Journal of
Water & Climate Change

IWA PUBLISHING

Journal of Water and Climate Change Vol 13 No 2, 522 doi: 10.2166/wcc.2021.174

Performance evaluation of CFSR, MERRA-2 and TRMM3B42 data sets in simulating river discharge of data-scarce tropical catchments: a case study of Manafwa, Uganda

Maria Theresa Nakkazi [a,b], Jotham Ivan Sempewo [a,*], Martin Dahlin Tumutungire[a] and Jimmy Byakatonda[c]

[a] Civil and Environmental Engineering Department, Makerere University, P O Box 7062, Kampala, Uganda
[b] Department of Civil Engineering, KU Leuven, Kasteelpark Arenberg 40, Bus 2448, Leuven 3001, Belgium
[c] Department of Biosystems Engineering, Gulu University, P.O.Box 166, Gulu, Uganda
*Corresponding author. E-mail: jotham.sempewo@mak.ac.ug

MTN, 0000-0002-3171-1687

ABSTRACT

Data scarcity has been a huge problem in modelling catchments especially in the tropical region. Satellite data and different statistical methods are being used to improve the quality of conventional meteorological data. However, their potential needs to be further investigated. This paper evaluates the performance of three datasets in simulating discharge of River Manafwa, Uganda. Two reanalysis datasets were selected for studying both rainfall and temperature, whereas a satellite algorithm was selected for studying rainfall alone. MERRA-2 data and CFSR were chosen as the reanalysis datasets whereas TRMM3B42 data were used as the satellite product in this study. The SWAT model was used to evaluate the performance of these datasets. The model performance indicators indicated that, at daily time steps, all the three datasets produced values of Nash -Sutcliffe Efficiency (NSE$<$0.4), coefficient of determination ($R^2$$<$0.4) and Percent Bias $+$25%. Despite a general underperformance compared to MERRA-2, CFSR performed better than TRMM. On applying generated bias corrections for precipitation and temperature climate data, overall results showed that the bias-corrected data outperformed the original data. We conclude that, in the absence of gauged hydro-meteorological data, bias-corrected MERRA-2, CFSR and TRMM data could be used for simulating river discharge in data-scarce areas.

Key words: hydrological modelling, Manafwa, reanalysis datasets, SWAT

HIGHLIGHTS

- Comparing CFSR, TRMM, MERRA-2 and gauged precipitation datasets.
- Investigating the potential of using the above reanalysis datasets to run the SWAT model.
- Computing bias correction factors for the datasets using Linear Scaling and Local Intensity Scaling methods.
- Using the bias-corrected data to drive the SWAT model.
- Comparing the performance of the SWAT models before and after application of bias correction factors.

INTRODUCTION

Over the past three decades or so, the lakes and rivers of East Africa have come under increasing and considerable pressure from a variety of interlinked factors, such as human interventions, industrial pollution, eutrophication and sedimentation (Olago & Odada 2007). Land-use changes, such as deforestation, desertification, urbanization and reclamation of wetlands, have resulted in continued occurrences of flooding and drought, land degradation related to soil erosion, reduced agricultural productivity and deterioration of fragile natural ecosystems (Lorup *et al.* 1998). Therefore, it is important to investigate the water resources and their fluxes not only considering the processes within the catchments but also the hydrological processes occurring within the entire catchment (Naschen *et al.* 2018). This can provide valuable information for the development of water resource management and land-use planning strategies.

The usage of multiyear global gridded representations of weather known as reanalysis datasets has become widely accepted for hydrological and climatological modelling for catchments (Nkiaka *et al.* 2017; Bahati *et al.* 2021; Byakatonda *et al.* 2021). National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (Tarana & Slobodan

2010), Climate Forecasting System Reanalysis (CFSR) (Daniel *et al.* 2014), European Center for Medium-Range Weather Forecasts (ECMWF) ERA-Interim (Peng *et al.* 2020) and Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Gelaro *et al.* 2017) are some of the widely used datasets. Reanalysis datasets have been used as input for hydrological modelling in many research studies (Kigobe *et al.* 2011; Kisembe *et al.* 2018; Jingrui *et al.* 2020). Satellite rainfall estimates such as Tropical Rainfall Measuring Mission (TRMM) (Liu *et al.* 2012) data are also being used to provide alternative inputs for rainfall–runoff modelling. Recent examples of applications of satellite-based precipitation estimates include Najmaddin *et al.* (2017), Ouatiki *et al.* (2017) and Rossi *et al.* (2017).

Different successes in the numerous studies carried out have been recorded. For example, Ganiyu *et al.*(2017) assessed the suitability of the Global Precipitation Climatology Project (GPCP) and ERA-Interim rainfall estimates for runoff modelling and concluded that river discharge simulated with the GPCP showed a higher correlation with observed rainfall than the ERA-Interim. Elias *et al.* used CFSR, (ECMWF) ERA-Interim datasets and one global meteorological forcing dataset WATCH Forcing Data methodology applied to ERA-Interim (WFDEI). Results showed that bias-corrected WFDEI outperformed the two reanalysis datasets and CFSR performed better than the ERA-Interim. It was, therefore, concluded that, in the absence of gauged hydro-meteorological data, WFDEI and CFSR could be used for hydrological modelling in data-scarce areas such as the Sudano-Sahel region. However, the potential of these rainfall products and reanalysis datasets in effective catchment runoff simulation is still a great research challenge in the tropical region.

The effective evaluation of water resources requires the availability of high-quality hydro-meteorological data, which are inadequate in the tropical and semi-arid regions (Mango *et al.* 2011). Obtaining representative meteorological data for watershed-scale hydrological modelling can be difficult and time-consuming (Nkiaka *et al.* 2017). The difficulty in collecting data can be attributed to the following reasons: (i) lack of reliable equipment; (ii) absence of good archiving system and software to store and process the data and (iii) lack of funds to organize data collection campaigns (Gorgoglione *et al.* 2016). It is also worth mentioning that once the data have been captured and archived, accessing it is quite costly (Liu *et al.* 2008). Weather stations based on the ground do not always adequately represent the weather occurring over a watershed, because they can have gaps in their data series and can be far from the watershed of interest or recent data are not available (Daniel *et al.* 2014). Hydrological models are, therefore, designed to fill some of these gaps, and their application to enhance the water resource management is widely acknowledged (Worqlul *et al.* 2014).

In addition, Demory *et al.* (2013) argued that driving hydrological models with reanalysis datasets to reproduce observed streamflow represents one of the most accurate ways to evaluate how the hydrological cycle is simulated by reanalysis forecast models. However, before their application into the hydrological models, it is highly recommended that the correlation between observed rainfall and reanalysis precipitation estimates should be assessed first (Monteiro *et al.* 2016).

The main objective of this study was to obtain bias correction factors for precipitation and temperature datasets that can be applied to specific reanalysis datasets (CFSR and MERRA-2) and the satellite rainfall product, TRM3B42 for their application in hydrological modelling specifically for the Manafwa Catchment under study. The specific objectives pertaining to this study were (1) application of the observed data and reanalysis datasets for continuous runoff modelling and comparing the performance using statistical metrics like NSE, R^2, etc. and (2) obtaining bias corrections on gauged data from the Manafwa Catchment and applying this corrected data to see how best it can simulate river discharge. The novelty of the present study is the comparison of these three rainfall models, i.e. CFSR, TRMM and MERRA-2, which has not been done before for the tropical catchments. The study goes on to evaluate the performance of these models when subjected to the generated bias correction factors from the observed rain gauge and temperature data by seeing how best they can improve the reanalysis datasets to produce models close to those obtained by the observed data.

This study will be useful in validating the use of reanalysis datasets in runoff modelling for data-scarce catchments especially in the tropical region. The study area is an important hydrological region in Uganda and no similar studies (to the authors' knowledge) have been conducted in the past on the Manafwa Catchment. The proposed methodology is believed to contribute to the bias correction of reanalysis datasets and hence, the hydrologic simulation modelling at the river basin scale with increased accuracy. All researchers, especially those working on the hydrological models in tropical catchments (Uganda to be exact), will be empowered with the findings to use the bias corrections on reanalysis datasets to run their models instead of having to buy climatic data from the Meteorological Department. Furthermore, this research will add to our understanding of the hydrology of the Manafwa basin and give additional data sources in data-scarce locations like Uganda. This will save time and money for researchers and hydrologists in terms of data availability. The Ministry of

Water and Environment in Uganda will be in a position to strengthen the catchment planning process, and this will be a platform for further studies to be carried out on other catchments in the country.

DATA AND METHODS

Description of the study area

The Manafwa sub-catchment region is located in the tropical Mt. Elgon climatic zone, which experiences a bi-model rainfall pattern. The average annual rainfall totals about 1,500 mm with rainy seasons occurring in the months of April till June and August till November. The mean annual maximum temperature is 27 °C and the annual minimum temperature is 15 °C. The forest of the Mt. Elgon National Park is the main ground water recharge source within the catchment. The infiltration rates vary from 1.2 to 363 cm/h depending on soil characteristics. The Manafwa River is served by eight permanent tributaries. From its source to the outlet of the sub-catchment at the Manafwa Waterworks, the river has a flow length of 43 km. The geology in the Mt. Elgon region comprises mainly Pre-Cambrian and Cainozoic rock formations including volcanices, granitites and sediments. The rocks on the slopes are deeply weathered; the dominant soils are Vertisols, regionally known as 'black cotton soils'. The relief is relatively undulating and hilly with the highest catchment elevation at 3,616 m a.s.l. and the lowest at 1,083 m a.s. l.

The 470 km^2 delineated catchment is highly influenced by past volcanic activities and the soil is very variable. Generally, the soils in the highlands are clays, while those in the midlands and the lowlands are clay loams or sandy. The sub-catchment is characterized by population growth, low-income generating activities and weak infrastructural and service facilities. Ongoing deforestation due to a high firewood demand and landslides during the wet season is the main risk. The decreasing Manafwa River base flow, which is essential for the fresh water supply of the sub-catchment, as well as the downstream-located Mbale City, is one of the major concerns in the catchment (https://www.geo.fberlin.de/en/v/iwmnetwork/living_laboratories/uganda/geointro/index.html) (Figure 1).

Data sources

The study mainly dealt with reanalysis datasets, namely CFSR and Modern-Era Retrospective Analysis for Research Application; Version 2 (MERRA-2), as well as the satellite rainfall product TRMM. The aim was to obtain bias corrections from the observed meteorological data to be used when working with these datasets to run hydrological models. The SWAT model was used which was then calibrated and validated using the SWAT-CUP.

The datasets used in this study included meteorological data such as precipitation and air temperature; hydrological data including runoff; and spatial data such as Digital Elevation Model (DEM). The data collection methods to be used were specifically for the SWAT model; land cover maps were obtained from Landsat images, DEM, soil maps from the available soil data and the weather data were obtained from weather stations, which were used as the reference material to correct the reanalysis datasets. The reanalysis datasets were obtained from the available websites.

The elevation grid map at a spatial resolution of 30-m was obtained from the SERVIR-ESA. This DEM was used to delineate the watershed and define hydrological response units (HRUs). The land-use and drainage maps were obtained at a scale of 300 m from the European Space Agency (ESA). Soil data were obtained from the Food and Agricultural Organization, Harmonize World Soil Database at a spatial resolution of 1 km. Together with land-use data, it was used to generate a curve number map. For the stream data, there is only one stream gauge in the watershed; the Manafwa river gauge (station ID 82212), which is located at the outlet of the delineated watershed at 34.16°E and 0.94°N. Daily and monthly stream flow data for Manafwa (1948–2016) were obtained from the Directorate of water resources management Entebbe but with some missing data during this period. Owing to the large gaps in observed station datasets, only periods (2000–2012) having full records of data were compared with reanalysis and satellite data. These discharge data were used for model calibration and validation in the SWAT-CUP. Observed monthly and daily climatic data (precipitation, temperature, dew point, wind speed, solar radiation and relative humidity) for the period 2000–2016 were obtained from the Meteorological Department housed at the Office of the Prime Minister.

Reanalysis datasets

Reanalysis data are a continually updated gridded dataset that represents the state of the atmosphere, incorporating observations and outputs of numerical weather prediction models from past to present day (Pijush 2015). A reanalysis project makes use of a consistent modern analysis system to produce a dataset, that to a certain extent, can be regarded as a

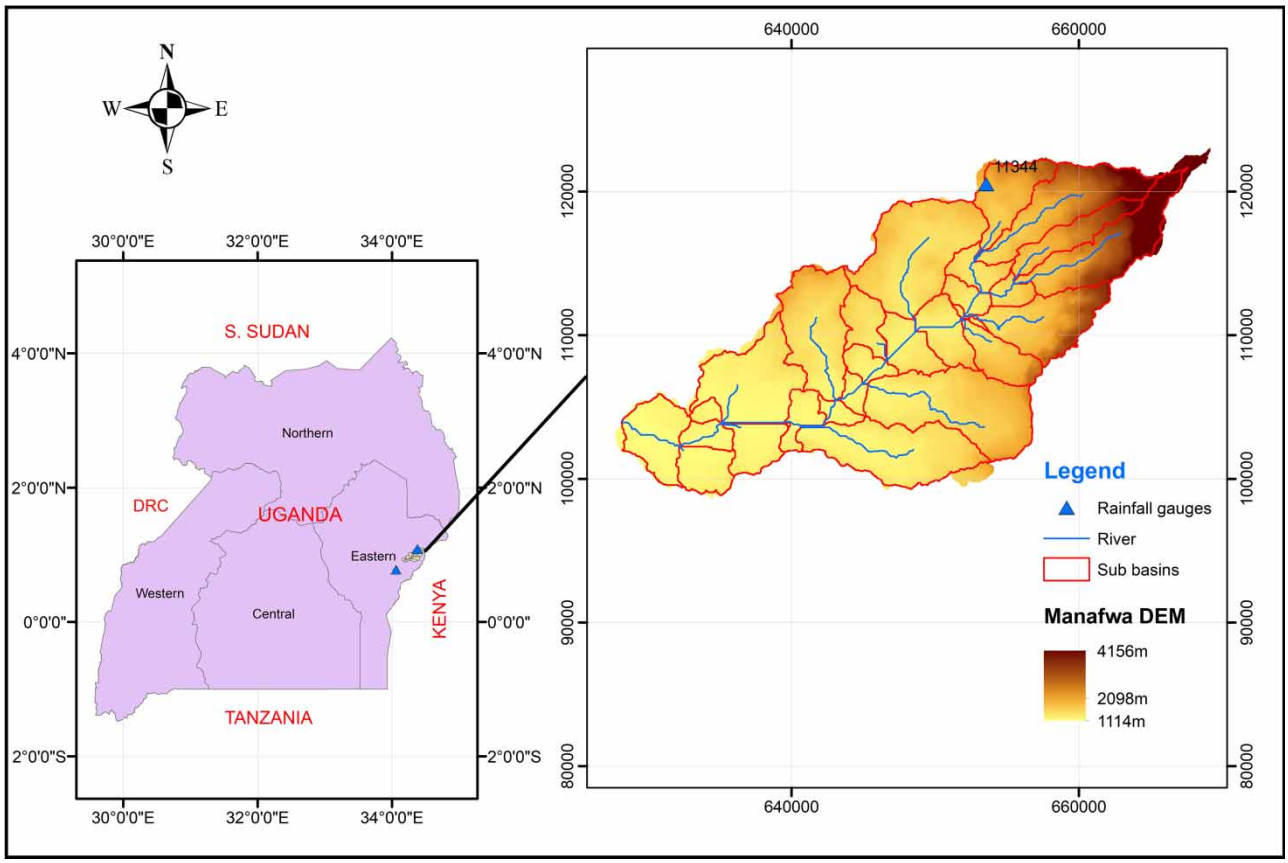

Figure 1 | Location of Manafwa watershed in Uganda.

'proxy' for observation with the advantage of providing coverage and time resolution often unobtainable with a normal observational network (Morse *et al.* 2013). It is generated with a data assimilation system combining observations with a numerical weather prediction model. The reanalysis provides a physical picture of the global climate over a period during which observational data are available.

Climate Forecasting System Reanalysis

The Climate Forecast System (CFS) is a model representing the global interaction between Earth's oceans, land and atmosphere (NCEI 2020). Produced by several dozen scientists under guidance from the NCEP, this model offers hourly and daily data with a horizontal resolution down to one-half of a degree (approximately 38 km) around Earth for many variables (Saha *et al.* 2014). CFS uses the latest scientific approaches for taking in, or assimilating, observations from data sources including surface observations, upper air balloon observations, aircraft observations and satellite observations. It spans the period from 1 January 1979 to 31 July 2014 with a spatial/temporal resolution of approximately 0.30°/6-hourly. For our study, two CFSR data-station gauges were used, i.e. Stn 8341 at 34.06°E, 0.78°N and 1103 m. The second station was Stn 8344 at 34.37°E, 0.78°N and 1349 m. The data were downloaded from https://global-weather.tamu.edu/.

Modern-Era Retrospective analysis for Research and Applications, Version 2

The MERRA-2 provides data beginning in 1980. It was introduced to replace the original MERRA dataset because of the advances made in the assimilation system that enable assimilation of modern hyperspectral radiance and microwave observations, along with GPS-Radio Occultation datasets. It also uses NASA's ozone profile observations that began in late 2004. Additional advances in both the GEOS model and the GSI assimilation system are included in MERRA-2. The spatial resolution remains about the same (0.5°×0.625° (lat×lon) or approximately 50 km) with temporal resolution up to hourly. Along

with the enhancements in the meteorological assimilation, MERRA-2 takes some significant steps towards GMAO's target of an Earth System reanalysis. MERRA-2 is the first long-term global reanalysis to assimilate space-based observations of aerosols and represent their interactions with other physical processes in the climate system. MERRA-2 includes a representation of ice sheets over (say) Greenland and Antarctica (NASA, Modern-Era Retrospective analysis for Research and Applications, Version 2; Project Overview, 2020). MERRA-2 reanalysis data were obtained from https://gmao.gsfc.nasa.gov/reanalaysis/MERRA-2/. Only one station was used for our study area located at 34.29°E, 1.00°N and an average elevation for 1/2°×1/2° lat/lon region of 1,568.84 m.

TRMM 3B42 satellite product

The TRMM is a joint space mission between NASA and Japan's National Space Development Agency designed to monitor and study tropical and subtropical precipitation and the associated release of energy. The mission uses five instruments: Precipitation Radar (PR), TRMM Microwave Imager (TMI), Visible Infrared Scanner, Clouds & Earths Radiant Energy System and Lightning Imaging Sensor. The TMI and PR are the main instruments used for precipitation. These instruments are used in an algorithm that forms the TRMM Combined Instrument (TCI) calibration dataset (TRMM 2B31) for the TRMM Multisatellite Precipitation Analysis (TMPA), whose TMPA 3B43 monthly precipitation averages and TMPA 3B42 daily and subdaily (3 h) averages are probably the most relevant TRMM-related products for climate research. 3B42 and 3B43 are available in 0.25° spatial resolution, covering 50°N to 50°S for 1998–present (NCAR 2020). TRMM 3B42 data were obtained from https://giovanni.gsfc.nasa.gov/giovanni. The data used in this study were covering a geographical range of '33.60°E, 0.17°N, 33.99°E, 0.64°N'.

Methodology

The main objective of this paper is to obtain bias correction factors for precipitation and temperature datasets that can be applied to CFSR, TRMM and MERRA-2 datasets for their application in hydrological modelling. Details of the methodology followed are presented in Figure 2.

For this study, the datasets are used to set up the SWAT models which are then calibrated and validated in the SWAT-CUP and the performance was evaluated using Nash–Sutcliffe efficiency (NSE), coefficient of determination (R^2) and Percent Bias (PBIAS) criteria. Bias correction factors are then generated for precipitation and temperature data and these are used on the original reanalysis and satellite data. The new SWAT models with bias-corrected data are set up and these are calibrated and validated. Their performance is also evaluated using the same criteria as earlier mentioned.

Prior to using these datasets for hydrological modelling, CFSR and TRMM 3B42 were resampled to horizontal 0.5×0.5 grid sizes using bilinear interpolation, a typical approach in meteorology and climate studies, to obtain a uniform spatial

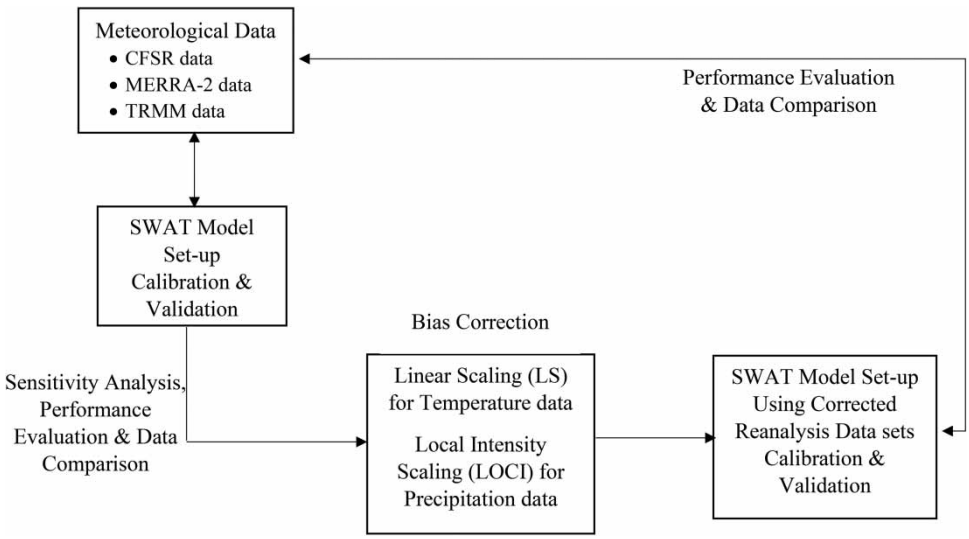

Figure 2 | Study methodology.

resolution (Bahati *et al.* 2021). Resampling the MERRA-2 reanalysis data would add inaccuracies due to the irregular spacing of x and y coordinates across the grid points; hence, this was not done. As a result, rain gauge data were compared directly to the nearest MERRA-2 grid points in the original resolution, without resampling.

Model setup

The impacts of corrected precipitation and temperature in hydrological processes are evaluated by the SWAT hydrological model. River discharge at the specified gauge station along the Manafwa River was simulated using the SWAT (Gassman *et al.* 2014) in the ArcSWAT interface. In the SWAT model, the cycle is divided into two steps. The first is the simulation of hydrologic transformation of rainfall in catchment, which is closely connected with the movements of water, sediments and substances from the slopes of watershed into streams. The second one is the simulation of hydraulic transformation in the streams, which solves the problems of the moving water, sediments and substances in reaching towards the outlet of the given sub-basin. In the land phase of the hydrological cycle, the SWAT simulates the hydrological cycle based on the water balance (Equation (1)),

$$\mathrm{SW}_t = \mathrm{SW}_0 + \sum_{i=1}^{t} (R_{\mathrm{day}} - Q_{\mathrm{surf}} - E_{\mathrm{a}} - W_{\mathrm{seep}} - Q_{\mathrm{gw}}) \tag{1}$$

where SW_t is the final soil water content (mm), SW_0 is the initial water content (mm), t is time (days), R_{day} is the amount of precipitation on day i (mm) Q_{surf} is the amount of surface water runoff on day i (mm), E_a is the amount of actual transpiration on day i (mm), W_{seep} is the amount of water entering the vadose zone from the soil profile on day i (mm) and Q_{gw} is the amount of return flow on day i (mm). Details of equations and methods used to estimate various hydrological components can be found in Neitsch *et al.* (2011).

The SCS (Soil Conservation Service) curve number procedure method was used for estimating the surface runoff using daily rainfall, the SWAT simulates surface runoff volumes and peak runoff rates for each HRU. In this study, the SCS curve number method was used to estimate the surface runoff. The SCS curve number is shown in the following equation

$$Q_{\mathrm{surf}} = \frac{(R_{\mathrm{day}} - 0.2S)^2}{R_{\mathrm{day}} + 0.8S)} \tag{2}$$

where Q_{surf} is the accumulated runoff or rainfall excess (mm), R_{day} is the rainfall depth for the day (mm), S is the retention parameter (mm).

SCS defines three antecedent moisture conditions: I – dry (wilting point), II – average moisture and III– wet (field capacity). The moisture condition I curve number is the lowest value the daily curve number can assume in dry conditions. The curve numbers for moisture conditions I and III are calculated with the following equations,, respectively.

$$\mathrm{CN}_1 = \mathrm{CN}_2 - \frac{20 \times (100 - \mathrm{CN}_2)}{(100 - \mathrm{CN}_2 + \exp[2.533 - 0.0636 \times (100 - \mathrm{CN}_2)])} \tag{3}$$

$$\mathrm{CN}_3 = \mathrm{CN}_2 \times \exp[0.00673(100 - \mathrm{CN}_2)] \tag{4}$$

where CN_1 is the moisture condition I curve number, CN_2 is the moisture condition II curve number and CN_3 is the moisture condition III curve number. The retention parameter is defined by the following equation;

$$S = 25.4\left(\frac{1000}{\mathrm{CN}} - 10\right) \tag{5}$$

During model development, the SWAT divides a catchment into sub-catchments using DEM data. The spatial distribution of hydrological processes over each sub-catchment is represented through hydrologic response units (HRUs), used to further divide the sub-catchments into smaller units. The HRU can be defined as a land area within a sub-catchment with the

same land-use class, soil type, slope class and management combinations. 171 HRUs and 27 sub-basins were created over a watershed area of 470.07 km^2. The observed weather data were for 12 years (2000–2012) with 2 warmup years. Three soil types were generated, namely Sandy clay loam-Fe47-2ab-501, Loam-Nh2-2c-848 and Loam-Tm9-2c-948. The land cover categories were also five, namely agricultural land generic, forest mixed, forest deciduous, range-grasses and Barren. Separate models were developed for each of the reanalysis datasets using the same threshold values.

Criteria for model calibration, performance evaluation and comparison

After setting up the model within the ArcSWAT 2012, the model was calibrated and validated using the Sequential Uncertainty Fitting ver. 2 (SUFI-2) algorithm in the SWAT-CUP (version 5.1.6.2) (Arnold *et al.* 2012), basically following the guidelines (Abbaspour 2015). Based on the Latin hypercube sampling method and the SUFI-2 method, we obtained a set of sensitive parameters and their ranges. This process can be done in the SWAT-CUP software with 95% prediction uncertainty (95PPU). Based on the sensitive parameters and their ranges, the reasonable values of the model parameters were able to be calibrated manually based on the literature.

From the literature, a total of 19 parameters were initially taken into consideration (Mutenyo *et al.* 2013). The first simulation (500) included all the datasets and after that, a sensitivity analysis was done. Less sensitive parameters were discarded. Table 1 shows the final model parameters considered.

Calibration was performed in two steps. First, auto-calibration was performed with the 19 most sensitive parameters as input using the parameter solution (Parasol) method. Model parameters were further improved by manual calibration using an approach suggested by Arnold *et al.* (2012). Two types of sensitivity analysis were generally performed: local, by changing values one at a time, and global, by allowing all parameter values to change. Among the 19 initial parameters initially identified from the literature and after 500 simulations, 10 parameters were formed to be sensitive during the local analysis and ranked using the global sensitivity analysis run for 500 simulations with the most sensitive having a *p*-value of 0.05. Calibration involved running SUFI-2 up to five iterations of 500 simulations each while adjusting the new parameter sets within reasonable and practical ranges. To validate the model, the model was run with 500 simulations using the parameter ranges that were determined in the best iteration during the calibration process and comparing the predictions to observed data not used in the calibration.

The NSE, R^2 and Percent-Bias (PBIAS) [Equations (6)–(8), respectively] were used to evaluate the performance of the model under the different rainfall inputs. The NSE is used to assess the predictive capacity of the model and measures how well the observed and simulated flows match. Its value ranges from $-\infty$ to 1 with values close to 1 indicating high model performance. The R^2 measures how well the observed data are correlated to the simulated data and varies from 0 to 1 with values closer to 1 also indicating high model performance. PBIAS indicates the average tendency of the simulated flows to be over/underestimated than observed flows with absolute low values indicating accurate model simulation. Positive values indicate model underestimation while negative values indicate overestimation (Nkiaka *et al.* 2017). Model calibration was aimed at achieving a satisfactory model efficiency of concurrently having NSE$\geq$0.5, PBIAS of $\pm$25% and R^2 of $\geq$0.5 for a

Table 1 | Monthly calibrated parameters for the SWAT

No.	Parameter	Description
1	CN2.mgt	SCS curve number for moisture condition II
2	ALPHA_BF.gw	Base flow alpha factor
3	GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)
4	CH_K2.rte	Effective hydraulic conductivity in the main channel alluvium
5	SOL_AWC(..).sol	Available water capacity of the soil (mm)
6	SOL_Z().sol	Soil depth (mm)
7	ESCO.hru	Soil evaporation compensation factor
8	REVAPMN.gw	Threshold depth of water in the shallow aquifer for 'revap' or percolation to the deep aquifer to occur (mm H_2O)

monthly time step as recommended by Moriasi *et al.* (2007)

$$\mathrm{NSE} = 1.0 - \frac{\sum_{i=1}^{n} (\mathrm{O}_i - \mathrm{P}_i)^2}{\sum_{i=1}^{n} (\mathrm{O}_i - \bar{\mathrm{O}})^2} \tag{6}$$

$$R^2 = \left\{ \frac{\sum_{\mathrm{i}=1}^{N} (\mathrm{O_i} - \bar{\mathrm{O}})(\mathrm{P_i} - \bar{\mathrm{P}})}{\left[\sum_{\mathrm{i}=1}^{N} (\mathrm{O_i} - \bar{\mathrm{O}})^2\right]^{0.5} \left[\sum_{\mathrm{i}=1}^{N} (\mathrm{P_i} - \bar{\mathrm{P}})^2\right]^{0.5}} \right\} \tag{7}$$

$$\mathrm{PBIAS} = \left[\frac{\sum_{\mathrm{i}=1}^{n} (\mathrm{O_i} - \mathrm{P_i}) * 100}{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{O_i}} \right] \tag{8}$$

where $\mathrm{O_i}$ is the observed data, $\mathrm{P_i}$ is the predicted data and $\bar{\mathrm{O}}$ is the mean of observed values and n is the total number of observations.

The degree of uncertainty in the calibrated model(s) was quantified using the *p*-factor and *r*-factor. The *p*-factor represents the percentage of observed streamflow bracketed by the 95PPU while the *r*-factor is the average width of the 95PPU. The 95PPU is calculated at the 2.5 and 97.5% confidence interval of observed streamflow obtained through Latin hypercube sampling. In SUFI-2, the goal is to minimize the width of the uncertainty band and enclose as many observations as possible because these observations are a result of all processes taking place in the catchment (Abbaspour 2015). The *p*-factor can range from 0 to 1, while the optimal *r*-factor value is 0, indicating that the model outputs are completely reliable. However, an *r*-factor of 0 indicates that the 95PPU band contains fewer flow measurements.

Bias correction of the reanalysis and satellite datasets

A product containing historical (1981–2010) precipitation data, which generally contains biases when compared to observations, is the best performing reanalysis (Bahati *et al.* 2021). Therefore, bias correction was carried out to correct the reanalysis and satellite precipitation and temperature data. Two bias corrections were used for this study. For precipitation, the Local Intensity Scaling (LOCI) method was used, whereas for temperature, the Linear Scaling (LS) method was used. Both are conducted on a monthly basis for each calendar month during the period 2000–2012.

The LS method implements a constant corrected factor that is estimated by the difference between original RCM simulations and observations for each calendar month. This method aims to perfectly match the long-term monthly mean of corrected values with those of the observed values (Maikel *et al.* 2020). This approach is capable of perfectly adjusting for climatic factors when monthly mean values are included (Teutschbein & Seibert 2013). Precipitation is adjusted with a multiplier and temperature is corrected by the additive term, as the following equations demonstrate respectively.

$$P_{hst,m,d}^{cor} = P_{hst,m,d} \times \left[\frac{\mu(P_{obs,m})}{\mu(P_{hst,m}} \right] \tag{9}$$

$$T_{hst,m,d}^{cor} = T_{hst,m,d} + [\mu(T_{obs,m}) - \mu(T_{hst,m})] \tag{10}$$

where $P_{hst,m,d}^{cor}$ and $T_{hst,m,d}^{cor}$ denote the corrected precipitation and temperature on the d-th day of the m-th month, respectively. $P_{hst,m,d}$, $T_{hst,m,d}$, respectively, denote the precipitation and temperature from the original RCM outputs during the relevant period; the subscripts d and m are specific days and months, respectively, whereas μ denotes the mean value. The LS method operates with monthly correction values based on the differences between the observed control and the uncorrected data. The LS method is the simplest bias correction method that corrects the climate data upon the basis of the difference between reanalysis datasets and observations. Although it can adjust mean values, it cannot be used in the analysis of extreme events because the unique scaling factor in a specific month often leads to heavy precipitation being greatly underestimated (Haerter *et al.* 2011).

LOCI method for precipitation

The LOCI method (Schmidli *et al.* 2006) aims to simultaneously correct the precipitation intensity and frequency. First, the rainfall intensity threshold ($P_{thres,m}$) for each month is confirmed. Accordingly, the number of wet days in RCM precipitation that exceed this threshold matches the number of days for which observed precipitation was determined (Teutschbein & Seibert 2013). The advantage of this bias-corrected method is that it effectively eliminates the drizzle effect. This is because too many drizzly days are often included in original RCM outputs (Fowler *et al.* 2007).

A scaling factor (s_m) as illustrated in the following equation is then calculated to ensure that the mean amounts of corrected precipitation are equal to observations.

$$\frac{\mu(P_{obs,m,d}|P_{obs,m,d} > 0)}{\mu(P_{hst,m,d}|P_{hst,m,d} > thres_m)} \tag{11}$$

The LOCI can adjust the mean as well as both wet-day frequencies and wet-day intensities of rainfall time series (Fang *et al.* 2015). For this study, a spreadsheet that uses the statistical downscaling method known as LOCI for rainfall by Shrestha (2015) was used.

The LOCI method is an extension of the LS method that effectively corrects the precipitation frequency. No bias can be found in wet-day frequency or intensity. When compared to the LS method, heavy precipitation is also partly corrected, and this is attributable to the correction in wet-day frequency.

RESULTS AND DISCUSSION

Comparison of the precipitation and temperature datasets

First, we evaluated the performance of the CFSR, MERRA-2 and TRMM rainfall datasets against the rain gauge data on a monthly scale. Meng *et al.* (2014) insisted that the rainfall comparison of the different reanalysis datasets should strongly be considered for the purpose of hydrological applications. The areal observed rainfall is, therefore, compared with the areal rainfall datasets for the Manafwa Catchment over the 2000–2012 period. Figure 3 shows the average monthly rainfall for the gauged rainfall, TRMM, CFSR and MERRA-2 datasets in the 2000–2012 period. Figures S-1 and S-4 (supplementary figures) illustrate the comparison between monthly and daily scale times of rainfall for all the weather datasets at the individual gauges before bias correction.

It can be seen that MERRA-2 and CFSR-8341 station estimates are close to those of the rain gauge. This is confirmed by the statistical values (including mean, standard deviation and maximum rainfall) shown in Table 2, as well as the R^2, NSE and PBIAS of the gridded rainfall data against rain gauges data during the period of analysis at monthly scales are shown in Table 4. It can be seen that MERRA-2 data have a quite similar trend as the rain gauges data, as shown in Figure 3. Aside from that, the CFSR-8341 data have also captured the rain gauges data for the study area satisfactory. However, the fit between CFSR-8344 data and rain gauge data is poor. TRMM datasets are generally low compared to all the other datasets.

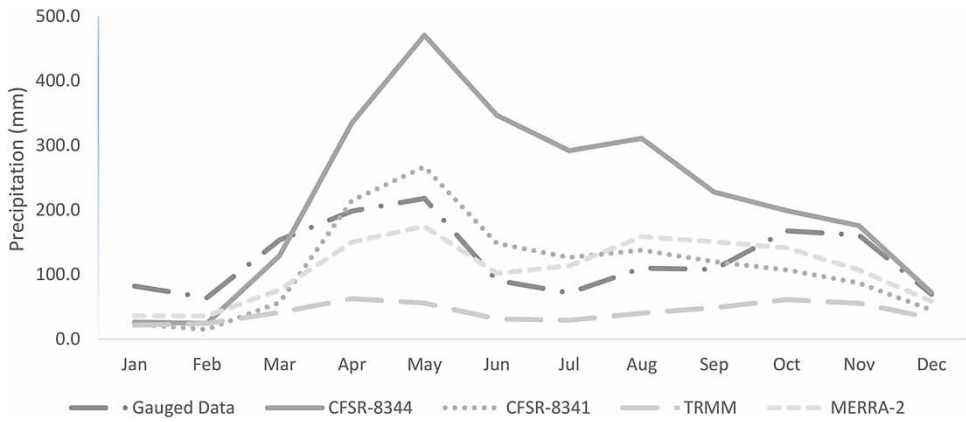

Figure 3 | Average monthly rainfall of rain gauges and reanalysis rainfall products in the 2000–2012 period.

Table 2 | Statistical values of areal averaged monthly rainfall of rain gauges data, CFSR AND MERRA-2 reanalysis datasets as well as TRMM satellite product for the period of 2000–2012

Rainfall dataset	Mean	SD	Max
Gauged data	124.5	50.9	217.9
CFSR-8344	217.5	133.4	470.6
CFSR-8341	112.4	72.1	267.0
TRMM	42.1	13.9	62.7
MERRA-2	108.8	46.2	174.8

To sum it all up, MERRA-2 and CFSR-8341 datasets showed good performance in capturing the rainfall pattern of rain gauges data in the 2000–2012 period.

A similar comparison was carried out for the maximum and minimum temperature values for CFSR, MERRA-2 and the measured datasets at the individual gauge points for both monthly and daily scales in Figures S-2, S-3, S-5 and S-6, respectively. It can be noted that for maximum temperature, station CFSR-8341 measured the highest temperatures compared to all the other stations followed by the station CFSR-8344. MERRA-2 values were the closest to the observed temperature values. The average monthly maximum values were within the range of 38–22 °C as shown in Figure S-2.

For the minimum temperature, the monthly average values were within the range of 18–22 °C as shown in Figure S-3. Station CFSR-8341 measured the highest minimum temperatures whereas MERRA-2 measured the lowest.

Model calibration, sensitivity and uncertainty analysis

The sufficiency of the modelling performance is highly dependent on the calibrated parameters. Table 3 shows the sensitivity ranks in decreasing order and final daily and monthly calibration values of the top eight selected model parameters considered. During the first simulation, the initial SWAT values were obtained from the observation file of the SWAT-CUP and were adjusted using the experience and literature. The best parameter range obtained in the first iteration was then substituted and used in the next iteration for each of the five iterations performed. After running the model for five iterations each with 500 simulations, ranges of the sensitive parameters were obtained. These ranges cut across all the datasets. The established SWAT model provides parameter adjustment to improve the accuracy of hydrology prediction through parameter sensitivity analysis and model calibration.

The NSE, PBIAS and R^2 were used to compare model performance and the results for the simulations are presented in Table 4.

The results show that NSE, R^2 and PBIAS fall in the acceptable range according to Moriasi *et al.* (2007); however, it can clearly be seen that the R^2 of the observed data are higher than that of the reanalysis datasets for the calibration and validation periods. In addition, it is worth noting that TRMM underperformed compared to MERRA-2 and CFSR but all provided sufficient accuracy to build the hydrological model. The poor performance of TRMM (R^2 of 0.54) dataset can be

Table 3 | The monthly calibrated SWAT parameter ranges and sensitivity analysis

Parameter	Min. value	Max. value	Rank
CN2.mgt	0.11	0.192	1
ALPHA_BF.gw	0.073	0.19	2
GWQMN.gw	0.00	101.17	3
CH_K2.rte	152.65	218.14	4
SOL_AWC(..).sol	−0.13	−0.10	6
SOL_Z().sol	−0.38	−0.10	7
ESCO.hru	0.59	0.87	8
REVAPMN.gw	389.41	454.87	9

Table 4 | Comparison of the SWAT daily simulation results for Manafwa catchment before bias correction

	Evaluation criteria	Gauged data	CFSR	TRMM	MERRA-2
Calibration period (2003–2007)	NSE	0.67	0.53	0.44	0.57
	R^2	0.77	0.61	0.51	0.60
	PBIAS (%)	15.19	9.40	−6.76	11.63
	p-factor	0.71	0.64	0.58	0.60
	r-factor	1.09	3.13	3.69	2.94
Validation Period (2009–2012)	NSE	0.53	0.50	0.46	0.54
	R^2	0.68	0.55	0.43	0.52
	PBIAS (%)	23.4	−10.6	16.9	17.0
	p-factor	0.78	0.55	0.61	0.65
	r-factor	1.72	2.5	3.54	1.99

attributed to the high variability in annual rainfall produced by this dataset compared to the other two datasets. Hence, there is a need to cautiously incorporate bias corrections in the reanalysis datasets before its application.

From Table 4, the *p*-factor values obtained indicate that >50% of the observed streamflow values for all the different SWAT models were bracketed within the 95PPU band at the daily time step for all the datasets although CFSR outperformed MERRA-2 and TRMM. At the daily time step, the gauged dataset outperformed the other two datasets with >70% of observed streamflow bracketed within the 95PPU band. Nevertheless, *r*-factor values obtained for CFSR and TRMM as shown in Table 3 indicate that the uncertainty band for these datasets was slightly wider compared to that of MERRA-2. This suggests that streamflow simulated using TRMM dataset had the lowest level of uncertainty followed by CFSR while MERRA-2 produced the highest uncertainty. Therefore, to a certain degree, the reanalyses of weather datasets representing the weather occurring over the watershed are able to drive the hydrological model.

The graphs in Figures 4 and 5 show the relationship between the measured discharge data and the simulated flow for a calibration period of 2003–2007 and a validation period of 2010–2013. The days are written in the format mm/dd/yy, e.g. 01/01/2000.

From the tendency of daily streamflow comparison, the simulated results from the CFSR and MERRA-2-driven SWAT models fit more closely to the observed data with R^2 >0.6 compared to that of TRMM with R^2 of 0.51. The simulated baseflow slightly matched the observed simulations, and there was a similar discrepancy between the peak streamflow as those of the observed simulations. The daily discharge of 90.374 m^3/s, with the largest water level in the available data for our chosen period, was recorded in May 2010. From the calibration results of the MERRA-2 and CFSR data, generally, the flows were overestimated though the peak flow was highly overestimated by MERRA-2 data. The TRMM weather flow simulations

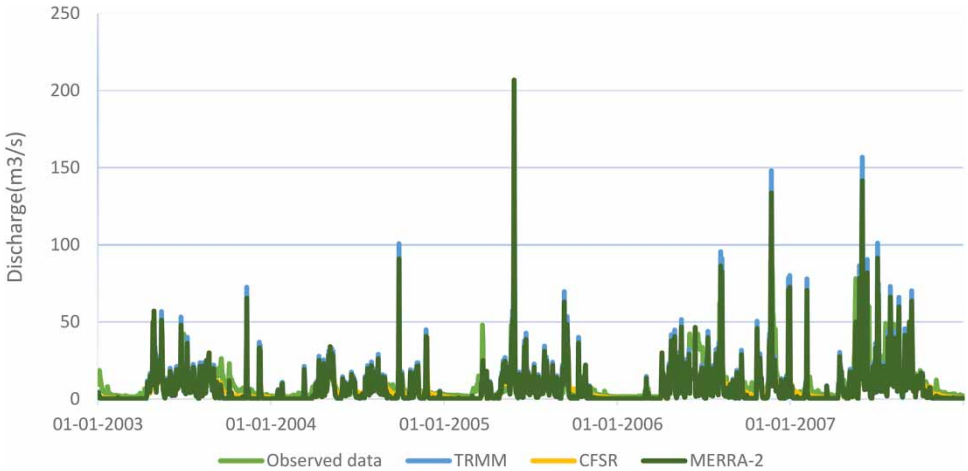

Figure 4 | Simulated flow of the SWAT model by TRMM, CFSR and MERRA-2 data plotted with the observed flow for Manafwa Catchment (calibration period: 2003–2007).

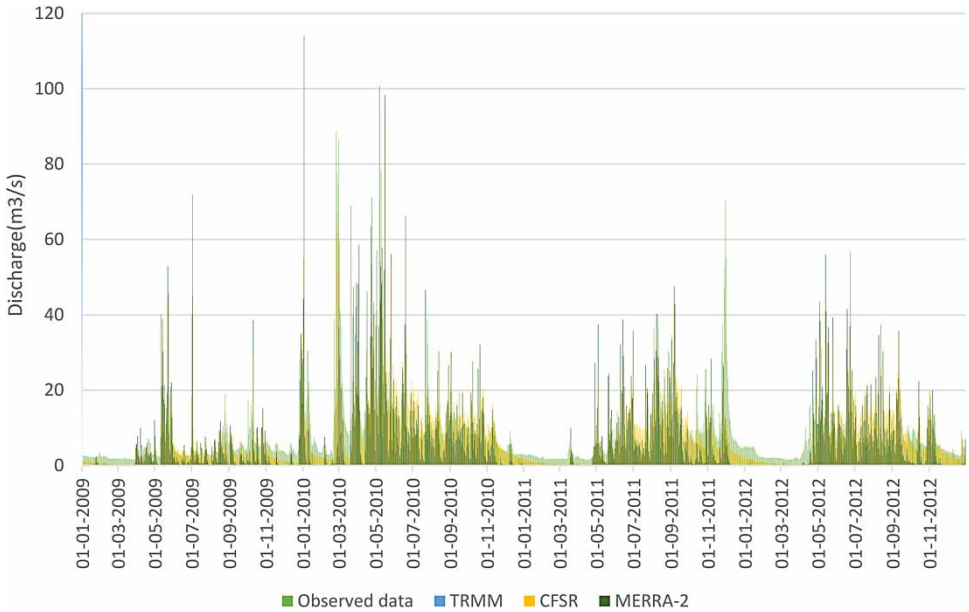

Figure 5 | Simulated flow of the SWAT model TRMM, CFSR and MERRA-2 data plotted with observed flow for Manafwa Catchment (validation period: 2009–2012).

were also generally overestimated and the pattern was similar to that of MERRA-2 just a bit higher. The peak flow was severely overestimated in all the reanalysis and satellite data as compared to that of the observed weather data flow simulation. This underestimation of peak flows may be due to specific characteristics of Manafwa Catchment or biases in observed precipitation, temperature and discharge. The underestimated precipitation in the upper reaches with higher elevation led to difficulty to reach the observed peak streamflow.

In Hurkmans *et al.* (2010), the bias-corrected precipitation and temperature were used to calibrate the model and do a climate impact study. However, the model needed other meteorological forcing data as well, such as wind speed, solar radiation, vapour pressure and specific humidity. Correcting precipitation and temperature only would violate the energy balance present in ERA15/REMO. Unfortunately, there were no observations for the other forcing variables available at the temporal and spatial resolution used in that study and they were, therefore, left uncorrected. The same was done for our study. This is okay, because for calibration purposes, it is expected that precipitation and temperature have the largest influence on the performance of the hydrological model.

All in all, the stream flows generated with the observed weather data are better than those simulated with CFSR, TRMM and MERRA-2 weather data, and the results simulated with all these weather data are reasonable.

Generating bias corrections

For rainfall data

The LOCI spreadsheet was used as earlier on mentioned. New time series were inserted using the option 'Paste Special as value'. The time series used was 2000–2013 (total of 14 years). The observed and raw data were entered in as well as the number of years. The spreadsheet was then run and the results below were generated in Tables 5 and 6.

From Table 6, it can clearly be seen that for the three reanalysis datasets, that the datasets are negatively biased, i.e. it is underestimated as the correction factors are generally positive. It is worth noting that the correction factors for TRMM are generally more than those of the other two datasets. The CFSR data from station 8344 presented the least correction factors which meant that its precipitation data are close to that of the rain gauges on the ground.

For temperature data

This was done for CFSR data for both the minimum and maximum temperatures. For MERRA-2 data, these corrections were also generated for the maximum and minimum average surface temperatures obtained at 2 m above the surface. As TRMM only measures the precipitation data, this was not done for it.

Table 5 | Calculation of the number of rainy days for each dataset

			Threshold			
Month	Obs	Raw	CFSR 8341	CFSR 8344	MERRA	TRMM
1	4.4	4.4	0.6	1.2	2.6	1.4
2	4.5	4.5	0.4	0.8	2.1	1.6
3	10.7	10.7	0.3	0.5	2.0	1.1
4	16.1	16.1	1.7	3.5	2.6	0.8
5	15.1	15.1	4.1	11.5	4.2	0.9
6	11.1	11.1	4.3	12.9	3.4	0.5
7	7.9	7.9	5.3	13.5	4.6	0.9
8	11.2	11.2	10.7	11.9	5.3	0.8
9	11.8	11.8	2.1	5.4	5.2	1.2
10	16.0	16.0	0.7	2.4	3.1	0.7
11	11.8	11.8	0.5	2.3	2.4	1.4
12	6.1	6.1	4.6	6.9	3.8	2.2

Table 6 | Calculation of correction factors for precipitation

		CFSR 8341		CFSR 8344		MERRA		TRMM	
Month	Obs	Raw	Factor	Raw	Factor	Raw	Factor	Raw	Factor
1	18.7	5.0	4.3	7.9	2.8	5.9	5.6	4.2	6.6
2	14.1	3.2	5.1	5.5	3.0	6.6	3.1	4.6	4.7
3	13.8	5.3	2.8	9.0	1.6	6.2	3.3	3.6	5.5
4	12.3	13.1	1.1	20.0	0.7	8.5	2.1	3.7	4.1
5	14.5	16.3	1.2	26.3	1.0	9.6	2.7	3.5	5.7
6	8.2	11.2	1.2	22.1	0.9	6.4	2.8	2.7	3.9
7	9.0	12.1	1.3	21.3	1.2	8.6	2.3	3.2	3.8
8	9.8	5.2	1.1	20.8	1.1	9.6	2.3	3.2	4.1
9	9.8	9.2	1.4	16.8	0.9	9.1	2.5	3.7	3.9
10	10.5	6.6	1.8	11.8	1.1	7.6	2.3	3.6	3.6
11	13.6	7.2	2.0	14.2	1.1	7.7	2.6	4.3	4.4
12	11.5	6.9	1.5	12.1	0.9	7.2	2.2	2.9	3.2

For CFSR temperature data

The correction factors of the mean monthly minimum and maximum temperatures were generated as per Equation (2). The bias correction factors in Tables 7–10 were obtained and this was done for the individual stations and for every single month in our study period (2000–2012).

For MERRA-2 data (T_{max})

The following bias corrections were generated using the observed maximum and minimum temperatures for the catchment, as well as the average maximum and minimum temperatures obtained from the MERRA-2 climate data. Using Equation (2), the following bias corrections were computed as shown in Tables 11 and 12.

Table 7 | Correction factors for CFSR 8341 T_{max}

Month	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	−5.46	−1.32	−2.47	−3.07	−4.22	−3.70	−7.10	−1.35	−4.34	−3.67	−2.24	−5.35	−3.74
2	−5.19	−5.61	−5.83	−6.11	−4.90	−7.84	−6.35	−2.47	−3.59	−3.84	−2.07	−7.04	−8.92
3	−6.67	−4.54	−2.91	−6.83	−5.12	−4.68	−3.70	−6.88	−2.63	−5.34	−0.66	−7.09	−8.66
4	−3.31	−2.51	−3.64	−2.71	0.12	−3.41	−1.16	−3.48	−3.27	−0.05	1.23	−6.33	−5.25
5	−3.51	−1.01	−0.97	0.08	−0.95	1.29	−1.09	−1.78	−2.59	0.75	3.12	−3.16	−0.59
6	−4.19	−0.60	−2.79	0.73	−0.30	0.35	−1.70	0.59	−0.21	−1.02	2.28	−2.97	0.46
7	−3.04	−1.27	−4.51	1.53	−0.90	−0.28	−0.64	0.68	−0.65	−0.35	1.20	−3.85	−1.38
8	−3.20	−0.79	−2.60	0.24	−0.89	−1.92	−1.93	0.38	−2.28	−2.79	0.45	−3.35	−2.51
9	−6.42	−3.14	−5.64	−3.22	−3.38	−2.25	−2.45	−0.76	−2.23	−3.30	0.74	−2.55	−3.44
10	−2.58	−1.42	−4.29	−5.76	−4.21	−4.68	−3.80	−4.95	−1.40	−3.87	−0.17	−4.70	−4.84
11	−3.87	−0.66	−3.36	−5.96	−3.35	−5.23	−1.49	−6.31	−2.34	−5.49	−3.31	−3.52	−4.76
12	−2.59	−4.02	−2.09	−3.75	−3.85	−8.12	0.82	−3.92	−5.50	−1.08	−2.58	−3.84	−4.72

Table 8 | Correction factors for CFSR 8341 T_{min}

Month	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	−0.75	0.75	−0.38	−0.72	−2.32	−0.02	−1.13	−0.73	−0.84	−1.55	−1.04	0.66	1.32
2	−2.03	−0.14	−0.82	−1.67	−1.22	−1.23	−1.41	−0.49	−0.29	−0.18	−2.02	0.71	1.61
3	1.26	−0.62	−0.25	−0.59	−0.66	−1.23	−0.41	−0.43	0.12	−0.64	−0.71	−0.32	1.40
4	−0.59	−0.15	−0.25	−0.26	−0.65	−0.61	−0.39	−0.09	−0.16	−0.34	−0.44	0.26	0.58
5	−1.06	−0.25	−0.03	−0.05	−0.30	−0.43	0.06	−0.41	−0.32	−0.13	−0.75	0.32	1.54
6	−0.42	−0.26	−1.04	−0.82	−1.08	−0.47	−0.76	−0.67	−0.12	−0.07	−0.72	−0.21	0.23
7	−0.29	−0.43	−1.44	−0.61	−0.71	−0.29	−0.83	−0.94	−0.30	−0.31	−0.53	0.18	0.06
8	0.55	−0.73	−1.21	−0.77	−1.76	−0.78	−1.64	−1.32	−1.26	−1.31	−0.80	−0.87	0.01
9	−2.29	−0.85	−1.87	−1.56	−0.90	−1.61	−1.51	−1.93	−1.22	−1.89	−0.98	−1.32	−0.27
10	−2.64	−0.64	−1.10	−1.24	−1.18	−0.40	−1.44	−0.88	−0.97	−1.17	−1.28	−0.69	0.09
11	−0.71	−0.31	0.13	−0.92	−1.27	−1.20	−0.44	−0.17	0.38	−0.84	−0.15	−0.38	0.37
12	−0.05	−0.47	−0.72	−0.57	−1.51	−0.74	−0.65	−0.06	0.15	−1.26	−0.02	0.24	0.97

From the computations above, it can be concluded that most of the CFSR temperature data were positively biased which means that generally, the temperature data were overestimated. For the MERRA-2 temperature data, we observe that most of the bias corrections are positive implying that the temperatures measured were generally underestimated.

The generated bias corrections were then subjected to the rainfall and temperature datasets at both daily and time scales and these were compared as shown in Figures S-7 up to Figures S-12 in the supplementary material. From Figure S-7, it can be seen that the bias-corrected monthly precipitation data for all the datasets are much closer to the observed rainfall. Although, it should be noted that CFSR-8344 rainfall data have the highest peaks of rainfall followed by CFSR-8341, the comparison on a daily scale in Figure S-10 showed that CSFR-8341 produced a significant high amount of rainfall at the start of 2010 of about 350 mm.

At the monthly scale, all the datasets are a perfect fit with the observed temperature values for both minimum and maximum temperatures as illustrated in Figures S-8 and S-9. This is so because we used the linear scale method for computing the bias corrections using the monthly mean data. On a daily scale, both the bias-corrected maximum and minimum temperatures as shown in Figures S-11 and S-12, respectively, all fell in the range as the initial temperature values of the observed values.

Table 9 | Correction factors for CFSR 8344 T_{max}

Month	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	−3.28	−1.51	−1.96	−1.68	−3.08	−2.66	−4.54	−1.25	−2.49	−2.47	−1.44	−5.64	−4.70
2	−3.32	−2.18	−2.81	−2.52	−2.06	−3.84	−2.95	−1.42	−2.30	−2.33	−1.73	−6.63	−5.78
3	−2.70	−1.82	−1.76	−3.61	−2.23	−2.72	−2.30	−3.13	−1.72	−2.49	−1.56	−6.08	−5.86
4	−2.17	−0.80	−2.04	−1.79	−0.78	−2.50	−0.95	−1.46	−1.12	0.31	−0.67	−6.85	−4.34
5	−2.58	−0.37	−0.90	0.07	−0.33	−0.44	−0.80	−0.73	−1.37	0.34	1.08	−4.35	2.37
6	−1.86	−1.43	−1.78	−0.57	−1.53	−1.13	−1.89	−1.15	−0.99	0.49	−0.78	−0.12	3.11
7	−2.33	−1.43	−1.72	−0.01	−1.10	−1.92	−0.60	−0.87	−1.66	−0.70	−0.98	−0.44	1.92
8	−1.90	−1.37	−1.18	−1.17	−0.93	−0.86	−2.16	−1.13	−1.25	0.51	−1.52	−1.05	−0.30
9	−1.72	−1.00	−2.09	0.10	−1.02	−0.61	−0.59	0.90	−1.14	0.72	−0.69	0.44	−0.26
10	−1.67	−0.63	−1.94	−2.52	−1.37	−1.41	−0.80	−0.79	−0.51	−1.02	−1.23	−1.78	−2.13
11	−1.53	−0.70	−2.47	−3.17	−2.20	−3.05	−2.02	−2.29	−1.82	−1.96	−2.33	0.01	−1.51
12	−1.69	−2.14	−2.82	−2.79	−3.18	−4.87	−1.04	−3.24	−3.15	−1.71	−2.28	−0.75	−1.89

Table 10 | Correction factors for CFSR 8344 T_{min}

Month	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	−1.37	1.07	−0.04	−0.55	−1.04	−0.13	−0.88	−0.59	−0.09	−1.44	−0.50	0.17	−0.23
2	−2.13	0.02	−0.59	−1.12	−0.46	−0.73	−1.15	0.02	0.20	0.35	−1.14	0.21	0.77
3	0.58	0.14	0.17	−0.13	−0.44	−0.20	0.28	−0.07	1.01	−0.49	0.20	0.20	1.12
4	0.78	1.20	0.84	0.65	0.81	0.41	0.81	0.53	1.00	1.21	0.48	0.99	1.58
5	0.12	1.34	1.27	1.37	1.29	0.83	1.15	0.71	1.39	1.34	0.44	1.88	2.96
6	0.94	1.21	0.37	0.32	0.43	0.79	0.66	1.23	1.18	1.50	0.24	0.93	0.99
7	0.72	1.16	−0.18	0.39	0.53	1.04	0.43	0.38	0.86	0.92	0.67	1.56	1.40
8	1.26	0.74	0.12	0.34	−0.33	0.53	−0.45	−0.45	0.08	−0.17	0.39	0.31	1.39
9	−1.12	0.54	−0.46	−0.02	0.32	−0.13	0.00	−0.59	−0.03	−0.68	0.09	0.19	1.15
10	−1.38	0.17	0.08	−0.51	−0.04	0.70	−0.42	0.11	0.33	0.11	0.01	0.68	0.75
11	0.58	0.66	0.80	−0.07	−0.18	−0.30	0.14	−0.04	0.39	0.14	0.55	0.60	1.29
12	0.43	−0.19	0.01	−0.12	−0.71	−0.53	0.08	0.29	0.68	−0.29	0.34	0.34	1.26

The bias-corrected data were used to make the SWAT models which were calibrated and validated on a daily scale in the SWAT-CUP and NSE, PBIAS and R^2 were obtained for each of the datasets as shown in Table 13. Figures 6 and 7 show daily simulated flows of the SWAT model after application of bias corrections to reanalysis datasets of TRMM, CFSR, MERRA-2 plotted with that of the gauged data for Manafwa Catchment.

Generally, it was observed that there was a considerable improvement in R^2 values after the application of the bias corrections for all datasets. For example, R^2 values for TRMM data improved from a range of 0.43–0.51 to 0.67–0.70 while that of CFSR and MERRA-2 improved from a range of 0.52–0.61 to 0.73–0.77. This shows that after the application of bias corrections, the corrected weather data have reasonable accuracy to represent the weather conditions occurring in the watershed. The PBIAS values obtained also showed that all the datasets both before and after the introduction of bias corrections were able to produce values that fell within the acceptable limits (PBIAS ± 25).

CONCLUSION

It is obvious that the different reanalysis datasets can be used to drive the hydrological models of various tropical catchments in the absence of the gauged data. This is based on their ability to drive hydrological models that are close to those of the

Table 11 | Correction factors for MERRA-2 T_{max}

Month	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	−0.66	2.17	2.35	0.62	1.09	1.21	−1.47	1.45	1.58	0.80	1.47	0.85	3.08
2	−0.88	−2.55	−0.66	−1.94	0.33	−1.59	−1.82	1.59	0.83	−0.41	1.37	−0.41	−0.55
3	−4.26	−0.31	−0.14	−2.24	−1.06	−1.00	−0.41	−1.08	0.51	−2.15	2.45	0.46	−2.37
4	−1.37	−0.11	−1.09	−0.18	2.12	−1.13	0.89	−1.15	−0.13	0.87	1.51	−0.54	2.44
5	2.78	1.74	1.31	1.96	1.67	2.89	1.75	1.19	0.65	1.88	3.14	2.33	4.40
6	2.36	3.94	1.49	3.66	2.86	1.99	1.73	3.72	2.59	0.01	3.21	4.08	5.88
7	4.19	3.54	0.28	3.54	2.46	2.70	2.02	3.83	3.21	2.06	3.41	2.70	6.05
8	4.68	3.79	2.72	4.05	3.13	2.08	2.27	4.11	2.67	1.50	3.50	4.47	5.33
9	1.48	2.84	0.95	2.51	2.39	1.58	1.61	4.92	3.69	1.18	3.83	4.50	5.35
10	3.85	3.59	1.67	1.71	1.35	0.68	0.50	2.61	4.23	2.15	3.24	3.95	3.91
11	1.40	3.26	0.86	0.72	1.49	1.02	2.51	0.66	3.77	0.03	1.70	4.81	4.15
12	0.72	1.32	2.21	1.82	1.16	−1.08	3.41	2.04	0.29	1.94	1.82	3.95	2.79

Table 12 | Correction factors for MERRA-2 T_{min}

Month	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	−0.28	0.47	0.64	0.59	−0.58	0.57	−0.97	0.43	0.86	−0.02	0.24	0.20	1.81
2	−0.82	0.65	1.22	−0.32	−0.11	0.39	−1.05	0.18	1.10	0.41	−0.67	0.69	2.47
3	0.16	0.59	0.27	−0.47	−0.15	−0.39	−0.19	0.45	0.50	−0.40	−0.04	0.28	1.50
4	−0.30	0.62	0.52	0.37	0.28	0.20	0.62	0.06	1.40	0.43	0.29	0.65	1.47
5	−0.55	−0.13	0.44	0.59	0.36	0.14	−0.12	−0.14	0.11	0.33	−0.38	0.17	1.68
6	0.25	0.51	0.10	0.22	0.13	−0.08	−0.32	0.40	0.17	−0.55	−0.20	−0.09	0.09
7	−0.32	0.72	−0.31	0.16	0.18	0.34	−0.56	0.05	0.70	−0.12	0.06	0.12	0.54
8	0.43	0.29	−0.69	0.08	−0.64	−0.01	−0.66	−0.19	0.02	−0.83	−0.37	−0.20	0.97
9	−0.82	−0.10	−0.52	−0.27	−0.53	0.33	−1.00	−0.26	0.17	−1.41	−0.72	−0.24	0.72
10	−0.54	0.36	−0.08	0.09	0.29	0.33	−0.50	0.49	0.53	−0.27	−0.18	0.11	0.89
11	0.92	0.92	0.77	0.38	0.43	0.35	0.00	0.54	1.73	0.72	0.56	0.74	1.26
12	0.83	1.08	−0.13	1.36	0.35	0.10	0.11	0.99	2.11	−0.40	0.78	1.46	1.10

Table 13 | Comparison of the SWAT simulation results for Manafwa catchment using the observed data as well as the reanalysis datasets after correction

	Evaluation criteria	CFSR	TRMM	MERRA-2
Calibration period (2003–2007)	NSE	0.65	0.62	0.68
	R^2	0.75	0.70	0.73
	PBIAS (%)	19.41	−0.17	15.64
Validation period (2009–2012)	NSE	0.62	0.66	0.64
	R^2	0.77	0.67	0.74
	PBIAS (%)	18.85	13.5	12.20

observed datasets. However, it is also true that hydrological modelling using these datasets is exposed to significant amount of model uncertainty due to the differences between them and the true observed data. Original reanalysis datasets are very biased, and this precludes their direct use in hydrological modelling of various catchments.

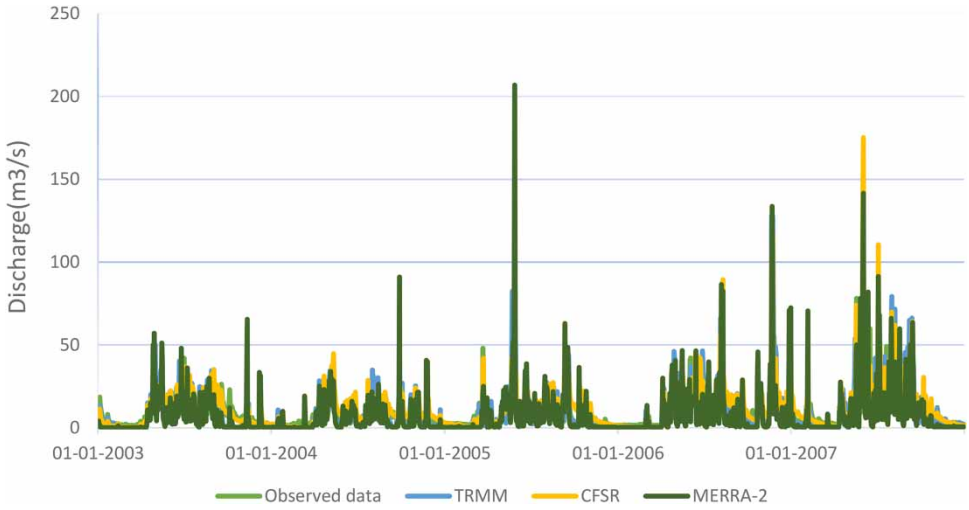

Figure 6 | Simulated flow of the SWAT model after application of bias corrections to reanalysis datasets of, TRMM, CFSR, MERRA-2 and gauged data plotted with that of the observed flow for Manafwa Catchment (calibration period: 2003–2007).

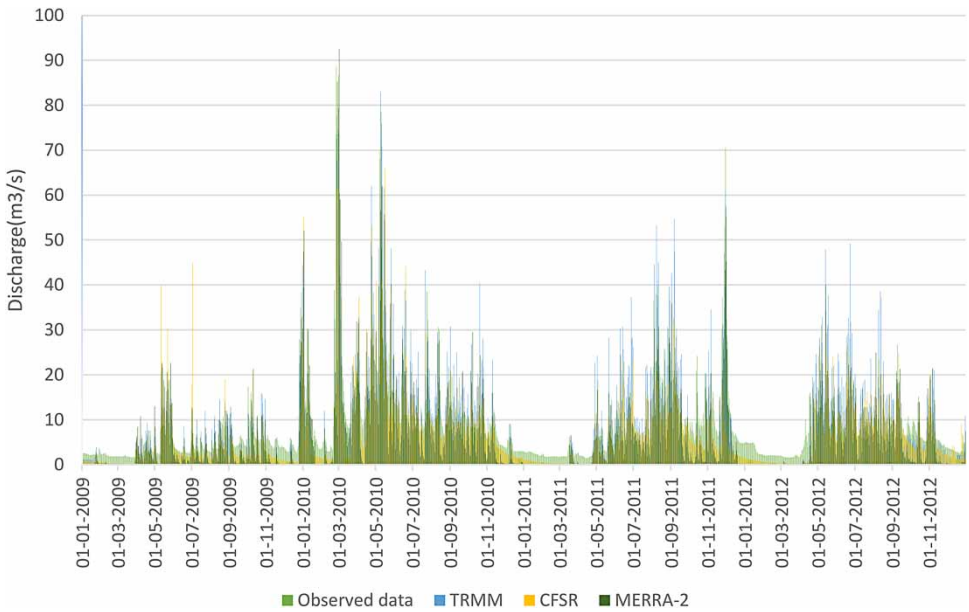

Figure 7 | Simulated flow of the SWAT model after application of bias corrections to reanalysis datasets of, TRMM, CFSR, MERRA-2 and gauged data plotted with that of the observed flow for Manafwa Catchment (validation period: 2010–2012).

The aim of this study was to evaluate the performance of reanalysis datasets, namely CFSR, MERRA-2 and TRMM in simulating river discharge in data-scarce tropical catchments: a case study of Manafwa, Uganda. The model was set up in the Arc SWAT after which it was calibrated and validated with SUFI-2 on a monthly resolution to see how best it can simulate river discharge. This was done for the observed data, uncorrected and bias-corrected reanalysis datasets. Bias correction was intended to reduce the possible errors with the reanalysis data. Model performance was evaluated and compared using statistical metrics like NSE, PBIAS and R^2.

The results of our study showed that the bias-corrected datasets significantly outperformed the non-corrected datasets in simulating streamflow in the study area. This highlights the importance of bias correcting the global reanalysis datasets before using them for hydrological modelling. It is concluded that from our study, the TRMM rainfall dataset matched best with the gauged data followed by CFSR and MERRA-2 reanalysis datasets.

In this study, the LOCI method was used to generate the monthly bias corrections for precipitation whereas the LS method was used to generate the bias corrections for temperature. The LOCI method was preferred for precipitation because it takes into account the number of wet days. Bias correction practices with the usage of reanalysis datasets are highly recommended to minimize the model uncertainties in hydrological models.

Results of model evaluation indices showed that all the three datasets performed better after the application of bias corrections. This is not surprising given that the corrected sets had the best rainfall and temperature input because of the reduced variability in these climate estimates. This demonstrates the importance of post-processing or bias-correcting the global reanalysis datasets before using them for hydrological modelling.

However, this study had its limitations as well. One of the datasets, i.e. TRMM satellite data do not include temperature measurements; therefore, ideally, its usage is limited to only precipitation data. It should be noted that the SWAT requires the input of both the maximum and minimum temperatures. Therefore, for the application of TRMM satellite data, an assumption was the other weather input requirements used were those of the observed gauged data. For future research, there is still a need to exhaustively investigate the potential of using bias corrections on reanalysis datasets for this catchment for other periods other than those used in this study. A comparison of models would greatly confirm how dependable the corrections can be when it comes to hydrological modelling. There is also another gap to try the datasets and their computed bias corrections using the LS and LOCI bias correction methods on other catchments to see how best the methods are.

To sum it all up, CFSR, MERRA-2 and TRMM data (in that order) can adequately be used in simulating river discharge of tropical catchments. As seen in the graphs in Table 11, the bias correction reduces the uncertainty in precipitation and temperature estimates used to drive the hydrological model which has a direct positive impact in reducing the overall uncertainty in the simulated streamflow. There was an increased correlation between the performance indicator values of the observed flows, as well as those of the corrected dataset. We can conclude that, in the absence of gauged hydro-meteorological data, bias-corrected CFSR, MERRA-2 and TRMM datasets could be used for hydrological modelling in data-scarce areas such as the Manafwa Catchment and other remote locations with poor data availability.

ACKNOWLEDGEMENTS

This work reported here was undertaken as part of the *Building Capacity in Water Engineering for Addressing Sustainable Development Goals in East Africa* (*CAWESDEA*) project which is part of the IDRC-funded programme on Strengthening Engineering Ecosystems in sub-Saharan Africa. CAWESDEA Project is led by Global Water Partnership Tanzania in collaboration with Makerere University (Uganda), Moi University (Kenya) and University of Dar es Salaam (Tanzania). We acknowledge the support from the Directorate of Water Resources Management; Ministry of Water and Environment for hosting the research herein.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Abbaspour, K. 2015 *SWAT-CUP 2012. SWAT Calibration and Uncertainty Programs*. EAWAG, Dübendorf, Switzerland.

Arnold, J., Kiniry, J., Srinivasan, R., Williams, J., Haney, E. & Neitsch, S. 2012 *Soil & Water Assessment Tool: Input/Output Documentation Version 2012; No. 439*. TexasWater Resources Institute: College Station, Texas, USA.

Bahati, H. K., Ogenrwoth, A. & Sempewo, J. I. 2021 Quantifying the potential impacts of land-use and climate change on hydropower reliability of Muzizi hydropower plant, Uganda. *Journal of Water and Climate Change*. https://doi.org/10.2166/wcc.2021.273

Byakatonda, J., Openy, G., Sempewo, J. I. & Mucunguzi, D. B. 2021 Over century evidence of historical and recent dryness/wetness in sub-humid areas: a Uganda, East African case. *Meteorological Applications* **28** (5), e2028.

Daniel, R., Todd, M., Charlotte, M., Arthur, T. & Zachary, M. E. 2014 *Using the Climate Forecast System Reanalysis as Weather Input Data for Watershed Models*. https://doi.org/10-1002/hyp.10073. Wiley Online Library.

Demory, M.-E., Eltahir, E. A. & Siam, M. S. 2013 Hydrological cycles over the Congo and Upper Blue Nile Basins: evaluation of general circulation model simulations and reanalysis products. *Journal of Climate* **26** (22), 8881–8894. © 2013 American Meteorological Society.

Fang, G., Yang, J., Chen, Y. N. & Zammit, C. 2015 Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China. *Hydrology and Earth System Sciences* **19** (6), 2547–2559.

Fowler, H., Ekström, M., Blenkinsop, S. & Smith, A. 2007 Estimating change in extreme European precipitation. *Journal of Geophysical Research, D: Atmospheres* **2** (112), D18104.

Ganiyu, T. O., Ibukunoluwa, O. F. & Olubunmi, A. D. 2017 Modeling runoff with satellite-based rainfall estimates in the Niger basin. *Cogent Food & Agriculture* **3** (1363340), 1–12.

Gassman, P., Sadegh, I. & Srinivasan, R. 2014 Applications of the SWAT model special section: overview and insights. *J. Environ. Qual.* **43**, 1–8.

Gelaro, R., Max, J., Will, M. & Ricardo, T. 2017 The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate* **30** (14), 5419–5454.

Gorgoglione, A., Gioia, A., Iacobellis, V., Piccinni, A. & Ranieri, E. 2016 A rationale for pollutograph evaluation in ungauged areas, using daily rainfall patterns: case studies of the Apulian Region in Southern Italy. *Applied and Environmental Soil Science*. Articlee ID: 9327614. http://dx.doi.org/10.1155/2016/9327614.

Haerter, J., Hagemann, S., Moseley, C. & Piani, C. 2011 Climate model bias correction and the role of timescales. *Hydrological Earth System Science* **15**, 1065–1079.

Hurkmans, R., Terink, W., Uijlenhoet, R., Torfs, P., Jacob, D. & Troch, P. 2010 Changes in streamflow dynamics in the Rhine basin under three high-resolution regional climate scenarios. *Journal of Cimate* **23**, 679–699. doi:10-1175/2009JCLI3066.1.

Jingrui, W., Ganming, L. & Chen, Z. 2020 Evaluating precipitation products for hydrologic modeling over a large river basin in the Midwestern USA. *Hydrological Sciences Journal* **65** (7), 1221–1238.

Kigobe, M., McIntyre, N., Wheater, H. & Chandler, R. 2011 Multi-site stochastic modelling of daily rainfall in Uganda. *Hydrological Sciences Journal* 17–33. https://doi.org/10.1080/02626667.2020.536548.

Kisembe, J., Favre, A., Dosio, A., Christopher, L., Sabiiti, G. & Alex, N. 2018 Evaluation of rainfall simulations over Uganda in CORDEX regional climate models. *Theoretical and Applied Climatology* **1117–1134** (2019). https://doi.org/10.1007/s00704-018-2643-x.

Liu, Y., Gupta, H., Springer, E. & Wagener, T. 2008 Linking science with environmental decision making: experiences from an integrated modeling approach to supporting sustainable water resources management. *Environmental Modelling and Software* **23**, 846–858.

Liu, Z., Ostrenga, D., Teng, W. & Kempler, S. 2012 Tropical rainfall measuring mission (TRMM) precipitation data and services for research and applications. *Bulletin of the American Meteorological Society* **93**, 1317–1325. https://doi.org/10.1175/BAMS-D-11-00152.1.

Lorup, J., Refsgaard, J. & Mazvimavi, D. 1998 Assessing the effect of land use change on catchment runoff by combined use of statistical tests and hydrological modelling: case study from Zimbabwe. *Journal of Hydrology* **205** (3–4), 147–163. http://dx.doi.org/10.1016/S0168-1176(97)00311-9.

Maikel, M., Maathuis, B., Hein-Griggs, D. & Alvarado-Gamboa, L.-F. 2020 Performance evaluation of bias correction methods for climate change monthly precipitation projections over Costa Rica. *Journal of Water* **8–12**, 482. doi:10.3390/w12020482. Available from: www.mdpi.com/journal/water.

Mango, L., Melesse, A., McClain, M., Gann, D. & Setegn, S. 2011 Land use and climate change impacts on the hydrology of the upper Mara River Basin, Kenya: results of a modeling study to support better resource management. *Hydrology and Earth Systems Sciences* **15**, 2245–2258.

Meng, J., Li, L., Hao, Z., Wang, J. & Shao, Q. 2014 Suitability of TRMM satellite rainfall in driving a distributed hydrological model in the source region of Yellow River. *Journal of Hydrology* **509**, 320–332.

Monteiro, J., Strauch, M., Srinivasan, R., Abbaspour, K. & Gücker, B. 2016 Accuracy of grid precipitation data for Brazil: application in river discharge modelling of the Tocantins catchment. *Hydrological Processes* **30** (9), 1419–1430.

Moriasi, D., Arnold, G. J., Van Liew, M. W., Bingner, R. L., Harmel, D. R. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy In watershed simulations. *American Society of Agricultural and Biological Engineers* **50** (3), 885–900. ISSN 0001−2351.

Morse, A., Caminade, C., Tompkins, A. & McIntyre, K. 2013 *The QWeCI Project (Quantifying Weather and Climate Impacts on Health in Developing Countries); Final Report*. University of Liverpool, UK.

Mutenyo, I., Pouyan, N. A., Sean, A. & Subhasis, G. 2013 Evaluation of SWAT performance on a mountainous watershed in Tropical Africa. *Hydrology Current Research*, 1–7.

Najmaddin, P. M., Whelan, M. & Balzter, H. 2017 Application of satellite-based precipitation estimates to rainfall-runoff modelling in a data-scarce semi-arid catchment. *Climate* **5** (2), 32.

Naschen, K., Diekkruger, B., Leemhuis, C., Steinbach, S., Seregina, S., Thonfeld, F. & Can der Linden, R. 2018 Hydrological modeling in data-scarce catchments: the Kilombero floodplain in Tanzania. *Water Journal*. doi:10.3390/w10050599. www.mdpi.com/journal/water.

NCAR 2020 *Tropical Rainfall Measuring Mission*. The National Center for Atmospheric Research; Climate Data Guide: Available from: https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission

NCEI 2020 *Climate Forecasting System (CRFS)*. National Centers for Environmental Information (National Centers and Atmospheric Administration-NCDC). Available from: https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/climate-forecast-system-version2-cfsv2

Neitsch, S., Arnold, J., Kiniry, J. & Williams, J. 2011 *Soil and Water Assessment Tool Theoretical Documentation Version 2009*. Texas Water Resources Institute, Temple, TX, USA.

Nkiaka, E., Nawaz, N. & Jon, C. 2017 Evaluating global reanalysis datasets as input for hydrological modelling in the Sudano-Sahel Region. *Journal of Hydrology*. https://doi.org/10.3390/hydrology4010013.

Ouatiki, H., Abdelghani, B., Yves, T., Lionel, J., Tarik, B., Lahoucine, H. & Abdelghani, C. 2017 Evaluation of TRMM 3B42 V7 Rainfall Product over the Oum Er Rbia Watershed in Morocco. *Climate* **5** (1), 1.

Peng, Z., Lu, G., Jianhui, W., Miaomiao, M., Haijun, D., Jianyun, G. & Xingwei, C. 2020 Evaluation of ERA-interim air temperature data over the Qilian Mountains of China. *Hindawi Journal*, 11. Article ID 7353482, https://doi.org/10.1155/2020/7353482.

Pijush, S. 2015 *Handbook of Research on Advanced Computational Techniques for Simulation-Based Engineering*. VIT University, India. Centre for Disaster Mitigation and Management.

Rossi, M., Kirschbaum, D., Valigi, D., Mondini, C. & Fausto, G. 2017 Comparison of satellite rainfall estimates and rain gauge measurements in Italy, and impact on landslide modeling. *Climate* **5** (4), 90.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y-T., Chuang, H-Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M. & Becker, E. 2014 The NCEP climate forecast system version 2. *International Journal of Climatology* **27**, 2185–2208.

Schmidli, J., Frei, C. & Vidale, P. 2006 Downscaling from gcm precipitation: a benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology* **26**, 679–689.

Shrestha, M. 2015 *Data Analysis Relied on Local Intensity Scaling (LOCI) Bias Correction (V.3.0)*. Microsoft Excel file.

Tarana, A. S. & Slobodan, P. S. 2010 National Centers for Environmental Prediction – National Center for Atmospheric Research (NCEP–NCAR) reanalyses data for hydrologic modelling on a basin scale. *Canadian Journal of Civil Engineering* **37** (4), 611–623. https://doi.org/10.1139/L10-005.

Teutschbein, C. & Seibert, J. 2013 Bias correction of regional climate model simulation for hydrological climate-change impact studies: review and evaluation of different methods. *Journal of Hydrology* **456–457**, 12–29.

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

First received 2 March 2021; accepted in revised form 7 November 2021. Available online 23 November 2021