Evaluation and application of alternative rainfall data sources for forcing hydrologic models in the Mara Basin

Tadesse Alemayehu, Fidelis Kilonzo, Ann van Griensven and Willy Bauwens

ABSTRACT

Accurate and spatially distributed rainfall data are crucial for a realistic simulation of the hydrological processes in a watershed. However, limited availability of observed hydro-meteorological data often challenges the rainfall–runoff modelling efforts. The main goal of this study is to evaluate the Climate Forecast System Reanalysis (CFSR) and Water and Global Change (WATCH) rainfall by comparing them with gauge observations for different rainfall regimes in the Mara Basin (Kenya/Tanzania). Additionally, the skill of these rainfall datasets to simulate the observed streamflow is assessed using the Soil and Water Assessment Tool (SWAT). The daily CFSR and WATCH rainfall show a poor performance (up to 52% bias and less than 0.3 correlation) when compared with gauge rainfall at grid and basin scale, regardless of the rainfall regime. However, the correlations for both CFSR and WATCH substantially improve at monthly scale. The 95% prediction uncertainty (95PPU) of the simulated daily streamflow, as forced by CFSR and WATCH rainfall, bracketed more than 60% of the observed streamflows. We however note high uncertainty for the high flow regime. Yet, the monthly and annual aggregated CFSR and WATCH rainfall can be a useful surrogate for gauge rainfall data for hydrologic application in the study area.

Key words | CFSR, evaluation, gauge rainfall, Mara Basin, SWAT, WATCH

Tadesse Alemayehu (corresponding author)
Ann van Griensven
Department of Hydrology and Hydraulic Engineering,
Vrije Universiteit Brussel (VUB),
Pleinlaan 2, Brussels 1050,
Belgium
and
Department of Water Science and Engineering,
IHE Delft Institute for Water Education,
Delft,
The Netherlands
E-mail: *t.abitew@un-ihe.org*

Fidelis Kilonzo
Department of Civil Engineering,
Kenyatta University,
Nairobi City,
Kenya

Willy Bauwens
Department of Hydrology and Hydraulic Engineering,
Vrije Universiteit Brussel (VUB),
Pleinlaan 2, Brussels 1050,
Belgium

INTRODUCTION

Rainfall is one of the most important forcing inputs for hydrologic models (Obled *et al.* 1994) and hence reliable quality data are required for good results (Boughton 2006). However, for catchments located in developing countries, the weather records are often scarce, incomplete and of questionable quality (Schuol & Abbaspour 2007; Li *et al.* 2013; Fuka *et al.* 2014). One alternative data source for poorly gauged and/or ungauged basins is gridded global rainfall data. The advancement in global data assimilation systems, and hence the availability of global reanalysis data, could partly alleviate the forcing rainfall data requirement for watershed models. Reanalysis involves the assimilation of historical observations (over extended period) including *in-situ* and remote sensing measurements using a single consistent assimilation scheme throughout (Saha *et al.* 2010; Dee *et al.* 2011).

The accuracy of reanalysis rainfall has been evaluated in two ways: by directly comparing the data with gauge rainfall (Hwang *et al.* 2013; Li *et al.* 2014; Worqlul *et al.* 2014; Blacutt *et al.* 2015; Roth & Lemann 2016; Xu *et al.* 2016) and by their ability to reproduce observed streamflow using hydrologic models (Li *et al.* 2013; Dile & Srinivasan 2014; Srivastava *et al.* 2014; Seyyedi *et al.* 2015; Radcliffe & Mukundan 2016; Roth & Lemann 2016; Xu *et al.* 2016). Hwang *et al.* (2013) reported that raw downscaled reanalysis predictions (of the Florida State University Center for Ocean-Atmospheric Prediction Studies CLARReS10 data) accurately reproduced the seasonal trends of precipitation, but generally

doi: 10.2166/nh.2017.081

overestimated the monthly mean and standard deviation of daily precipitation. Worqlul *et al.* (2014) also observed a good performance of daily Climate Forecast System Reanalysis (CFSR) rainfall with gauge observations and, on average, for 38 stations, 86% of the observed rainfall variation is explained by CFSR data in the Upper Blue Nile Basin (Ethiopia). Li *et al.* (2014) compared the Water and Global Change (WATCH) forcing rainfall against gauge rainfall across several regions in India with varying seasonality. Their results showed underestimations of the WATCH rainfall but a fair general agreement. In contrast, Roth & Lemann (2016) noted a poor performance of daily CFSR rainfall for bimodal and unimodal rainfall climates in north-western Ethiopia. The divergent performances of the global rainfall products necessitate more research in several hydro-climatic regions to further explore their usability as a surrogate for observation in data-scarce regions.

Several studies have applied weather data from global reanalysis products to drive hydrological models at watershed and regional scale (El-Sadek *et al.* 2011; Wang *et al.* 2011; Li *et al.* 2013; Bastola & Misra 2014; Dile & Srinivasan 2014; Fuka *et al.* 2014; Seyyedi *et al.* 2015; Radcliffe & Mukundan 2016; Roth & Lemann 2016; Xu *et al.* 2016). A study by Srivastava *et al.* (2014) showed a poor performance of dynamically downscaled daily precipitation from ERA interim using the conceptual probability distribution model (PDM) in simulating the observed streamflow for the Brue catchment (UK). Similarly, Roth & Lemann (2016) reported unsatisfactory daily streamflow simulations skill of the Soil and Water Assessment Tool (SWAT) model forced by daily CFSR rainfall in the upper Blue Nile Basin (Ethiopia). In contrast, several studies reported a good skill of reanalysis rainfall in simulating streamflow (Wang *et al.* 2011; Dile & Srinivasan 2014; Fuka *et al.* 2014; Radcliffe & Mukundan 2016). Fuka *et al.* (2014) evaluated the use of CFSR rainfall for streamflow simulation using the SWAT for catchments with varying area, from 20 to 1,200 km^2. In their study they found that the SWAT model forced by CFSR rainfall showed a slightly better model performance as compared to a model forced by gauge rainfall. Li *et al.* (2013) simulated the water balance of 22 gauged basins in southern Africa using the water and snow balance modelling system (WASMOD-D) forced by WATCH rainfall and observed a good reproduction of the observed flow duration curves (FDCs) for the majority of the basins. Given the spread of the reported performances of global rainfall products for forcing hydrological models, it is imperative to further investigate the streamflow simulation skills.

Rainfall monitoring networks are declining in developing countries. According to the Kenyan meteorology department (KMD), the number of rainfall measuring stations declined from 2,000 (in year 1977) to 700 (at present) (KMD 2014). Despite all the evaluation studies of rainfall in the region, the knowledge about the skills and limitations of global rainfall products in the Mara Basin is still poor. Therefore, evaluating the performance of global rainfall products using available historical gauge observation is important in order to assess the usability of such products in the region. The main goal of this paper is to evaluate CFSR and WATCH rainfall: (i) by comparing reanalysis rainfall with gauge rainfall and (ii) by assessing their skill in simulating the observed flow at the catchment outlet. This will provide an insight into the quality of the rainfall products for hydrological applications in this data-scarce region.

MATERIALS AND METHODS

The study site

The Mara River Basin, a transboundary river basin shared by Kenya and Tanzania, is located between 1.29° and 0.37° S latitude, and between 33.97° and 35.86° E longitude (Figure 1). This river originates from the forested Mau Escarpment (about 3,000 m.a.s.l) and flows through agricultural and rangelands before entering the Masai-Mara Game Reserve in Kenya and the Seregenti National Park in Tanzania and finally joins Lake Victoria. Rainfall is highly variable in the basin with the highest and lowest mean annual rainfall being 1,750 mm yr^{-1} (Mau region) and 600 mm yr^{-1} (south-east part), respectively (WREM 2008). This is mainly due to its equatorial location and its range of land forms, including high mountains, expansive plains and a large inland lake. The rainfall pattern in most parts of the basin is bi-modal, with a short rainy season (October–December) driven by convergence and southward migration of the intertropical convergence zone (ITCZ) and long rainy

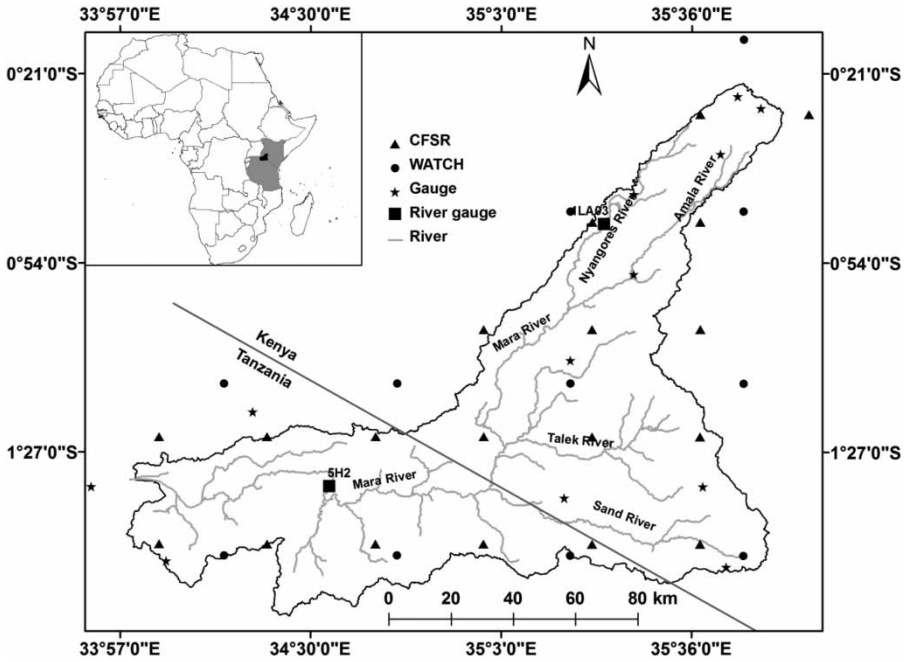

Figure 1 | Location of the Mara River Basin along with the location of the local gauges, and the CFSR (38 km) and WATCH (50 km) grid points.

season (March–May) driven by south-easterly trade winds. The mean annual temperature is approximately 25.5°C and, in general, the temperature increases southwards.

Data availability

Observed gauge rainfall data (1977–1990) for several stations (shown in Figure 1) were acquired from the Kenyan and Tanzanian Meteorological Agency. The rainfall time series has, on average, 15% gaps and thus the gaps were filled using available same day measurements from a nearby gauge. The lack of complete data will affect the calibration results; however, the quality is sufficient for evaluation purposes. A total of ten gauge stations were used in the study.

The WATCH forcing dataset has been produced under the European Water and Global Change project (http://www.eu-watch.org). The data are derived from the ERA-40 reanalysis product via sequential interpolation to half-degree resolution, based on monthly observations of the Climate Research Unit (CRU) and GPCC. The data period spans from 1950 to 2001 (Weedon *et al.* 2011). The WATCH grids in and around the study area from 1977 to 1990 were used in this study.

The CFSR dataset of the National Centers for Environmental Prediction in the USA (NCEP) has been developed as part of the Climate Forecast System project (Saha *et al.* 2010). The CFSR has a global horizontal resolution of c. 38 km. The SWAT team in Texas A&M University provides precipitation, minimum and maximum temperature, humidity, wind speed and solar radiation data (Global-weather 2013). Therefore, CFSR grids falling in and around the Mara Basin from 1979 to 1990 were used in this study. There are about 16 and 10 grids falling in and around the basin for CFSR and WATCH, respectively.

The streamflow data for the Mara River at Mines were obtained for 1980–1990. However, these data have a great deal of discontinuity and hence only periods with a relatively better completeness, i.e., 1980–1982 and 1988–1990, were used.

The development of the SWAT model and calibration scheme

The Mara River Basin was delineated using a high resolution (30 m) digital elevation model (DEM) (NASA 2014) in ArcSWAT2012. The basin was subdivided into 57 sub-basins to

spatially differentiate areas of the basin dominated by different land use and/or soil with dissimilar impact on hydrology. Each sub-basin was further discretized into several HRUs, which represent unique combinations of soil, land use and slope classes. The land cover classes for the basin were obtained from the FAO-Africover project (www.africover.org). Generally speaking, the dominant portion of the basin is covered by natural vegetation including savanna grassland and shrubland. We extracted the soil classes for the basin from the Harmonized Global Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC 2009). A soil properties database for the Mara River Basin was established using the soil water characteristics tool (SPAW, http://hydrolab.arsusda.gov/soilwater).

We developed three SWAT models that differ only with regard to the forcing rainfall (such as gauge, CFSR and WATCH). To each sub-basin, a rainfall time series was assigned based on the proximity to the centroid of the sub-basins. To derive the reference evapotranspiration (ET_r) we used the Hargreaves method, based on measured maximum and minimum temperatures at four locations.

We calibrated the models from 1980 to 1982 excluding three years of warm-up period using the sequential uncertainty fitting (SUFI-2) algorithm (Abbaspour *et al.* 2004). The model results were validated for the 1988 to 1990 period. We used 14 SWAT parameters to calibrate the models by comparing observed daily streamflow for the Mara River at Mara Mines. The Kling–Gupta efficiency (KGE), that measures the linear correlation, the bias and the variability of stream flow (Gupta *et al.* 2009) was used as objective function to calibrate the SWAT models. The SUFI-2 algorithm (Abbaspour *et al.* 2004, 2015) maps all uncertainties (parameter, model structure and input, among others) on the parameters. In an iterative process, the model tries to capture most of the measured data within the 95% prediction uncertainty (95PPU). The 95PPU is calculated at the 2.5% and 97.5% levels of the cumulative distribution of an output variable obtained through Latin hypercube sampling.

Evaluation approach

Direct evaluation

The WATCH and CFSR rainfall are compared with gauge observations representing mono-modal and bi-modal rainfall regimes (1980–1990). This means rainfall estimates from WATCH and CFSR – with 50 km and 38 km grids size, respectively – are compared with the closest point gauge observations. It is acknowledged here that the spatial mismatch and therefore the rainfall datasets are compared at basin level to reduce the spatial representative effect. In SWAT, each sub-basin is assigned one gauge/grid depending on proximity to the centroid of the sub-basin. Thus, the basin areal rainfall is area weighted aggregate of sub-basins' rainfall.

To quantify and illustrate the degree of agreement, we use the Taylor diagram (Taylor 2001), which provides a way of graphically summarizing how closely a pattern (or a set of patterns) matches observations using their correlation, their centred root-mean-square difference and their standard deviations.

Indirect evaluation

The reliability of WATCH and CFSR rainfall for forcing hydrological models can be evaluated indirectly using skills in predicting the observed streamflow. This approach, in fact, measures the total error on the simulated streamflow stemming from the forcing rainfall and the hydrologic model. Several hydrologic modelling studies, in general, agree on the fact that multiple parameter combinations can result in equally acceptable streamflow (Beven 2006). To account for this issue, we use multiple statistical measures: the p-factor, r-factor and the KGE. The p-factor is the fraction of measured data (plus its error) bracketed by the 95% prediction uncertainty (95PPU) band and varies from 0 to 1, where 1 indicates 100% bracketing of the measured data within model prediction uncertainty (i.e., a perfect model simulation considering the uncertainty). SUFI-2 includes a measured discharge uncertainty of 10% in the analysis of the p-factor (Abbaspour *et al.* 2004, 2015). The r-factor, on the other hand, is the ratio of the average width of the 95PPU band and the standard deviation of the measured streamflow. The desirable value for r-factor is less than 1.5 (Abbaspour *et al.* 2004). It is assumed reliable rainfall estimates results in low prediction uncertainty.

RESULTS AND DISCUSSION

The assessment of CFSR and WATCH rainfall

Figures 2 and 3 illustrate the performance of CFSR and WATCH rainfall for grid scale – mono-modal and bi-modal rainfall regimes – and for basin scale using Taylor diagrams. The reference point – the circle mark on the x-axis – represents a perfect fit between global rainfall data and gauge rainfall. The position of each rainfall dataset is determined by the values of r, the centred RMSE and the normalized standard deviation. The daily CFSR and WATCH rainfall correlates poorly (i.e., less than 0.3 correlation) with gauge rainfall (Figure 2) for both rainfall regimes, albeit with slight improvements for basin scale comparison. The daily CFSR rainfall has an average bias of −14.9% and the bias WATCH rainfall is about −10.5% for mono-modal rainfall regimes. The CFSR rainfall exhibits substantial bias (about 52%) compared to gauge rainfall for bi-modal rainfall regime, while we note minimal bias (−2.8%) for WATCH rainfall. At basin scale, the daily CFSR and WATCH rainfall have low biases (i.e., less than 1.1%).

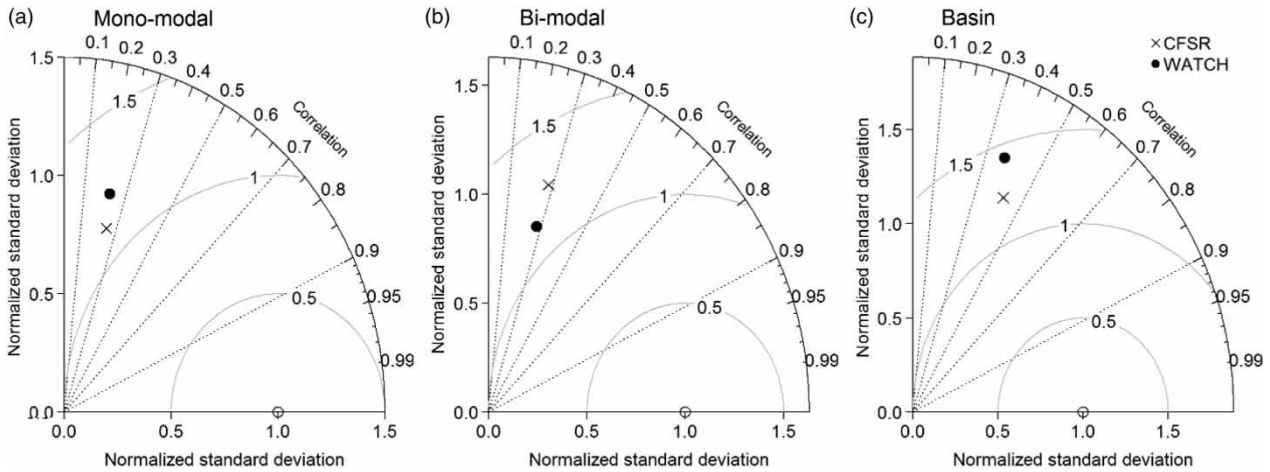

Figure 2 | The Taylor diagrams of daily CFSR and WATCH rainfall against gauge observations for 1980–1990.

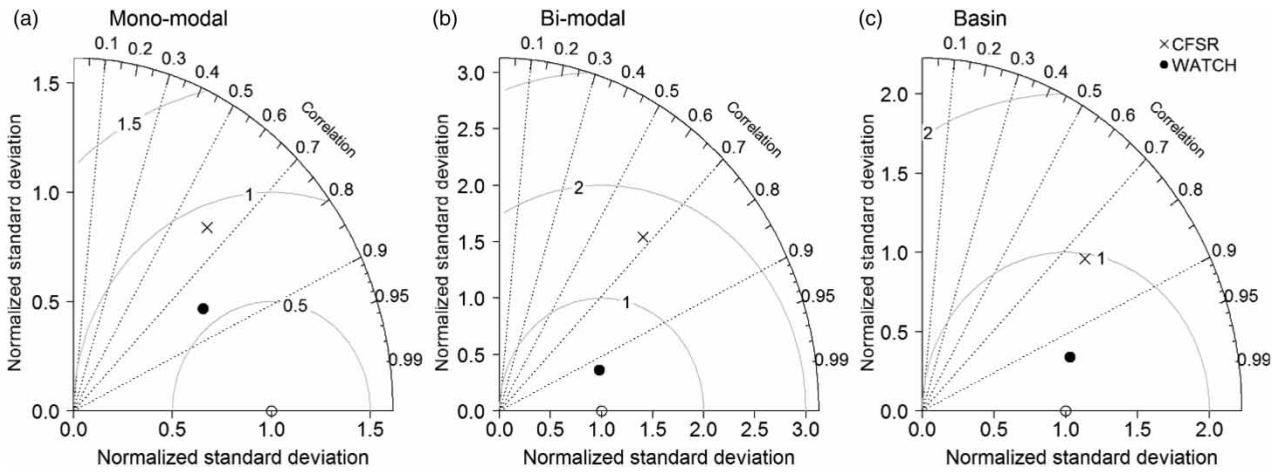

Figure 3 | The Taylor diagrams of monthly CFSR and WATCH rainfall against gauge observations for 1980–1990.

Unlike the poor performances of CFSR and WATCH rainfall at a daily scale, the monthly aggregated rainfall from these global products show a good agreement with the monthly gauge rainfall (Figure 3). The monthly WATCH rainfall shows a correlation of 0.81 and 0.94 and with gauge observations for the mono-modal and bimodal rainfall regimes, respectively. The performances of monthly CFSR rainfall are also improved, but not as good as WATCH rainfall. The WATCH rainfall consistently outperforms the CFSR rainfall at basin and at grid scale, albeit coarse scale. The good performance of the monthly WATCH rainfall is partly ascribed to the adjustments using CRU and GPCC rainfall (Weedon *et al.* 2011).

Figure 4 presents the comparison of mean monthly rainfall for gauge, CFSR and WATCH rainfall (1980–1990). The

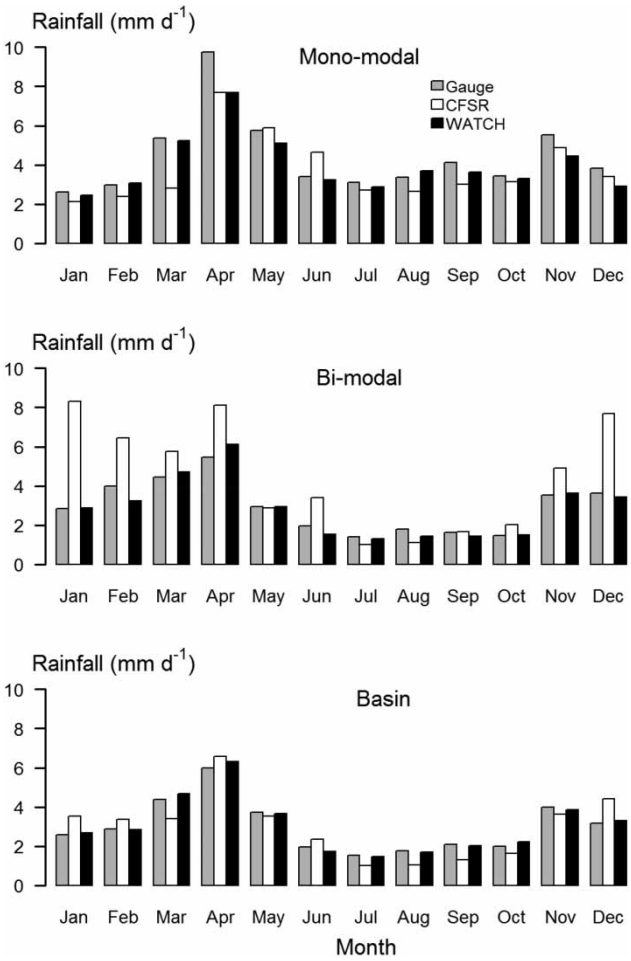

Figure 4 | Comparison of the mean monthly WATCH and CFSR rainfall with the gauge rainfall (1980–1990).

WATCH rainfall shows a fair agreement with average bias of 5% compared to gauge rainfall at basin scale and we note a general match both in magnitude and seasonal pattern. The CFSR rainfall shows high average biases – up to 5.5 mm d^{-1} during wet months – for the bi-modal rainfall regime, but a fair performance for the mono-modal regime. Nevertheless, both rainfall products show a fair agreement with gauge rainfall at a basin level.

Performance of the SWAT model driven by CFSR, WATCH and gauge rainfall

Table 1 summarizes the performance of the SWAT models forced by daily gauge, CFSR and WATCH rainfall in reproducing the daily observed streamflow for the calibration and validation periods. With the exception of CFSR rainfall, the SWAT models forced by WATCH and gauge rainfall exhibit good skills in reproducing the observed streamflow with KGE of 0.75 and 0.81, respectively, during the calibration period. During the validation period, the performance of CFSR rainfall improves considerably while WATCH and gauge rainfall shows slight changes. The improved performance of CFSR rainfall during the validation period is attributed to the underestimation of rainfall (6%) during this period compared to gauge rainfall and thus reduced the overestimation tendency of the model for peak streamflow. This is noted with the comparable standard deviation of CFSR (0.72 mm d^{-1}) and observation (0.79 mm d^{-1}). Figures 5 and 6 present the observed streamflow time series along with the 95PPU for the calibration and validation periods, respectively. It is clear from the illustration that a substantial part of the observed streamflow is captured in the 95PPU; however, the uncertainty band is high for peak flows. The SWAT models forced by WATCH and gauge rainfall are able to capture, respectively, 81% and 75% of the observed streamflow within the 95PPU, for the calibration period. These data sources also show similar performances during the validation period. The CFSR rainfall is able to capture less observed streamflow within the 95PPU as compared to the gauge and WATCH rainfall, for both calibration and validation periods. Despite a generally good p-factor for gauge, WATCH and CFSR rainfall, we note a high r-factor during the calibration period. The r-factor improves to some extent during the validation

Table 1 | Performance summary for streamflow simulation skills

	Calibration (1980–1982)					Validation (1988–1990)				
	p-factor	r-factor	KGE	Mean[a]	STD[a]	p-factor	r-factor	KGE	Mean[a]	STD[a]
Observation	–	–	–	0.33	0.49	–	–	–	0.49	0.79
Gauge	0.84	1.54	0.72	0.35	0.45	0.72	1.07	0.67	0.53	0.70
CFSR	0.68	1.35	0.43	0.33	0.66	0.59	0.89	0.63	0.48	0.72
WATCH	0.75	1.24	0.63	0.32	0.47	0.75	1.5	0.66	0.49	0.91

[a]in mm d^{-1}; KGE and STD denote Kling and Gupta efficiency and standard deviation, respectively.
p- and r-factors quantify the percentage of observed streamflow bracketed in the 95% prediction uncertainly (95PPU) along with the average band width, respectively

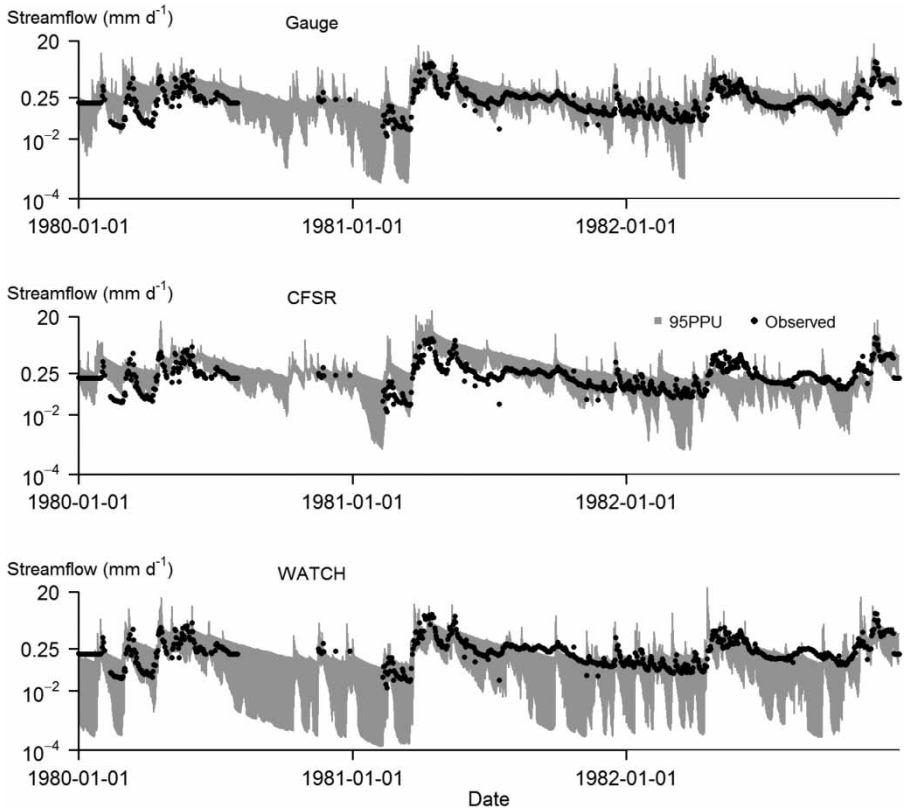

Figure 5 | The illustration of the 95% prediction uncertainty (95PPU) along with the measured streamflow at Mara Mines for the calibration period (1980–1982). Note that the y-axis is in log scale.

period, particularly for CFSR and yet all the models' r-factors indicate a somewhat wide 95PPU band. Abbaspour *et al.* (2015) recommends a p-factor value larger than 0.7 and an r-factor smaller than 1.5 for a satisfactory model performance. Based on these criteria, the three SWAT models performed satisfactorily, yet we visually note a high uncertainty for the high flows. For instance, we note high uncertainties for the flows simulated with the CFSR rainfall during the rainy season of 1981. Similarly, high uncertainties are observed when using the WATCH rainfall during the validation period. The errors noted in WATCH and CFSR rainfall are propagated and amplified by the nonlinear processes in the SWAT model. As discussed in the section 'The assessment of CFSR and WATCH rainfall',

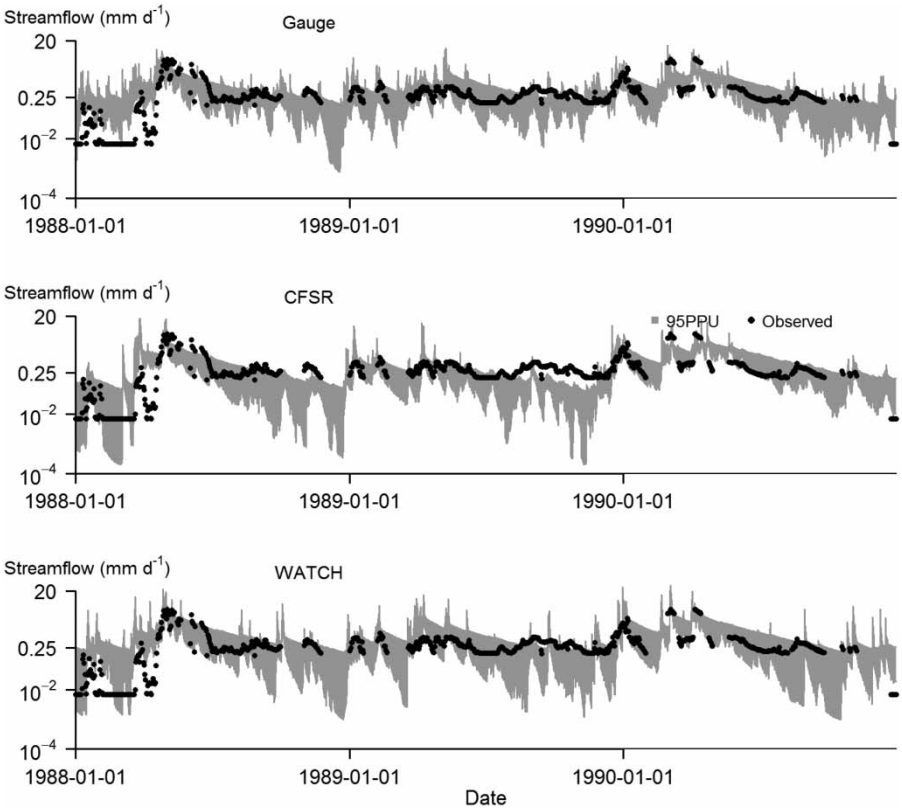

Figure 6 | The illustration of the 95% prediction uncertainty (95PPU) along with the measured streamflow at Mara Mines for the validation period (1988–1990). Note that the y-axis is in log scale.

the daily CFSR and WATCH rainfall are poorly correlated with gauge observations. Errors in the temporal variability of rainfall are more influential than errors in mean precipitation in causing errors in mean streamflow (Hwang *et al.* 2013). In conclusion, the CFSR and WATCH rainfall have strong limitations for simulating daily streamflow, particularly for flooding periods. This is in agreement with Xu *et al.* (2016), who demonstrated the limited skill of WATCH rainfall to simulate flood discharge in Xiangjiang River basin (China). They also noted the poor performance of reanalysis datasets at short temporal resolution. Seyyedi *et al.* (2015) found strong biases in flood simulation using GLDAS rainfall in the USA.

Evaluation of simulated streamflow consistency

To have a further insight into the prediction ability of the global rainfall data for long-term analysis, we reran the calibrated SWAT model for the period 1980–1990. Figure 7 presents the FDCs derived from SWAT simulated daily streamflow using gauge, CFSR and WATCH rainfall. Overall, the simulated streamflow reveals a similar flow regime pattern, whereby the flashiness of the Mara Basin is captured well. The daily simulated streamflow using gauge, CFSR and WATCH as input that is exceeded and/or equalled 50% of the time is about 0.14, 0.18 and 0.14 mm d^{-1}, respectively. Also, as shown in Table 1, the mean and standard deviation of the simulated daily streamflow using the different rainfall data are comparable with those of the observed streamflow during the calibration and validation periods. This indicates that reanalysis rainfall data can be used for understanding the streamflow regime using aggregated indices.

The simulated monthly streamflow (1980–1990) based on the different rainfall data is compared in Figure 8. It is apparent that for a monthly temporal resolution, the CFSR and WATCH rainfall data lead to flows that are comparable to those simulated with the gauge rainfall, albeit that the WATCH-based streamflows show an overestimation from

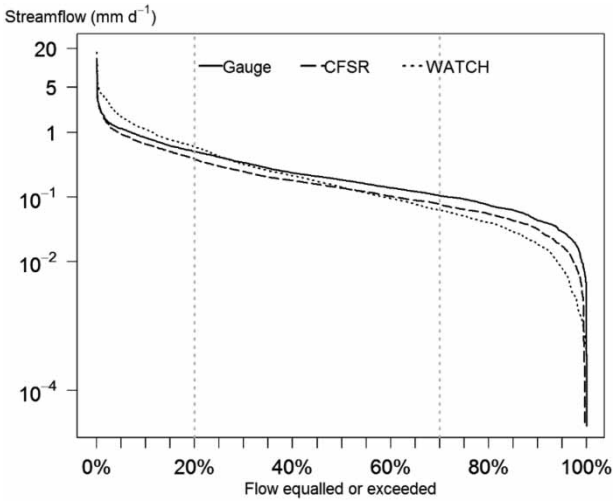

Figure 7 | Comparison of simulated flow duration curves (FDC) using gauge, CFSR and WATCH rainfall for the Mara River at Mines (1980–1990). Note that the y-axis is in log scale.

1988 to 1990. This is attributed to the relatively higher rainfall (on average 110 mm month^{-1}) compared to the average 85.5 mm month^{-1} rainfall during 1980–1987. The good match on monthly simulated streamflow variability can also be observed in Figure 9. The SWAT model forced by the WATCH rainfall better reproduced the streamflow pattern simulated by SWAT using gauge rainfall. However, the streamflow in April and May are overestimated by more than 100%.

The water budget analysis

A synthesis of the water budget components is crucial for having an overview on how the input rainfall is partitioned in a river basin. Figure 10 shows the average of the major water budget components in the study area, as simulated by the calibrated SWAT models using gauge, CFSR and WATCH forcing. The mean annual input rainfall to the SWAT models varies from 1,084 mm yr^{-1} (CFSR) to 1,106 mm yr^{-1} (WATCH). Depending on the input rainfall, the simulated actual evapotranspiration (ET) and total water yield (WYLD) differs over the study area. The SWAT simulated mean annual ET ranges between 83% (WATCH) and 89% (gauge) of the input rainfall, whereas the total water yield (WYLD) ranges between 13% (gauge) to 17% (WATCH) of the input rainfall. These results are in agreement with the modelling results of Dessu & Melesse (2012) for the study area.

The effect of input rainfall on the parameter sensitivity

The sensitivity rank of each SWAT model parameter is analysed using SUFI-2 global sensitivity analysis algorithm to highlight the effect of input rainfall on the sensitivity of the model parameters. Table 2 presents the most sensitive SWAT parameters (not all the parameters used in the calibration) using gauge, CFSR and WATCH rainfall as an input. The SCS curve number (CN2), that primarily controls the surface runoff generation, and the soil evaporation compensation factor (ESCO) are the most sensitive parameters, irrespective of the input rainfall. However, the sensitivity rank of the remaining SWAT parameters varies depending on the driving rainfall data. This indicates the importance of input rainfall selection from varying data sources while parameterizing watershed processes. However, we note

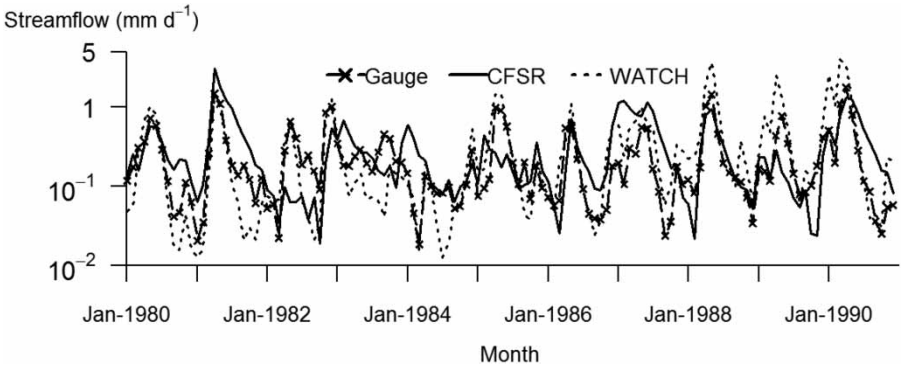

Figure 8 | The simulated monthly streamflows for the Mara River (1980–1990).

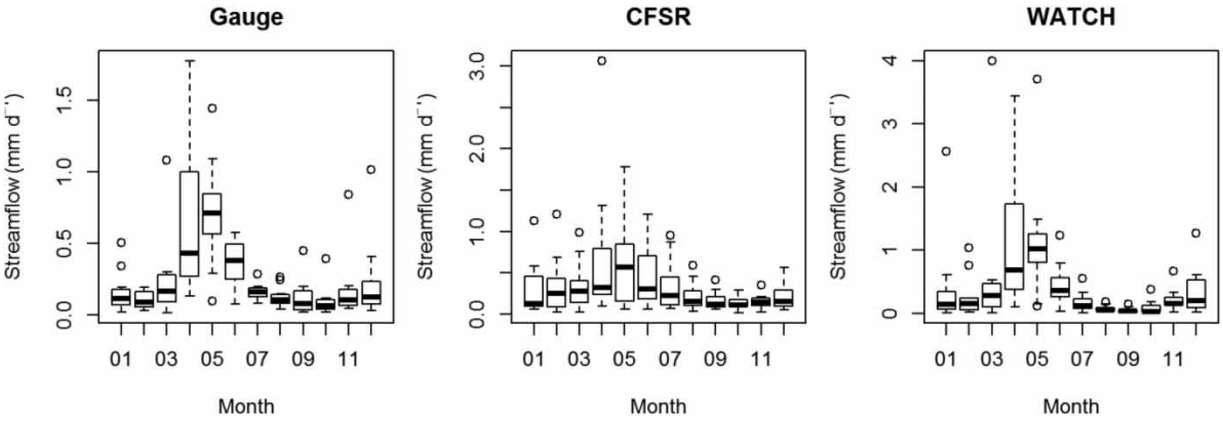

Figure 9 | Monthly variability of simulated seasonal streamflow for the Mara River at Mines (1980–1990). The boxes represent 1.5 times the interquartile range, with the bottom and top part indicating the 25th and the 75th quartiles, respectively. The circles denote extreme values (outliers).

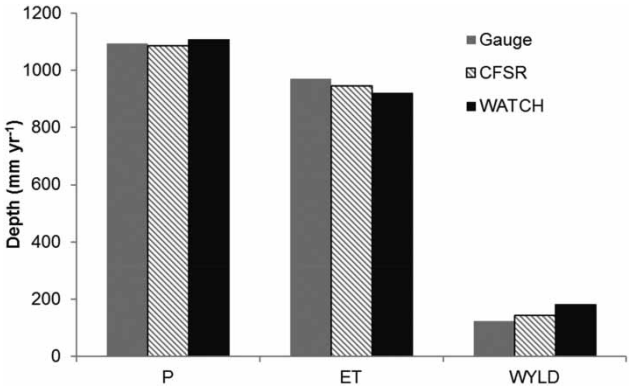

Figure 10 | The mean annual water balance components of the Mara Basin (1980–1990) using gauge, CFSR and WATCH rainfall are comparable.

Table 2 | The top sensitive parameters with varying input rainfall

Sensitivity rank	Gauge	CFSR	WATCH
1	CN2	CN2	CN2
2	ESCO	ESCO	ESCO
3	SOL_Z	SOL_Z	ALPHA_BF
4	SOL_AWC	GW_DELAY	SOL_AWC
5	GW_DELAY	ALPHA_BF	REVAPMN

CN2, SCS curve number II (–); ESCO, soil evaporation compensation factor (–); SOL_Z (mm), soil layer depths; SOL_AWC, soil available water (mm); GW_DELAY, delay time for aquifer recharge (days); ALPHA_BF, baseflow recession constant (days); REVAPMN, depth of water in the aquifer for revap (mm).

here the need for caution when interpreting evaluation results since parameter values can compensate for the quality of input rainfall and use of multi-parameter simulation is important, as suggested in Bastola & Misra (2014).

CONCLUSIONS

In this study two globally available rainfall data – the CFSR and the WATCH data – are evaluated by direct comparisons with gauge observations and with regard to their capability for simulating streamflow in the Mara Basin.

For the comparison with the rainfall gauge data, we distinguished stations with a mono-modal and stations with a bi-modal rainfall regime. The daily CFSR and WATCH rainfall data show poor performances at grid and basin spatial scales when compared with gauge observations, regardless of the rainfall regime pattern. However, their performance improved substantially at a monthly temporal resolution, whereby we noted a correlation of up to 0.76 (CFSR) and up to 0.95 (WATCH). Despite its coarser areal resolution, the WATCH rainfall generally outperforms the CFSR rainfall.

The SWAT models for the River Mara were built and calibrated using daily CFSR, WATCH and gauge rainfall data. The 95% prediction uncertainty (95PPU) of the streamflows simulated using the CFSR and WATCH rainfall bracketed more than 60% of the observed daily streamflow. However, considering the high uncertainty range on the streamflows – particularly for the high flows – their utility for daily streamflow hydrograph simulation is strongly limited, and that is consistent with the poor performance of daily WATCH and CFSR rainfall. On the other hand, it has been shown that CFSR and WATCH rainfall could be a useful surrogate for observations in order to investigate the aggregated streamflow behaviour, e.g., using FDCs or

to simulate the monthly streamflow which provides crucial information for water resources management and planning.

The SWAT simulated mean annual ET ranges between 83% (WATCH) and 89% (gauge) of the input rainfall, whereas the total water yield (WYLD) ranges 13% (gauge) to 17% (WATCH) of the input rainfall. Despite the slight differences in absolute magnitudes, the partitioning of the rainfall by the different SWAT models is comparable, showing that CFSR and WATCH rainfall-driven models can be used for the analysis of the annual water balance in the study area.

We recommend future studies to bias-correct the WATCH and CFSR rainfall using gauge observations to improve the daily rainfall variability and hence improve daily streamflow simulation skills.

ACKNOWLEDGEMENTS

The authors thank the Ministry of Water and Irrigation in Kenya and Tanzania for providing the observed hydro-climatological data for the Mara Basin.

REFERENCES

Abbaspour, K. C., Johnson, C. A. & van Genuchten, M. T. 2004 Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zo. J.* **3**, 1340–1352. doi:10.2113/3.4.1340.

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H. & Kløve, B. 2015 A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. *J. Hydrol.* **524**, 733–752. doi:10.1016/j.jhydrol.2015.03.027.

Bastola, S. & Misra, V. 2014 Evaluation of dynamically downscaled reanalysis precipitation data for hydrological application. *Hydrol. Process.* **28**, 1989–2002. doi:10.1002/hyp.9734.

Beven, K. 2006 A manifesto for the equifinality thesis. *J. Hydrol.* 320, 18–36. doi:10.1016/j.jhydrol.2005.07.007.

Blacutt, L. A., Herdies, D. L., de Gonçalves, L. G. G., Vila, D. A. & Andrade, M. 2015 Precipitation comparison for the CFSR, MERRA, TRMM3B42 and combined scheme datasets in Bolivia. *Atmos. Res.* **163**, 117–131. doi:10.1016/j.atmosres.2015.02.002.

Boughton, W. 2006 Calibrations of a daily rainfall-runoff model with poor quality data. *Environ. Modell. Softw.* **21**, 1114–1128. doi:10.1016/j.envsoft.2005.05.011.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. & Vitart, F. 2011 The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597. doi:10.1002/qj.828.

Dessu, S. B. & Melesse, A. M. 2012 Modelling the rainfall-runoff process of the Mara River basin using the soil and water assessment tool. *Hydrol. Process.* **26**, 4038–4049. doi:10.1002/hyp.9205.

Dile, Y. T. & Srinivasan, R. 2014 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**, 1226–1241. doi:10.1111/jawr.12182.

El-Sadek, A., Bleiweiss, M., Shukla, M., Guldan, S. & Fernald, A. 2011 Alternative climate data sources for distributed hydrological modelling on a daily time step. *Hydrol. Process.* **25**, 1542–1557. doi:10.1002/hyp.7917.

FAO/IIASA/ISRIC/ISSCAS/JRC 2009 *Harmonized World Soil Database (version 1.1)*. FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Fuka, D. R., Walter, M. T., MacAlister, C., Degaetano, A. T., Steenhuis, T. S. & Easton, Z. M. 2014 Using the climate forecast system reanalysis as weather input data for watershed models. *Hydrol. Process.* **28**, 5613–5623. doi:10.1002/hyp.10073.

Globalweather 2013 NCEP Climate Forecast System Reanalysis (CFSR) [WWW Document]. http://globalweather.tamu.edu/ (accessed January 2013).

Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. 2009 Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *J. Hydrol.* **377**, 80–91. doi:10.1016/j.jhydrol.2009.08.003.

Hwang, S., Graham, W. D., Adams, A. & Geurink, J. 2013 Assessment of the utility of dynamically-downscaled regional reanalysis data to predict streamflow in west central Florida using an integrated hydrologic model. *Reg. Environ. Chang.* **13**, 69–80. doi:10.1007/s10113-013-0406-x.

KMD 2014 Kenya Meteorological Department [WWW Document]. URL http://www.meteo.go.ke/data/ (accessed 20 September 2014).

Li, L., Ngongondo, C. S., Xu, C. & Gong, L. 2013 Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in southern Africa. *Hydrol. Res.* **44**, 770–788. doi:10.2166/nh.2012.175.

Li, L., Xu, C.-Y., Zhang, Z. & Jain, S. K. 2014 Validation of a new meteorological forcing data in analysis of spatial and temporal variability of precipitation in India. *Stoch. Environ. Res. Risk Assess.* **28**, 239–252. doi:10.1007/s00477-013-0745-7.

NASA 2014 United States Geological Survey Earth Explorer. Available online: http://earthexplorer.usgs.gov/ (accessed 9 September 2015).

Obled, C., Wendling, J. & Beven, K. 1994 The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *J. Hydrol.* **159**, 305–333. doi:10.1016/0022-1694(94)90263-1.

Radcliffe, D. E. & Mukundan, R. 2016 PRISM vs. CFSR precipitation data effects on calibration and validation of SWAT models. *J. Am. Water Resour. Assoc.* **53**, 89–100. doi:10.1111/1752-1688.12484.

Roth, V. & Lemann, T. 2016 Comparing CFSR and conventional weather data for discharge and soil loss modelling with SWAT in small catchments in the Ethiopian Highlands. *Hydrol. Earth Syst. Sci.* **20**, 921–934. doi:10.5194/hess-20-921-2016.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Van Den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G. & Goldberg, M. 2010 The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* **91**, 1015–1057. doi:10.1175/2010BAMS3001.1.

Schuol, J. & Abbaspour, K. C. 2007 Using monthly weather statistics to generate daily data in a SWAT model application to West Africa. *Ecol. Modell.* **201**, 301–311. doi:10.1016/j.ecolmodel.2006.09.028.

Seyyedi, H., Anagnostou, E. N., Beighley, E. & McCollum, J. 2015 Hydrologic evaluation of satellite and reanalysis precipitation datasets over a mid-latitude basin. *Atmos. Res.* **164–165**, 37–48. doi:10.1016/j.atmosres.2015.03.019.

Srivastava, P. K., Han, D., Rico-Ramirez, M. A. & Islam, T. 2014 Sensitivity and uncertainty analysis of mesoscale model downscaled hydro-meteorological variables for discharge prediction. *Hydrol. Process.* **28**, 4419–4432. doi:10.1002/hyp.9946.

Taylor, K. E. 2001 Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **106**, 7183–7192. doi:10.1029/2000JD900719.

Wang, F., Wang, L., Koike, T., Zhou, H., Yang, K., Wang, A. & Li, W. 2011 Evaluation and application of a fine-resolution global data set in a semiarid mesoscale river basin with a distributed biosphere hydrological model. *J. Geophys. Res.* **116**, D21108. doi:10.1029/2011JD015990.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O. & Best, M. 2011 Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J. Hydrometeorol.* **12**, 823–848. doi:10.1175/2011JHM1369.1.

Worqlul, A. W., Maathuis, B., Adem, A. A., Demissie, S. S., Langan, S. & Steenhuis, T. S. 2014 Comparison of TRMM, MPEG and CFSR rainfall estimation with the ground observed data for the Lake Tana Basin, Ethiopia. *Hydrol. Earth Syst. Sci. Discuss.* **11**, 8013–8038. doi:10.5194/hessd-11-8013-2014.

WREM 2008 *Mara River Basin Monograph, Mara River Basin Transboundary Integrated Water Resources Management and Development Project, Final Technical Report.* WREM International Inc., Atlanta, GA.

Xu, H., Xu, C.-Y., Chen, S. & Chen, H. 2016 Similarity and difference of global reanalysis datasets (WFD and APHRODITE) in driving lumped and distributed hydrological models in a humid region of China. *J. Hydrol.* **542**, 343–356. doi:10.1016/j.jhydrol.2016.09.011.

First received 28 April 2017; accepted in revised form 27 August 2017. Available online 14 November 2017