Reanalysis dataset-based hydrologic predictions for ungauged basins

Anghesom Ghebrehiwot[*], and *Dmitry* Kozlov

National Research Moscow State University of Civil Engineering, Moscow, Russia

Abstract. In the present study, Soil and Water Assessment Tool (SWAT) is employed to simulate streamflows from watershed with a semi-arid climate, using Climate Forecast System Reanalysis (CFSR) as forcing data input. To this end, two streamflow simulation scenarios, with and without readjustment of the reanalysis datasets, were investigated depending on available ground information. The findings indicate that the performance of the model is slightly improved when the former scenario, with readjustment of precipitation, is considered. Despite improvement in the overall model prediction, uncertainties during calibration and validation partially remained far less than the permissible limits. The reason seems to be associated with the mismatch between in-situ data and reanalysis datasets with respect to time and space. Irrespective of the sources of prediction uncertainties, the use of readjusted reanalysis datasets are deemed to be the best option for streamflow simulations in poorly gauged or ungauged watersheds. However, to underpin the findings with supportive and sound evidence, further investigation on the reanalysis datasets for hydrological predictions from similar regions with sufficient and reliable ground information becomes imminent. The study also underscores the need for undertaking pre-emptive measures to reverse the quantitative decline of hydrometric networks and existing management practices in the region.

1 Introduction

Rainfall-runoff models are extensively applied as predictive tools for generating hydrological responses during water resources development and management studies. Despite the recent extraordinary developments in rainfall-runoff predictions, there are still water management problems that existing models are incapable of addressing with sufficient confidence. The reasons are related to limitations in fully representing complex hydrologic systems and inadequacy of the available measurement techniques. Conspicuously, the performance of a rainfall-runoff model is intimately linked to data availability [1–3]; performance is lower and uncertain in data-scarce regions and vice versa in gauged regions. In addition to the fact that drainage basins in many parts of the world are ungauged, there are two issues with regard to gauged watersheds that are worth noting; limitations of high-quality hydrometric data and decline of hydrometric networks [3–7].

---

[*]Corresponding author: bahghi2012@gmail.com

The decline in hydrometric stations means that at a time when global warming may be exacerbating weather extremes and water shortages, scientists are less able to monitor water supplies, predict droughts, and forecast floods. The causes for the global decline of ground-based hydrometric networks are varied but include insufficient funding, inadequate institutional frameworks, a lack of appreciation of the worth of long term hydrological data, and the turmoil caused by wars and other disasters. Studies carried out in the East African region [8] emphasized that the political and socio-economic history of the region for the last several decades have not been conducive to the management of hydrometeorological records. Furthermore, these problems are also compounded by the impacts of human-made changes to the land surface and climate, occurring at the local, regional and global scales. The regions that suffer the greatest human impacts are usually those where hydrometric networks are least established. This is true of regions in developing countries where lack of hydrometric data coupled with climate and land use changes has led to a depletion of water resources and environmental degradation. Numerous investigations on the potentiality of satellite-based measurements for hydrological and environmental studies have been conducted [9–11] to circumvent these problems. A bulk of the available evidence points to the fact that a wide range of remote sensing data-based models' effectiveness and accuracy.

Recent studies on applications of global climate reanalysis datasets for streamflow simulations [12–15] demonstrated that overall model performance depends on the source and resolution of input datasets and the location (climate) of the study area. Most of these studies utilized physically-based watershed models. Analyzing these literatures, the successes and failures of reanalysis datasets-based simulations seem to be balanced. For example, the CFSR and ERA-Interim datasets - ERA-interim stands for European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis - were used simulate daily and monthly flows, using the SWAT model of a watershed in the central Sudano-Sahelian region [16]. Model performance indicators showed that ERA-Interim produced better results than CFSR. In similar studies conducted in South Africa [12] and Eritrea [14] on streamflow simulations using SWAT and precipitation-runoff model (MIKE 11-NAM), respectively, a lower statistical agreement between CFSR precipitation and ground-based rainfall data as well as overall imbalances in water were reported. The influence of spatial resolution of mesoscale atmospheric analysis system for surface variables, namely SAFRAN (Système d'analyse fournissant des renseignements atmosphériques à la neige) with different grid sizes and CFSR was also investigated over the Garonne basin in France [13]. They found that the difference in the representation of the climate was more influential than its spatial resolution. The application of CSFR for hydrological modelling of five watersheds with different hydroclimate settings in Ethiopia and the USA was also assessed [10]. They revealed that the forcing variables from CFSR provided streamflow simulations as good as ground weather stations. Most of the above works underscored that a considerable deviation of forcing variables, which in turn had a remarkable effect on other derived variables, may lead to model outputs being incongruent with observed data. However, while conventional ground hydrometric data remain the most accurate and reliable data sources, the use of reanalysis datasets as an alternative model input in ungauged or poorly gauged regions have been suggested.

Physically-based watershed models describe the hydrological processes and their interaction as and where they occur in the catchment and therefore offer the prospect of remedying the inadequacies of the conventional rainfall-runoff models. Probably the most sophisticated and widely used physically-based model is the SWAT model. It is "a conceptual, continuous-time model developed to assist water resource managers in assessing water supplies and non-point source pollution on watersheds and large river basins" [17] and operates at a daily time step. The advantages of SWAT are manifold: models watershed with no monitoring data; uses readily available inputs greatly facilitated

by the development of GIS-based interfaces; computationally efficient; enables users to study long-term impacts; and freely available in the public domain. As such, the application of SWAT has expanded worldwide both at watershed and continental scales. It has been applied for evaluating and assessing current and potential future influences of anthropogenic, climate change and others on a wide range of water resources. For example, river discharges in the United States of America [18], the hydrology of the African continent [19] and the impact of climate change on freshwater availability in Africa [20] were researched using SWAT and global climate models. Nevertheless, the inadequacy of the model's performances had been described in some studies [21], especially when comparisons of simulated output were made with time series of measured daily flow and/or pollutant loss data.

SWAT being comprehensive and semi-distributed, a large number of input parameters, which complicates model parameterization and calibration, are required. Manual and automated procedures that use the shuffled complex evolution method and other common techniques have been developed. Recently, calibration and uncertainty procedures (SWAT-CUP) was developed [22] to support a decision-making framework that integrates a semi-automated approach. It allows users to conduct manual and automated calibration, sensitivity and uncertainty analyses. On the other hand, some users have endeavoured to minimize weaknesses in SWAT by creating component modifications [21], which support more accurate predictions of specific processes, or by interfacing SWAT with other models. Calibration and uncertainty analysis of distributed watershed models are affected by parameterization of watershed models, the definition of what is a "calibrated watershed model" and what are the limits of its use, the conditionality of a calibrated watershed model, calibration of highly managed watersheds where natural processes play a secondary role, and uncertainty and non-uniqueness problems [23]. In this regard, SWAT-CUP standalone program [24] was developed specifically to handle these problems, containing five different calibration procedures, validation, and sensitivity. Previous studies described that different optimization procedures generate a different solution at different locations in the parameter spaces with more or less the same flow rate results [23]. However, the SUFI-2 calibration procedure [25], among them, was reported to be quite efficient for time-consuming and large-scale models.

Significant spatio-temporal variations of climate and physical characteristics, intensive human-induced land use and climate changes of watersheds in Eritrea are quite evident. Besides, most of the watersheds are not only ungauged or poorly gauged, but also the available data in gauged watersheds are scanty and unreliable. Under such circumstances, design and planning of water resources studies have been comprehensibly challenging [26]. A recent survey on the overall management of the hydrometric data collection and processing shows that the current practices and efforts at national level do not seem to be promising, rather it is worsening. Despite the aforementioned shortcomings in the hydrometric data monitoring system, there are a lot of ongoing nationwide water resources related development activities [26]. Considering the unavailability of data on one hand and the ongoing development activities, on the other hand, it is high time to investigate the suitability of satellite-based data and catchment geomorphology for hydrologic predictions in the region. To this end, the authors recently evaluated the applicability of lumped and conceptual rainfall-runoff models in the Mereb-Gash river basin: continuous simulations [14], using forcing data input from CFSR; and single-event-based simulations [27] from catchment geomorphologic and physiographic characteristics. Besides, we also intercompared SWAT with MIKE 11-NAM, with the former producing better performance in simulating streamflows. However, the performance did not fully fulfil the acceptance criteria. So, this precedence directed us to hypothetically assume that the adjustment of reanalysis datasets about the available field data can produce better outputs. Cognizant of

the interdependence of the various climate variables, the adjustment considered in this study was made only to precipitation on account of two reasons: statistical evaluation and the possibility of generating (simulating) most of the other variables. The statistical evaluation of annual precipitation indicated that most of the satellite-based stations overestimated precipitation than ground observations. The 32 global stations within and near the Mereb-Gash basin represented 66%, 28%, and 6% of overestimation, underestimation, and no chance of precipitation, respectively. Moreover, the annual rainfalls were almost double in 31% of the satellite-based stations. Therefore, the objective of the study aimed at using SWAT to establish a hydrological model of Debarwa watershed at subbasin level as well as monthly time intervals and evaluate the degree of improvement in model performance upon readjusting CFSR-based precipitation during streamflow simulations.

2 Materials and Methods

Gash river basin originates. The watershed area is approximately 200 km2, with an altitudinal range from 1905 m to 2550 m above mean sea level. About 50% of the catchment has a slope of less than 10%. The soil type is dominated by Eutric Nitosols of clay soils (75%), followed by Humic Cambisols of clay-loam (25%). According to infiltration-based soil classification, both soil categories fall under the third hydrologic group (C); that is, the soils have slow infiltration and water transmission rates when thoroughly wetted and a layer that impedes downward movement of water or, has moderately fine to fine texture. The land use is predominantly grassland (54%) and agricultural (41%). The watershed lies in the moist highlands zone where temperature varies from 0°C to 32°C and average annual rainfall of 547 mm. The climate in the watershed can be characterized as moderate, with December-January being the coldest and March-April the hottest. Maximum rainfall occurs in the summer season, specifically in the months of July and August, with a monthly mean rainfall of 185 mm and 175 mm, respectively. The watershed consists of a global weather station and a flow measuring station.

The SWAT model for the study area was predominantly established using freely available data. Other than the topographic - digital elevation model (DEM), soil properties and land use data, SWAT requires climatic data at daily or sub-daily time steps. In our case, the daily weather data that include precipitation, maximum and minimum air temperature, relative humidity, wind speed and solar radiation were used. Besides, these data were complemented by additional sources from the Ministry of Land, Water and Environment, Department Water Resources, Eritrea.

In the SWAT program, a watershed is divided into multiple subbasins, which are then further subdivided into hydrologic response units (HRUs). HRUs are defined as "lands with similar spatial datasets, namely topography, land use, and soil types and all the components of the soil water balance could be determined on an HRU basis, with the assumption that similar HRUs would have similar hydrologic characteristics" [22]. Simulation of a hydrological watershed is done in the land phase, which controls the amount of water, sediment, nutrient, and pesticide loadings to the main channel in each subbasin, and in the routing phase, which is the movement of water, sediments, etc., through the streams of the subbasins to the outlets. The climate-driven hydrological cycle provides moisture and energy inputs [22], such as precipitation, air temperature, solar radiation, etc., that control the water balance. In this respect, global CFSR datasets from the National Centre for Atmospheric Research (USA) were utilized. The soil processes include lateral flow from the soil, return flow from shallow aquifers, and tile drainage, which transfers water to the river; shallow aquifer recharge, and capillary rise from the shallow aquifer into the root

zone; and finally deep aquifer recharge, which removes water from the system. As noted earlier, SWAT is designed to simulate the various components of the hydrologic cycle; for example, rainfall-runoff process uses the soil conservation service curve number (CN) method, the potential evapotranspiration is estimated using the Penman-Monteith equation, and the channel water routing is simulated by the variable storage routing, etc. Furthermore, QSWAT 2012 interface was used to set up and parameterize the model. On the basis of DEM and the stream network, a threshold drainage area of 3 km2 was chosen to discretize the watershed into 13 subbasins, which were further subdivided into 61 HRUs based on soil, land use, and slope. Each HRU is normally considered to be a uniform unit where water balance calculations are made.

Upon feeding the forcing data input, two different scenarios were followed: (i) using the CFSR data as they are (without any modification), and (ii) with partial modification. In the latter case, the precipitation for the selected global stations was recalculated by superimposing on the long-term isohyetal map of Eritrea (Figure 1) to balance the underestimation and/or overestimation of reanalysis values. As such, down-scaling (up-scaling) factors for corresponding 32 stations inside and nearby the Mereb-Gash basin were determined. But, the scope of the modelling work was limited only to the Debarwa watershed for reasons, as mentioned in the preceding section. In both scenarios, the total simulation period was set to run from 1994 to 2010, including three-year arbitrarily chosen warming up period. Roughly 2/3 of the data were used for calibration and the remaining for validation. The warm-up period that mitigates the initial conditions were excluded from the analysis. The simulated data obtained from the SWAT model and daily and monthly observed flow data were employed to perform the initial and final runs with SUFI-2. The results obtained from both approaches were analyzed and evaluated using different statistical indicators. Stations with nitrate data, point sources of pollution, and detailed operational information on all types of reservoirs within the watershed were missing; hence, outlets that could be directly influenced by the same structures and activities were ignored.
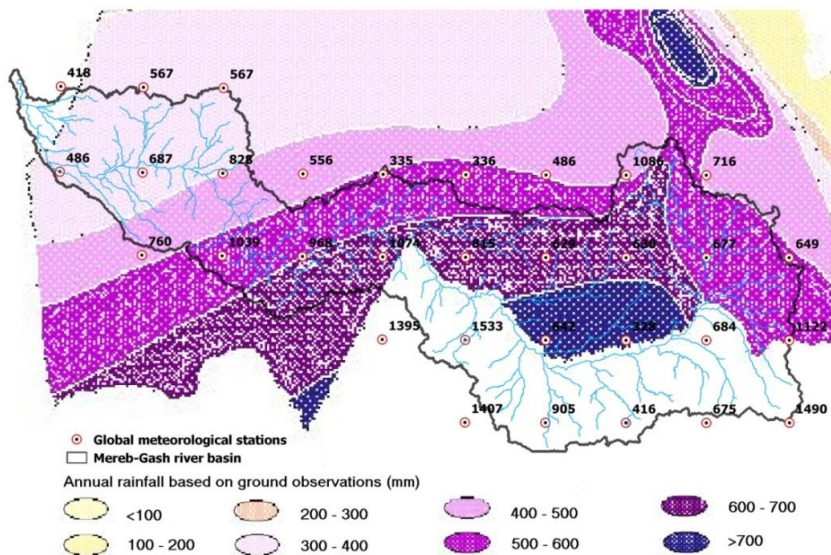

Fig. 1. CFSR-based annual precipitation superimposed over the long-term isohyetal map of Eritrea.

The three important concepts that require proper understanding are parameterization, objective function definition, and non-uniqueness (uncertainty) in calibration. In parameterization, we should answer the parameters to use and how to regionalize the parameters. The initial selection of parameters depends on the behaviour of the initial

model result before any calibration [28]. The SWAT-CUP program has the provision of eleven objective function options. While ten of them are independently evaluated, the remaining one is a multi-objective function, combining two or more objective functions. As has been clearly articulated in various literature [23–25], the outputs corresponding to each objective function are normally unique, leading to the conditionality of objective functions. As such, a multi-objective function has been suggested to overcome the problem of conditionality. But, due to limitation to access the multi-objective algorithm, our application was limited to individual functions.

Model uncertainty could be minimized if and only if we clearly identify the sources of uncertainty. Possible sources of uncertainty in hydrologic modelling can be categorized into five classes: (i) input data, including parameters, constants, and driving dataset; (ii) assumptions and simplifications; (iii) the science underlying the model; (iv) stochastic uncertainty also known as variability; and (v) code uncertainty, such as numerical approximations and undetected software errors. It would be unrealistic to expect a perfect model performance because of the aforesaid sources of errors and the many activities that happen in the watershed. Thus, our expectation from the model should be proportional to our knowledge of the model and model parameters, the quality of data and processes we put in the model, the quality of the effort we put into model building and calibration.

The successful application of hydrologic models is highly dependent on the calibration and sensitivity analysis of parameters. The calibration and validation processes are only efficiently engaged with observed data; streamflow data particularly plays a critical role in this procedure [24]. As such, available monthly streamflow data at the outlet of the Debarwa watershed were utilized for this purpose. As noted earlier, these procedures are mostly simplified with the help of the SUFI-2 algorithm. The algorithm provides plots of all uncertainties such as parameter, conceptual model output, input, etc., expressed as uniform distributions or ranges, and tries to detect most of the observed data within the 95% prediction uncertainty (95PPU) of the model in an iterative process. The 95PPU is determined at the 2.5% and 97.5% levels of the cumulative distribution of an output variable obtained through Latin hypercube sampling. P-factor and R-factor are suggested for the goodness of fit, as two bands are compared (the 95PPU for model simulation and the band representing observed data plus its error) [29]. The P-factor is the fraction of measured data plus its error enclosed by the 95PPU band and ranges between 0 and 1. While 1 represents 100% enclosure of the observed data within model prediction uncertainty, i.e., a perfect model simulation considering the uncertainty. P-factor>0.70 is considered to be adequate for streamflow. The P-factor value is influenced by the scale of the watershed and the sufficiency of the input and calibrating data. On the other hand, the R-factor is the ratio of the average width of the 95PPU band and the standard deviation of the observed variable and is represented by equation (1). An R-factor<1.50, again depending on the situation of the watershed, would be desirable for this index [29].

$$R - factor_j = \frac{\frac{1}{n_j}\sum_{t_i=1}^{n_j}\left(x_s^{t_i,97.5\%} - x_s^{t_i,2.5\%}\right)}{\sigma_{oj}} \tag{1}$$

where $x_s^{t_i,97.5\%}$ and $x_s^{t_i,2.5\%}$ are the upper and lower boundary of the 95PPU at time-step t and simulation i; n_j is the number of data points; and σ_{oj} is the standard deviation of the jth observed variable.

The *P*-factor and *R*-factor indices help us to justify the strength of the calibration and validation procedures. Once acceptable values of *R*-factor and *P*-factor are achieved in the final iteration, the parameter ranges can be assumed as the calibrated parameters. As

mentioned earlier, SUFI-2 allows the usage of different objective functions, out of which *NS* was chosen for calibration and validation procedures. Besides, the coefficient of determination (R^2), modified coefficient of determination (bR^2), per cent bias (*PBIAS*), and ratio of the root mean squared error to the standard deviation of measured data (RSR), whose corresponding equations are represented by eq. (2-6), were also additional criteria for statistical model evaluations.

$$R^2 = \frac{\left[\sum_i \left(Q_{m,i} - \overline{Q}_m\right)\left(Q_{s,i} - \overline{Q}_s\right)\right]^2}{\sum_i \left(Q_{m,i} - \overline{Q}_m\right)^2 \sum_i \left(Q_{s,i} - \overline{Q}_s\right)^2} \tag{2}$$

$$bR^2 = \begin{cases} |b| R^2 & if \quad |b| \le 1 \\ |b|^{-1} R^2 & if \quad |b| > 1 \end{cases} \tag{3}$$

$$NS = 1 - \frac{\sum_i \left(Q_m - Q_s\right)_i^2}{\sum_i \left(Q_{m,i} - \overline{Q}_m\right)^2} \tag{4}$$

$$PBIAS = 100 \times \frac{\sum_i^n \left(Q_m - Q_s\right)_i}{\sum_i^n Q_{m,i}} \tag{5}$$

$$RSR = \sqrt{\sum_{i=1}^n \left(Q_m - Q_s\right)_i^2} \; / \; \sqrt{\sum_{i=1}^n \left(Q_{m,i} - \overline{Q}_m\right)^2} \tag{6}$$

where Q is a variable (e.g., discharge); *m* and *s* stand for observed and simulated variables; *b* is slope of the regression line between the observed and simulated variables, and *i* is the ith observed or simulated data.

3 Results and Discussion

The choice of parameters that represent hydrological processes of a watershed from a long list of parameters (~800) is a quite difficult task. Parameterization procedure also depends on numerous factors such as topography, geology, soil, land use, size, location, the presence or absence of snow and glaciers, etc. Each factor is also further compounded by a multitude of parameters. However, selected parameters must reflect the basic hydrologic processes, such as surface overland flow, groundwater flow, soil and other watershed characteristics. As noted earlier, the study area is a poorly gauged watershed with limited observed data. The limited available data were considered in the calibration and sensitivity analysis procedures. During parameterization, SWAT-CUP provides two different methods of sensitivity analysis: one-at-a-time and global. In our case, the latter method was arbitrarily applied, where all selected parameters change at a time and use multi-regression computation. The global sensitivity uses the P-value and t-stats for analyzing the sensitivity of selected parameters to prioritize them; larger t-stat and lower P-value indicate higher parameter sensitivity and vice versa. Accordingly, parameters that are highly sensitive were selected and calibrated. Global sensitive parameters listed based on their level of sensitivity as analyzed by SWAT-CUP, are shown in Figure 2. After several trials, the baseflow alpha factor (ALPHA_BF), CN2, and effective hydraulic conductivity in main channel alluvium (CH_K2) are the topmost sensitive parameters among the 16 parameters considered during

calibration. Finally, the best simulated values along with recommended maximum and minimum thresholds were determined (Table 1).
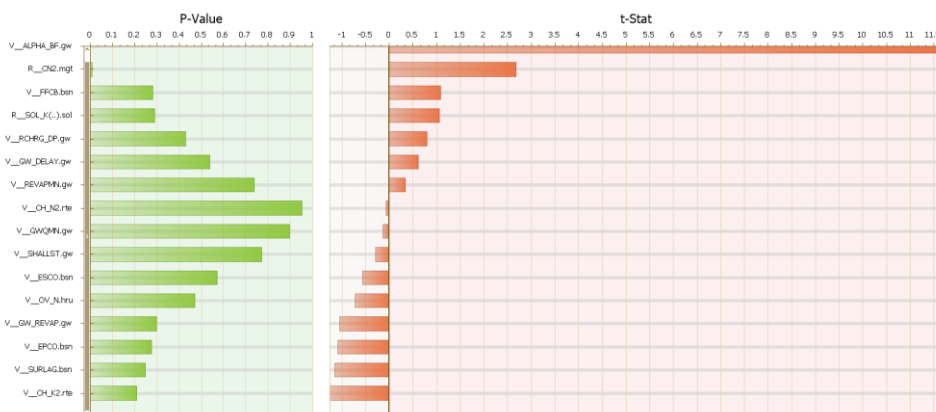

Fig. 2. Sensitivity analysis for monthly streamflows during calibration.

Table 1. Best simulated parameters' values during calibration procedure

Parameter	Description	Simulated values		
		Fitted	Min	Max
ALPHA_BF	Baseflow alpha factor (1/days)	0.58	0	1
CN2*	Soil conservation service (SCS) curve number for moisture condition II	0.08	-0.20	0.20
CH_K2	Effective hydraulic conductivity in main channel alluvium (mm/hour)	14.58	0	500
SURLAG	Surface runoff lag coefficient	1.25	0.05	24
EPCO	Plant uptake compensation factor	0.68	0	1
FFCB	Initial soil water storage expressed as a fraction of field capacity water content	0.29	0	1
SOL_K()*	Saturated hydraulic conductivity (mm/hour)	0.99	0	1
GW_REVAP	Groundwater "revap" coefficient	0.08	0.02	0.20
RCHRG_DP	Deep aquifer percolation fraction	0.27	0	1
OV_N.hru	Manning's "n" value for overland flow	0.21	0	1
GW_DELAY	Groundwater delay time (days)	466	0	500
ESCO.bsn	Soil evaporation compensation factor	0.16	0	1
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur (mm of H2O)	490	0	500
SHALLST	Initial depth of water in the shallow aquifer (mm of H2O)	1116	0	50000
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur (mm of H2O)	4130	0	5000
CH_N2	Manning's "n" value for the main channel	0.24	0	0.30

The change is relative, whereas the change in all other parameters is replacement with other value

As mentioned earlier, our expectation from a watershed model should be proportional to our knowledge of the model and model parameters, the quality of data and processes we put in the model, and the quality of the effort we put into model building and calibration.

Keeping this in mind and putting in maximum effort during model setup, the graphical visualization of CFSR rainfall and corresponding simulated daily streamflow from SWAT model prior to calibration in SWAT-CUP program and observed streamflow at the outlet of the watershed, were scrutinized. Based on these initial results and considering the dynamics and radical uncertainty of daily flows, calibration was limited to only monthly flows. Accordingly, the performance of the best parameter sets chosen during the sensitivity analysis was evaluated by two statistical evaluations: prediction uncertainty and performance evaluation. Uncertainty analysis refers to the propagation of all model input uncertainties to model outputs, which stem from the lack of knowledge of physical model inputs to model parameters and model structure. Identification of all acceptable model solutions in the face of all input uncertainties can provide us with model uncertainty in SWAT-CUP as 95PPU. Once the model is parameterized, and the ranges are assigned, the model is normally run some 300–1000 times [24]. After all, simulations are completed, the provision of the post-processing option in SWAT-CUP calculates the objective function and the 95PPU for all observed variables in the objective function.

The prediction uncertainty for the two scenarios, represented by the shaded regions for calibration (Figure 3 and 4) and validation (Figure 5 and 6) procedures, respectively, were achieved. As a result, P-factor values corresponding to reanalysis datasets without modification were estimated to be 0.34 and 0.43 for calibration and validation, respectively (Table 2). In other words, only 34% and 43% of the observed streamflows are bounded by the 95PPU during calibration (1997–2001 and 2007-2010) and validation periods (2002–2006), respectively. R-factor values are equal to 2.56 and 3.48 for calibration and validation periods, respectively. The calibrated and validated values of the P-factor and R-factor for the first scenario are outside of the recommended ranges (P-factor and R-factor shall be >0.70 and <1.50, respectively) [30]. Besides, the performance indicators for the calibration period (R2, bR2, and NS>0.70, and RSR<0.60) are in fairly acceptable ranges [24]. So, the statistical indices indicate that there is a good agreement between the observed and simulated streamflows. But, the corresponding model performance indicators for validation (R2 and bR2<0.40, NS<0.50, and RSR>0.70) are evaluated as unsatisfactory. PBIAS measures the average tendency of the simulated data to be larger or smaller than their observed counterparts [31]. Positive values represent model underestimation bias, and negative values indicate model overestimation bias. So, PBIAS based model performance when the first scenario was considered during calibration can be evaluated as unsatisfactory (PBIAS>±25), whereas that of validation is acceptable (PBIAS<±10). PBIAS show model overestimation by 42% and 9.8% during calibration and validation, respectively.

In the second scenario, some noticeable improvements to the overall performance of the SWAT model during calibration and validation procedures were observed. The downsizing factor was applied because the satellite-based precipitation was found to be overestimated for the area under consideration. R-factor shows an extraordinary improvement because it decimates from 2.56 and 3.48 to 0.84 and 0.97 for calibration and validation, respectively, while the ideal R-factor is 1. The implications of these values can be clearly visualized in the reduced range of the prediction spectrum (green shaded region). For instance, the reader can compare Figure 3 and Figure 4 and easily realize the changes in R-factor. Despite the occurrence of significant achievements in prediction uncertainty (R-factor) during the second scenario, the same could not be achieved in 95PPU (P-factor). While there was no change in calibration, it declines from 0.43 to 0.25 invalidation. This reinforces the broad consensus that the improvement of the R-factor is caused at the expense of the P-factor.

<div align="center">**Table 2.** Best simulated parameters' values during calibration procedure.</div>

Process	Uncertainty prediction		Objective function				
	P-factor	*R*-factor	R^2	bR^2	*NS*	*PBIAS*	*RSR*
Calibration	0.34	2.56	0.80	0.79	0.73	-42.0	0.52
	0.34	0.84	0.83	0.73	0.83	5.1	0.41
Validation	0.43	3.48	0.32	0.18	0.12	-9.8	0.94
	0.25	0.97	0.33	0.12	0.32	15.9	0.83

N.B: Italicized values represent model performance measures after readjustment of CFSR-based precipitation.
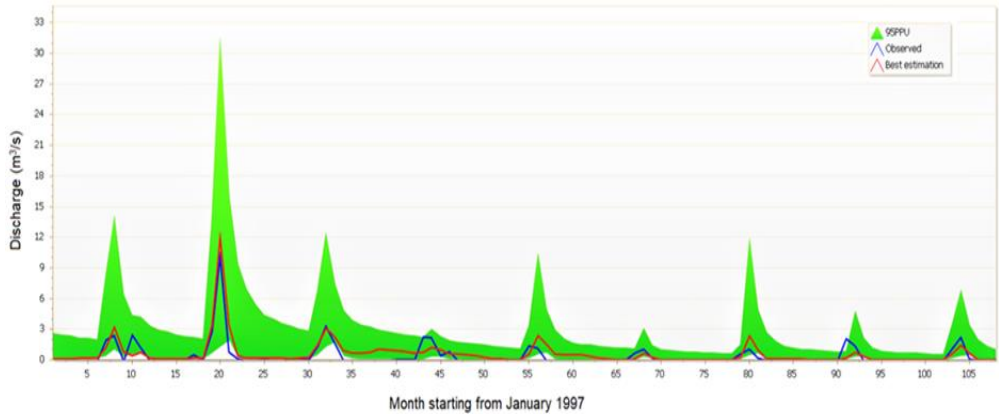

Fig. 1. Comparison of observed and simulated monthly streamflows during the calibration period (1997–2001 and 2007-2010) before readjustment of precipitation.
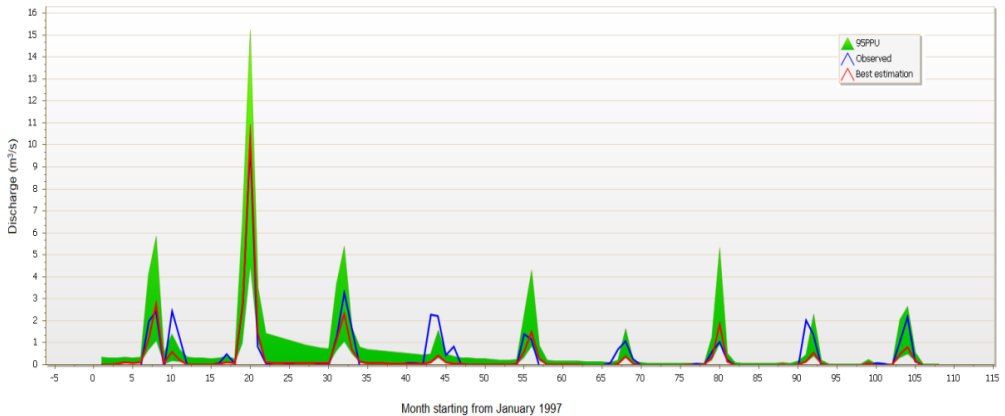

Fig. 2. Comparison of observed and simulated monthly streamflows during the calibration period (1997–2001) after readjustment of precipitation.
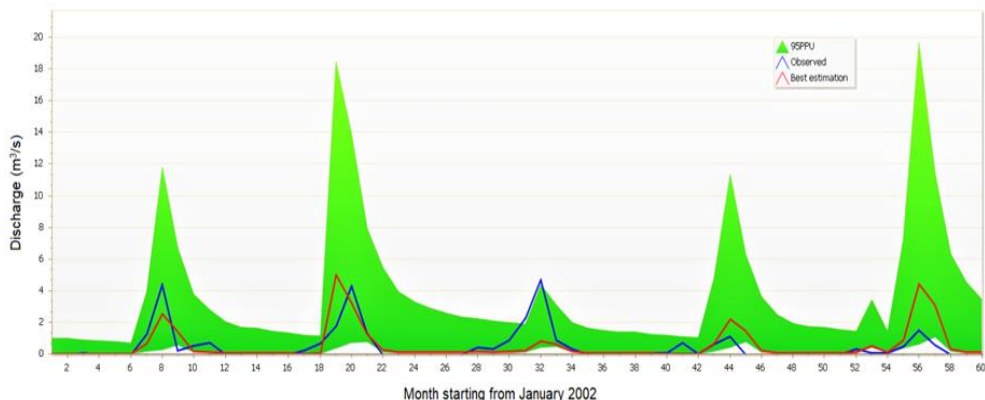
Fig. 3. Comparison of simulated and observed monthly streamflows during the validation period (2002–2006) before readjustment of precipitation.
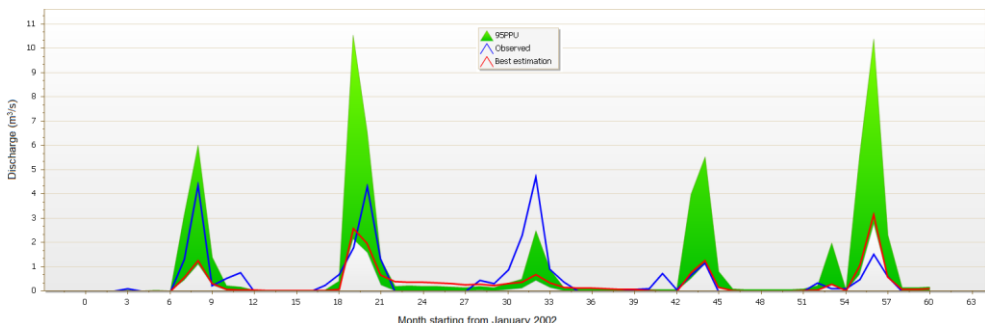

Fig. 4. Comparison of simulated and observed monthly streamflows during the validation period (2002–2006) after readjustment of precipitation.

Errors due to input climate data have an understandably considerable influence on the unacceptable model efficiency. Besides, significant mismatch in spatial and temporal distributions of CFSR-based precipitation compared to ground observations has been reported in the authors' recent works [8, 14]. As clearly articulated in the introduction section, it was this fact that led us to think of readjusting the reanalysis datasets, depending on the available ground-based information. As a first step towards this direction, we established the SWAT model for the study area and analyzed the degree of improvements in the performance of the SWAT model when readjusted CFSR-based precipitation was used as a forcing variable. As we expected, the readjustment of precipitation produced major changes to the modelling outputs; most of the statistical performance indicators showed improvements.

Nonetheless, the unsatisfactory performance of the SWAT model was specifically recognized in prediction uncertainty during both calibration and validation procedures. The performance indicators of uncertainty, especially the P-factor, failed partially to satisfy the allowable statistical requirements. In the absence of reliable and adequate ground information, it is difficult to pinpoint the actual error generating sources and generalise with great certainty. However, the encouraging achievements in the overall efficiency of the SWAT model, augmented by intensive research in similar regions, it is fair to say that it is possible to narrow down errors due to reanalysis datasets and come up with reasonable streamflow predictions in ungauged or poorly gauged watersheds.

Sensitivity analysis shows the portion of parameters in the model output uncertainties. More sensitive parameters have a higher share of model uncertainties than less sensitive

ones in the model output if that parameter is left uncalibrated. Therefore, sensitivity analysis is the first step that should be taken into consideration in model calibration. However, not all sensitive parameters may be calibrated in ungauged catchments. In this study, there were no available measured parameters. Therefore, it is recommended that further efforts should be made to use all available data sources of the watershed in future studies. This helps us to exclude less sensitive parameters from calibration and avoid unnecessary and arbitrary adjustments of parameters.

4 Conclusions

The suitability and non-suitability of reanalysis datasets for streamflow simulations in the light of evidence are well documented. Our recent work underpins this fact to some extent, and hence, an investigation on the impact of down-scaled (up-scaled) reanalysis datasets for hydrologic predictions from ungauged or poorly gauged watershed of a semi-arid climate becomes inevitable. In this regard, a hydrologic model, with an emphasis on monthly streamflow predictions, for the Debarwa watershed was established using SWAT and reanalysis datasets. Here two scenarios were considered: with and without readjustment of precipitation. Findings demonstrated that the SWAT model relatively produced better results when readjusted precipitation (in this case, down-scaled) was used as model input. While most of the objective functions proved acceptable model efficiency, prediction uncertainty (expressed by P-factor and R-factor) did not fully meet the relevant standards during both calibration and validation procedures. R-factor showed a great improvement following the readjustment procedure, whose values were within the recommended threshold values, whereas P-factor was far less than the standards. So, the lacks of balance among these two indices are yet to be worked out. Among the different sources of model errors, we believe errors due to forcing data input and unreliability of ground hydrometric data are equally likely accounted for. However, this does not mean that the plausible scenario that other sources of errors could influence the model's performance is totally out of consideration.

Even though reanalysis datasets have apparently great advantage over ground observations in terms of their simplicity, the findings from this study underscored the need for critical re-examination of both data sources in the context of Eritrea and possibly in regions with similar climatic settings. To minimize model uncertainty and improve the prediction capabilities, the declining tendency of existing ground data collection systems needs to be reversed. Besides, the suitability of CFSR datasets to simulate hydrological responses in similar regions but with reliable and data-rich regions would be the best course of action as it helps to substantiate the current findings with supportive and sound evidence. Therefore, due consideration shall be given to the scenarios addressed in the present study and other mechanisms of handling reanalysis datasets. Otherwise, the uses of reanalysis datasets at face value may generate erroneous hydrological predictions in the regions.

References

1. Ghebrehiwot A. A. and Kozlov D. V., *Hydrological modelling for ungauged basins of arid and semi-arid regions: review*, Vestn. MGSU, **14** (8), pp. 1023–1036, (2019).
2. Hrachowitz M. et al., *A decade of Predictions in Ungauged Basins (PUB)-a review*, Hydrological Sciences Journal, **58** (6), pp. 1198–1255, (2013).
3. Sivapalan M. *et al.,* IA*HS Decade on Predictions in Ungauged Basins (PUB)*, 2003-2012: Shaping an exciting future for the hydrological sciences, Hydrol. Sci. J., **48** (6), pp. 857–880, 2003.

4. Fekete B. M. and Vörösmarty C. J., *The current status of global river discharge monitoring and potential new technologies complementing traditional discharge measurements*, in Predictions in Ungauged Basins: PUB Kick-off (Proceedings of the PUB Kick-off meeting held in Brasilia,20–22 November 2002), IAHS publication, 2007.

5. Vörösmarty C. et al., Global water data: *A newly endangered species*, Eos, **82** (5), (2001).

6. Shiklomanov A. I., Lammers R. B. and Vörösmarty C. J., Widespread decline in hydrological monitoring threatens Pan-Arctic research, Eos (Washington. DC), **83** (2), (2002).

7. Shiklomanov A. I. and Lammers R. B., *Record Russian river discharge in 2007 and the limits of analysis*, Environ. Res. Lett., **4** (045015), p. 9, (2009).

8. Ghebrehiwot A. A. and Kozlov D. V., *Statistical and spatial variability of climate data in the Mareb-Gash river basin in Eritrea*, Vestn. MGSU, **15** (1), pp. 85–99, (2020).

9. Dile Y. T. and Srinivasan R., *Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: An application in the blue nile river basin*, J. Am. Water Resour. Assoc., **50 (**5), pp. 1226–1241, (2014).

10. Fuka D. R., Walter M. T., Macalister C., Degaetano A. T., Steenhuis T. S. and Easton Z. M., *Using the Climate Forecast System Reanalysis as weather input data for watershed models*, Hydrol. Process., **28** (22), pp. 5613–5623, (2014).

11. Auerbach D. A., Easton Z. M., Walter M. T., Flecker A. S. and Fuka D. R., *Evaluating weather observations and the Climate Forecast System Reanalysis as inputs for hydrologic modelling in the tropics*, Hydrol. Process., **30** (19), pp. 3466–3477, (2016).

12. Mararakanye N., Le Roux J. J. and Franke A. C., *Using satellite-based weather data as input to SWAT in a data poor catchment*, Phys. Chem. Earth, **117**, (2020).

13. Grusson Y., Anctil F., Sauvage S., and Pérez J. M. S., *Testing the SWAT model with gridded weather data of different spatial resolutions*, Water (Switzerland), **9** (1), (2017).

14. Ghebrehiwot A. A. and Kozlov D. V., *Assessment of applicability of MIKE 11-NAM hydrological module for rainfall runoff modelling in a poorly studied river basin*, Vestn. MGSU, **15** (7), pp. 1030–1046, (2020).

15. Refsgaard J. C. and Knudsen J., *Operational validation and intercomparison of different types of hydrological models*, Water Resour. Res., **32** (7), pp. 2189–2202, (1996).

16. Nkiaka E., Nawaz N. R., and Lovett J. C., *Evaluating global reanalysis datasets as input for hydrological modelling in the Sudano-Sahel region*, Hydrology, **4** (1), (2017).

17. Arnold J. G., Srinivasan R., Muttiah R. S., and Williams J. R., *Large area hydrologic modeling and assessment*. Part I: Model development, J. Am. Water Resour. Assoc., **34** (1), pp. 73–89, (1998).

18. Arnold J. G., Srinivasan R., Muttiah R. S., and Allen P. M., *Continental scale simulation of the hydrologic balance*, J. Am. Water Resour. Assoc., **35** (5), (1999).

19. Schuol J., Abbaspour K. C., Yang H., Srinivasan R., and Zehnder A. J. B., *Modeling blue and green water availability in Africa*, Water Resour. Res., **44** (7), (2008).

20. Schuol J., Abbaspour K. C., Srinivasan R., and Yang H., *Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model*, J. Hydrol., **352** (1–2), (2008).

21. Gassman P. W., Reyes M. R., Green C. H., and Arnold J. G., *The soil and water assessment tool: Historical development, applications, and future research directions*, Transactions of the ASABE, **50** (4). (2007).

22. Arnold J. G., et al., *SWAT: Model use, calibration, and validation*, Trans. ASABE, **55** (4), (2012).

23. Abbaspour K. C., SWAT-CUP 2012: SWAT Calibration and Uncertainty Programs - A User Manual, Sci. Technol., (2014).

24. Abbaspour K. C., Rouholahnejad E., Vaghefi S., Srinivasan R., Yang H., and Kløve B., *A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model*, J. Hydrol., **524**, pp. 733–752, (2015).

25. Yang D. and Musiake K., *A continental scale hydrological model using the distributed approach and its application to Asia*, Hydrol. Process., **17** (14), (2003).

26. Kozlov D. V. and Ghebrehiwot A. A., *Efficacy of digital elevation and Nash models in runoff forecast*, Mag. Civ. Eng., **3** (87), pp. 103–122, (2019).

27. Gehbrehiwot A. and Kozlov D., *GIUH-Nash based runoff prediction for Debarwa catchment in Eritrea*, E3S Web Conf., **97** (05001), pp. 1–11, (2019).

28. Abbaspour K. C., Vaghefi S. A., Yang H., and Srinivasan R., *Global soil, landuse, evapotranspiration, historical and future weather databases for SWAT Applications*, Sci. data, **6** (1), p. 263, (2019).

29. Abbaspour K. C., Johnson C. A., and van Genuchten M. T., E*stimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure*, Vadose Zo. J., **3** (4), pp. 1340–1352, (2004).

30. Kouchi D. H., Esmaili K., Faridhosseini A., Sanaeinejad S. H., Khalili D., and Abbaspour K. C., *Sensitivity of calibrated parameters and water resource estimates on different objective functions and optimization algorithms*, Water (Switzerland), **9** (6), (2017).

31. Gupta H. V., Sorooshian S., and Yapo P. O., *Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration*, J. Hydrol. Eng., **4** (2), pp. 135–143, (1999).