, M. T., and Kalyanapu, A. J.: PyGEE-SWToolbox: A Python Jupyter Notebook Toolbox for Interactive Surface Water Mapping and Analysis Using Google Earth Engine, Sustainability, 14, 2557, https://doi.org/10.3390/su14052557, 2022.

Polanco, E. I., Fleifle, A., Ludwig, R., and Disse, M.: Improving SWAT model performance in the upper Blue Nile Basin using meteorological data integration and subcatchment discretization, Hydrology and Earth System Sciences, 21, 4907–4926, https://doi.org/10.5194/hess-21-4907-2017, 2017.

775 QGIS, D. T.: QGIS Geographic Information System, Open Source Geospatial Foundation, 2009.

Roth, V. and Lemann, T.: Comparing CFSR and conventional weather data for discharge and soil loss modelling with SWAT in small catchments in the Ethiopian Highlands, Hydrology and Earth System Sciences, 20, 921–934, https://doi.org/10.5194/hess-20-921-2016, 2016.

Salama, A., ElGabry, M., El-Qady, G., and Moussa, H. H.: Evaluation of Grand Ethiopian Renaissance Dam Lake Using Remote Sensing
780 Data and GIS, Water, 14, 3033, https://doi.org/10.3390/w14193033, 2022.

Scaramuzza, P. and Barsi, J.: Landsat 7 scan line corrector-off gap-filled product development, in: Proceeding of Pecora, vol. 16, pp. 23–27, 2005.

Schmidt, E. and Zemadim, B.: Expanding sustainable land management in Ethiopia: Scenarios for improved agricultural water management in the Blue Nile, Agricultural Water Management, 158, 166–178, https://doi.org/10.1016/j.agwat.2015.05.001, 2015.

785 Seibert, J.: HBV light, User's manual, Uppsala University, Institute of Earth Science, Department of Hydrology, Uppsala, 1996.

Seibert, J. and Vis, M. J.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrology and Earth System Sciences, 16, 3315–3325, https://doi.org/10.5194/hess-16-3315-2012, 2012.

Setegn, S. G., Srinivasan, R., Melesse, A. M., and Dargahi, B.: SWAT model application and prediction uncertainty analysis in the Lake Tana Basin, Ethiopia, Hydrological Processes: An International Journal, 24, 357–367, https://doi.org/10.1002/hyp.7457, 2010.

790 Sinshaw, B. G., Moges, M. A., Tilahun, S. A., Dokou, Z., Moges, S., Anagnostou, E., Eshete, D. G., Kindie, A. T., Bekele, E., Asese, M., et al.: Integration of SWAT and remote sensing techniques to simulate soil moisture in data scarce micro-watersheds: a case of Awramba micro-watershed in the Upper Blue Nile Basin, Ethiopia, in: International Conference on Advances of Science and Technology, pp. 294–314, Springer, https://doi.org/10.1007/978-3-030-43690-2_20, 2020.

Stakhiv, E. Z.: Pragmatic approaches for water management under climate change uncertainty 1, JAWRA Journal of the American Water
795 Resources Association, 47, 1183–1196, https://doi.org/10.1111/j.1752-1688.2011.00589.x, 2011.

Sterl, S., Fadly, D., Liersch, S., Koch, H., and Thiery, W.: Linking solar and wind power in eastern Africa with operation of the Grand Ethiopian Renaissance Dam, Nature Energy, 6, 407–418, https://doi.org/10.1038/s41560-021-00799-5, 2021.

Sultana, R., Dhungana, S., and Bhatta, N.: Application of SWAT Model for Estimating Runoff in Upper Nile River Basin, European Journal of Engineering Science and Technology, 2, 20–35, https://doi.org/10.33422/EJEST.2019.09.35, 2019.

800 Survey, U. U. S. G.: Landsat Collection-1 Level-2, https://earthexplorer.usgs.gov/, accessed July 4, 2022, 2022.

Sutcliffe, J. V., Parks, Y. P., et al.: The hydrology of the Nile, IAHS Special Publication, 5, 1999.

Tegegne, G., Park, D. K., and Kim, Y.-O.: Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin, Journal of Hydrology: Regional Studies, 14, 49–66, https://doi.org/10.1016/j.ejrh.2017.10.002, 2017.

Tekleab, S., Mohamed, Y., and Uhlenbrook, S.: Hydro-climatic trends in the Abay/upper Blue Nile basin, Ethiopia, Physics and Chemistry
805 of the Earth, Parts A/B/C, 61, 32–42, https://doi.org/10.1016/j.pce.2013.04.017, 2013.

Tekleab, S., Uhlenbrook, S., Savenije, H. H., Mohamed, Y., and Wenninger, J.: Modelling rainfall–runoff processes of the Chemoga and Jedeb meso-scale catchments in the Abay/Upper Blue Nile basin, Ethiopia, Hydrological Sciences Journal, 60, 2029–2046, https://doi.org/10.1080/02626667.2015.1032292, 2015.

Teklesadik, A. D., Alemayehu, T., Van Griensven, A., Kumar, R., Liersch, S., Eisner, S., Tecklenburg, J., Ewunte, S., and Wang, X.: Inter-model comparison of hydrological impacts of climate change on the Upper Blue Nile basin using ensemble of hydrological models and global climate models, Climatic Change, 141, 517–532, https://doi.org/10.1007/s10584-017-1913-4, 2017.

Teshome, F., Moges, S. A., and Hailu, D.: Evaluation of globally available water resources reanalysis (WRR-1) runoff products for assessment and management water resources in the Upper Blue Nile basin: A data scarce major subbasins of the Nile basin, in: Extreme Hydrology and Climate Variability, pp. 165–173, Elsevier, https://doi.org/10.1016/B978-0-12-815998-9.00014-2, 2019.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V., and De Roo, A.: Validation of satellite-based precipitation products over sparsely gauged African river basins, Journal of Hydrometeorology, 13, 1760–1783, https://doi.org/10.1175/JHM-D-12-032.1, 2012.

Uhlenbrook, S., Mohamed, Y., and Gragne, A.: Analyzing catchment behavior through catchment modeling in the Gilgel Abay, upper Blue Nile River basin, Ethiopia, Hydrology and Earth System Sciences, 14, 2153–2165, https://doi.org/10.5194/hess-14-2153-2010, 2010.

Vu, D. T., Dang, T. D., Galelli, S., and Hossain, F.: Satellite observations reveal 13 years of reservoir filling strategies, operating rules, and hydrological alterations in the Upper Mekong River basin, Hydrology and Earth System Sciences, 26, 2345–2364, https://doi.org/10.5194/hess-26-2345-2022, 2022.

Wale, A., Rientjes, T., Dost, R., and Gieske, A.: Hydrological Balance of Lake Tana Upper Blue Nile Basin, Ethiopia, ITC The Netherlands, 2008.

Wamucii, C. N., van Oel, P. R., Ligtenberg, A., Gathenya, J. M., and Teuling, A. J.: Land use and climate change effects on water yield from East African forested water towers, Hydrology and Earth System Sciences, 25, 5641–5665, https://doi.org/10.5194/hess-25-5641-2021, 2021.

Wang, C., Si, J., Li, Z., Zhao, C., Jia, B., Celestin, S., He, X., Zhou, D., Qin, J., and Zhu, X.: Evaluation of three gridded potential evapotranspiration datasets for streamflow simulation in three inland river basins in the arid Hexi Corridor, Northwest China, Journal of Hydrology: Regional Studies, 44, 101 234, https://doi.org/10.1016/j.ejrh.2022.101234, 2022.

Wannasin, C., Brauer, C., Uijlenhoet, R., van Verseveld, W., and Weerts, A. H.: Daily flow simulation in Thailand Part I: Testing a distributed hydrological model with seamless parameter maps based on global data, Journal of Hydrology: Regional Studies, 34, 100 794, https://doi.org/10.1016/j.ejrh.2021.100794, 2021.

Warner, J. and Zawahri, N.: Hegemony and asymmetry: Multiple-chessboard games on transboundary rivers, International Environmental Agreements: Politics, Law and Economics, 12, 215–229, https://doi.org/10.1007/s10784-012-9177-y, 2012.

Wheeler, K. G., Basheer, M., Mekonnen, Z. T., Eltoum, S. O., Mersha, A., Abdo, G. M., Zagona, E. A., Hall, J. W., and Dadson, S. J.: Cooperative filling approaches for the grand Ethiopian renaissance dam, Water International, 41, 611–634, https://doi.org/10.1080/02508060.2016.1177698, 2016.

Wheeler, K. G., Hall, J. W., Abdo, G. M., Dadson, S. J., Kasprzyk, J. R., Smith, R., and Zagona, E. A.: Exploring cooperative transboundary river management strategies for the Eastern Nile Basin, Water Resources Research, 54, 9224–9254, https://doi.org/10.1029/2017WR022149, 2018.

Woldesenbet, T. A., Elagib, N. A., Ribbe, L., and Heinrich, J.: Hydrological responses to land use/cover changes in the source region of the Upper Blue Nile Basin, Ethiopia, Science of the Total Environment, 575, 724–741, https://doi.org/10.1016/j.scitotenv.2016.09.124, 2017.

Woldesenbet, T. A., Elagib, N. A., Ribbe, L., and Heinrich, J.: Catchment response to climate and land use changes in the Upper Blue Nile sub-basins, Ethiopia, Science of the Total Environment, 644, 193–206, https://doi.org/10.1016/j.scitotenv.2018.06.198, 2018.

845 World Population Review: Sudan Population 2022 (Demographics, Maps, Graphs), https://worldpopulationreview.com/countries/sudan-population, accessed July 4, 2022, 2022.

Worqlul, A. W., Yen, H., Collick, A. S., Tilahun, S. A., Langan, S., and Steenhuis, T. S.: Evaluation of CFSR, TMPA 3B42 and ground-based rainfall data as input for hydrological models, in data-scarce regions: The upper Blue Nile Basin, Ethiopia, Catena, 152, 242–251, https://doi.org/10.1016/j.catena.2017.01.019, 2017.

850 Wosenie, M. D.: Water balance modelling of the Lake Tana area, Ethiopia, Ph.D. thesis, Ghent University, 2015.

Wubneh, M. A., Fikadie, F. T., Worku, T. A., Aman, T. F., and Kifelew, M. S.: Hydrological impacts of climate change in gauged sub-watersheds of Lake Tana sub-basin (Gilgel Abay, Gumara, Megech, and Ribb) watersheds, Upper Blue Nile Basin, Ethiopia, Sustainable Water Resources Management, 8, 1–14, https://doi.org/10.1007/s40899-022-00665-6, 2022a.

Wubneh, M. A., Worku, T. A., Fikadie, F. T., Aman, T. F., and Kifelew, M. S.: Climate change impact on Lake Tana water storage, Upper
855 Blue Nile Basin, Ethiopia, Geocarto International, pp. 1–23, https://doi.org/10.1080/10106049.2022.2032397, 2022b.

Yang, G. and Block, P.: Water sharing policies conditioned on hydrologic variability to inform reservoir operations, Hydrology and Earth System Sciences, 25, 3617–3634, https://doi.org/10.5194/hess-25-3617-2021, 2021.

Ymeti, I.: Rainfall estimation by Remote Sensing for conceptual rainfall-runoff modeling in the Upper Blue Nile basin, ITC Enschede, The Netherlands, 2007.

860 Zelelew, D. G. and Melesse, A. M.: Applicability of a spatially semi-distributed hydrological model for watershed scale runoff estimation in Northwest Ethiopia, Water, 10, 923, https://doi.org/10.3390/w10070923, 2018.

Zhang, S., Gao, H., and Naz, B. S.: Monitoring reservoir storage in South Asia from multisatellite remote sensing, Water Resources Research, 50, 8927–8943, https://doi.org/10.1002/2014WR015829, 2014.

Ziese, M., Rauthe-Schöch, A., Becker, A., Finger, P., Rustemeier, E., and Schneider, U.: GPCC Full Data Daily
865 Version 2020 at 1.0°: Daily land-surface precipitation from rain-gauges built on GTS-based and historic data, https://doi.org/10.5676/DWD_GPCC/FD_D_V2020_100, 2020.

Zimale, F., Mogus, M., Alemu, M., Ayana, E., Demissie, S., Tilahun, S., and Steenhuis, T.: Calculating the sediment budget of a tropical lake in the Blue Nile basin: Lake Tana, Soil Discussions, pp. 1–32, https://doi.org/10.5194/soil-2015-84, 2016.