Assessment of global reanalysis precipitation for hydrological modelling in data–scarce regions: a case study of Kenya

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).
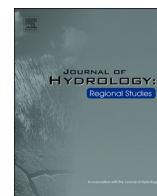
To link to this article DOI: http://dx.doi.org/10.1016/j.ejrh.2022.101105

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: A case study of Kenya

Maureen A. Wanzala [a,*], Andrea Ficchi [a,b], Hannah L. Cloke [a,c,d,e,f], Elisabeth M. Stephens [c,g], Heou M. Badjana [a], David A. Lavers [e,h]

[a] *Department of Geography & Environmental Science, University of Reading, Reading, UK*
[b] *Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy*
[c] *Department of Meteorology, University of Reading, Reading, UK*
[d] *Department of Earth Sciences, Uppsala University, Uppsala, Sweden*
[e] *European Centre for Medium-Range Weather Forecasts, Reading, UK*
[f] *Centre of Natural Hazards and Disaster Science, CNDS, Uppsala, Sweden*
[g] *Red Cross Red Crescent Climate Centre, The Hague, The Netherlands*
[h] *School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK*

ARTICLE INFO	ABSTRACT
	Study region: 19 flood prone catchments in Kenya, Eastern Africa *Study focus:* Flooding is a major natural hazard especially in developing countries, and the need for timely, reliable, and actionable hydrological forecasts is paramount. Hydrological modelling is essential to produce forecasts but is a challenging task, especially in poorly gauged catchments, because of the inadequate temporal and spatial coverage of hydro-meteorological observations. Open access global meteorological reanalysis datasets can fill in this gap, however they have significant errors. This study assesses the performance of four reanalysis datasets (ERA5, ERA-Interim, CFSR and JRA55) over Kenya for the period 1981–2016 on daily, monthly, seasonal, and annual timescales. We firstly evaluate the reanalysis datasets by comparing them against observations from the Climate Hazards group Infrared Precipitation with Station. Secondly, we evaluate the ability of these reanalysis datasets to simulate streamflow using GR4J model considering both model performance and parameters sensitivity and identifiability. *New hydrological insights for the region:* While ERA5 is the best performing dataset overall, performance varies by season, and catchment and therefore there are marked differences in the suitability of reanalysis products for forcing hydrological models. Overall, wetland catchments in the western regions and highlands of Kenya obtained relatively better scores compared to those in the semi-arid regions, this can inform future applications of reanalysis products for setting up hydrological models that can be used for flood forecasting, early warning, and early action in data scarce regions, such as Kenya.

1. Introduction

Precipitation is arguably the most important driver of catchment hydrological response (e.g. MacLeod et al., 2021), but it is

---

challenging to get accurate information on the amount, duration, and intensity of rainfall events (Beck et al., 2017a; Tapiador et al., 2012), due to the high spatio-temporal variability (Nicholson et al., 2019; Vischel et al., 2011). This is compounded by a low spatial coverage and a net decline in the number of ground gauge stations in the historical climatological observation network, especially in developing countries such as Kenya (Menne et al., 2018; Tarek et al., 2020, 2021; Zaitchik et al., 2011). Unreliable or incomplete datasets are unable to correctly identify seasonal or short-range temporal patterns (e.g., Gosset et al., 2013; Le et al., 2017).

Other sources of precipitation data such as those from satellite remote sensing are now available, but they come with their own errors, including random and systematic (see Beck et al., 2021; Beck et al., 2017a; Beck et al., 2017b; Fortin et al., 2015; Sun et al., 2018). Another freely available source of precipitation data are meteorological reanalysis products which are becoming increasingly promising due to upgrades in their spatial resolution and improved representation of atmospheric processes in global models (Hersbach, 2018). Reanalysis data combine a wide range of remotely sensed observations with a dynamical–physical coupled numerical model to produce the best estimate of the state of the atmosphere. Reanalysis is not reliant on the density of surface observational networks and can give surface variables in locations with little to no surface coverage. As a result, they can generate several variables both at the land surface and on vertical atmospheric levels, and hence have been applied in several studies both for climatological and hydrological purposes across the world (e.g., Beck et al., 2017a; Chen et al., 2018; Emerton et al., 2017; Essou et al., 2017). Several different reanalysis products exist but they are known to vary in quality with recurrent upgrades. It is important to evaluate them carefully both to inform the users and the developers of the datasets. The developers of these products can only work on improving their updates when there is a complete feedback loop between applications and developments Therefore, ground validation of reanalysis precipitation is very important but very challenging, particularly where the rain gauge networks are sparse.

Several studies attempt to quantify and account for the sampling errors comparing reanalysis data with observations in different parts of the world (e.g., Guo et al., 2018; Tang et al., 2020; Xu et al., 2020; Zaitchik et al., 2011), at a global scale (Beck et al., 2017a; 2017b), at regional or basin scale (e.g., Acharya et al., 2019; Nkiaka et al., 2017; Tarek et al., 2020) and at a national scale (e.g. Arshad et al., 2021; Gleixner et al., 2020; Koukoula et al., 2020; Lakew et al., 2020; Shayeghi et al., 2020; Tesfaye et al., 2017). However, the findings of these studies were mixed. Differences in approaches, regions, and time scales resulted in inconsistency in product performance, implying that site-specific performance evaluation may be required. Existing studies also aimed at analysing a single product or a few products for short periods of time, thus their estimated errors may not reflect long-term behaviour.

Additionally, the temporal dynamics of rainfall are very important as they play an important role in the total accumulated rainfall on daily and monthly timescales (Ficchì et al., 2016), thus influencing the bimodal seasonality observed over Kenya. Also the highly-variable temporal dynamics are key in explaining the nonlinear nature of infiltration process (Blöschl and Sivapalan, 1995), such as the peak discharge value (Gabellani et al., 2007) and runoff volume (Viglione et al., 2010) in hydrological modelling. Thus, the above highlights the need to consider different temporal scales, when evaluating the reanalysis precipitation relative to observation.

In Kenya, there were 20 major floods from 1964 to 2020 which were driven by precipitation falling in the seasonal rains. More than 160,000 people were displaced countrywide by floods in October 2019 (ReliefWeb, 2019a; 2019b; Opere, 2013). Annual average economic loss from flooding is estimated to be 5.5% of gross domestic product (Njogu, 2021). Thus, understanding the best representation of precipitation in flood models which can be used for forecasting or risk analysis is of great societal importance. Kenya has a widely varying physical geography resulting in great variability of river catchment characteristics across the country. Thus, it is essential not only to understand the representation of precipitation at a country scale, but also on a catchment-by-catchment basis (Golian et al., 2021; Meresa et al., 2021). Previous evaluation of reanalysis products in capturing Kenyan rainfall show varied levels of agreement in spatio-temporal variability relative to observations (e.g., Alemayehu et al., 2018; Dile and Srinivasan, 2014; Gleixner et al., 2020; Khan et al., 2011). Moreover, studies employing hydrological modelling generally used discharge observations from a small number of catchments (e.g., Alemayehu et al., 2018; Bitew et al., 2012; Langat et al., 2017; Le et al., 2017; Worqlul et al., 2017) and did not quantify uncertainties associated with each reanalysis (e.g., Alemayehu et al., 2018), leading to combined rainfall and model uncertainty that is not easily interpreted. Hence, there is a notable gap in the literature associated with evaluating the accuracy of multiple reanalysis products across different catchments, accounting for both model and input errors, especially in data-scarce regions like in Kenya, and this gap was an important motivation for the present study. This paper evaluates four reanalysis precipitation products with respect to observations and assesses their suitability for use in hydrological modelling in 19 Kenyan catchments. We assess their performance in reproducing the most important features of rainfall events and regimes, and in simulating catchment streamflow through answering the following research questions:

- How well do the precipitation datasets compare in terms of temporal dynamics at the basin scale? Which product is the most accurate compared to observations?
- How well do precipitation datasets compare in terms of spatial patterns? Which product shows consistency in spatial heterogeneity compared to observations?
- How does the general hydrological model performance vary with different datasets?
- How does the sensitivity of a rainfall runoff model (GR4J) vary with alternative rainfall forcing?

We consider both model performance and parameter uncertainty and compute a Model Suitability Index (MSI) by coupling the results of model performance statistics and Global Sensitivity Analysis. We compare four reanalysis datasets using the GR4J model across 19 Kenyan catchments with varied climate and morphological characteristics, to investigate which input data is suitable or require caution, when used in the place of observation dataset in different regions. This work is a steppingstone and an essential guide for hydrological applications of global reanalysis datasets because it compares several reanalysis products to observations on daily, monthly, and seasonal scales, and unveils the propagation of uncertainty from different reanalysis when used as model inputs. All the

above reviewed studies looked at the performance of at most one reanalysis dataset in simulating streamflow and only over one catchment; but none looked at such a country-scale performance. To our knowledge, this is the very first evaluation of the different reanalysis products over Kenya for simulating streamflow coupled with sensitivity analysis.

2. Study area and catchment characteristics

The study is undertaken in 19 Kenyan catchments (Fig. 1) with varying characteristics (Table 1). These were selected due to the frequency and magnitude of the impacts of floods, as well as the availability of river flow observations (Table 1). Kenya mainly experiences a bimodal rainfall pattern, occurring in the seasons of March - April - May (MAM) and October -November - December (OND) (Ayugi et al., 2016; Yang et al., 2015), which are commonly known as the 'long' and 'short' rains respectively. The rainfall seasonality and the migration of the precipitation zone is mainly influenced by the north-south movement of the inter-tropical convergence zone (ITCZ) (Black et al., 2003; Ongoma et al., 2015). The rainfall season migrates northward at a slower rate than it migrates southward, hence the two different names – 'long rains' and 'short rains' respectively (Nyenzi, 1988). The rainfall exhibits high spatiotemporal and interannual variability (Ongoma and Chen, 2017) and is strongly influenced by perturbations in global Sea Surface Temperature (SSTs) especially in the Pacific and Indian Oceans with the El-Niño Southern Oscillation (ENSO) (Black et al., 2003; Ogallo, 1993) and the

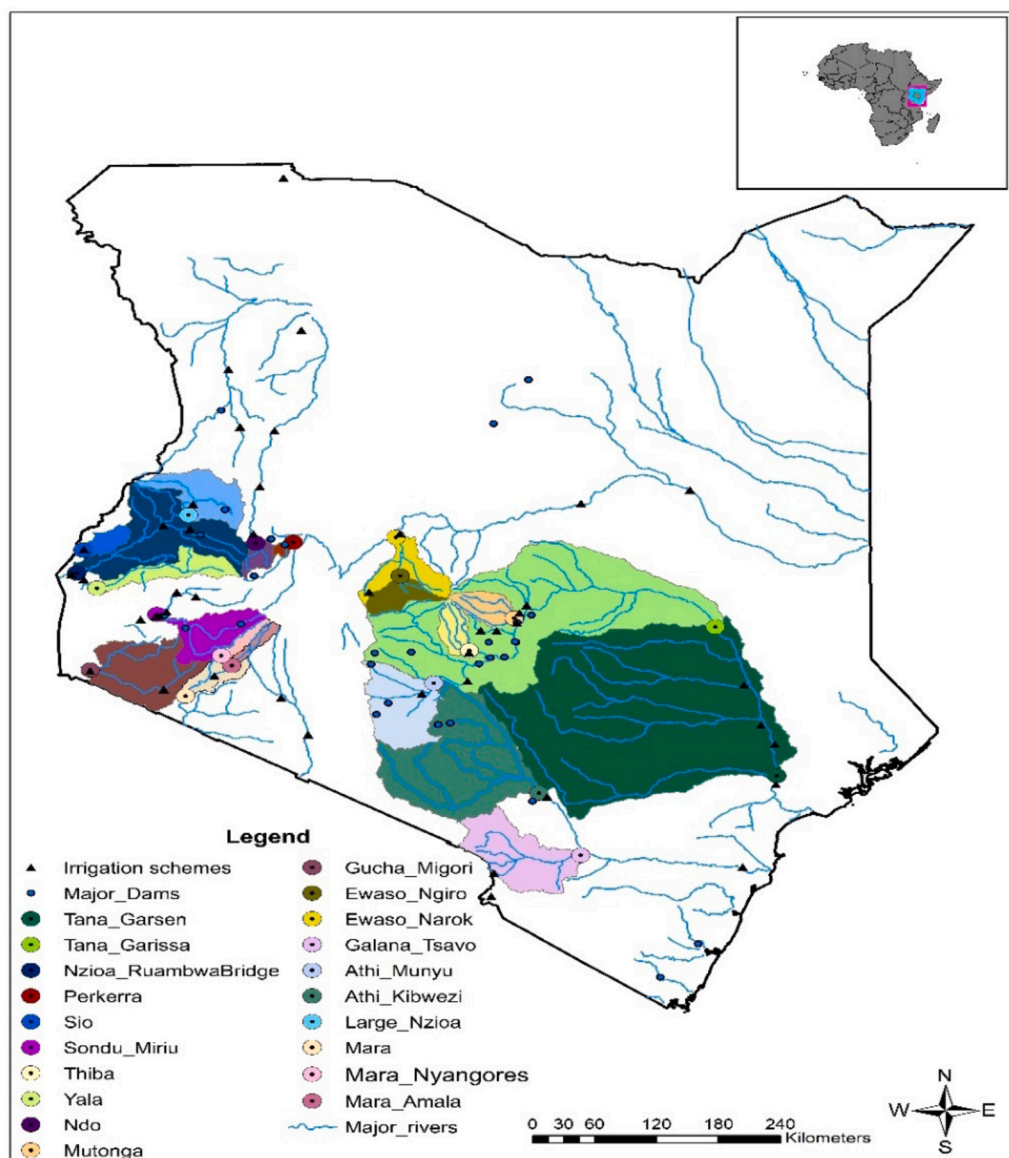**Fig. 1.** Study catchments across Kenya, with locations of the outlet river gauges (show by cirloured circled dots) used in this study and the main irrigation schemes (black trianges) and major dams (blue circles) provided by Kenya Water Authrority (WRA).

Table 1
Summary of the catchments considered, their characteristics, and the main human influences, including number of dams and water abstraction activities (Source: WRA-K).

River Name	Catchment Outlet point	Station ID	Lon	Lat	Drainage Area (km²)	Mean Elevation (m.a.s.l)	Mean Annual Rainfall (mm)	Annual Discharge (m³s⁻¹)	Catchment Characteristics	Human Influence Dams	Irrigation schemes	First & Last year of record	Record length (years)	Amount missing (%)
Tana	Tana Garsen	4G02	40.11	-2.28	80 760	720	672	135.8	Semi-arid plains	9	11	1981–2016	36	58.2
Tana	Tana Garissa	4G01	39.7	-0.45	32,695	870	868	169.3	Highlands on the upstream & semi-arid plains in the lowland	8	7	1981–2018	38	14.2
Nzioa	RuambwaBridge	1EF01	34.09	0.12	12,643	1740	1488	151.2	Dense forest cover (highlands) & low trees & bushes (lower reaches)	2	4	1981–2018	38	13.6
Galana	Galana Tsavo	3G02	38.47	-2.99	6560	930	628	3.3	Semi-arid savannah plains	3	1	1981–2015	35	59.6
Gucha	Gucha Migori	1KB05	34.21	-0.95	6 310	1650	1435	45.0	Eastern lowlands with dense vegetation cover	0	2	1981–2015	35	47.8
Athi	Athi Munyu	3DA02	37.19	-1.09	5689	1730	822	18.8	Highlands and forest cover	3	1	1981–2017	37	21.6
Nzioa	Large Nzioa	1BD02	35.06	0.76	3878	1720	1267	15.3	Dense forest cover	1	1	1981–2011	31	28.8
Sondu	Sondu Miriu	1JG04	34.80	-0.33	3444	2017	1614	53.9	Low lying plains (western) & highland (Eastern)	2	2	1981–2018	38	64.4
Mara	Mara	1LA04	35.04	-1.23	2977	2100	1262	11.8	Low lying shrubs, semi-arid	0	1	1981–2015	35	77.7
Yala	Yala	1FG02	34.27	0.04	2700		1696	40.8	Swampy	0	0	1981–2019	39	59.6
Ewaso	Ewaso Narok	5AC10	36.73	0.43	2597	1600	880	5.3	Low lying shrubs & mainly semi-arid	0	2	1981–2018	38	26.5
Tana	Mutonga	4EA07	37.89	-0.38	1867	1830	1427	35.5	Highlands and forest cover	0	1	1981–2016	36	44.2
Ewaso	Ewaso Ngiro	5BC04	36.91	0.09	1837	1700	972	20.6	Low lying shrubs & mainly semi-arid	0	0	1981–2019	39	35.0
Sio	Sio	2EE07A	34.14	0.39	1011	1650	1822	15.5	Low trees & bushes & swampy in lower reaches	0	1	1981–2018	38	18.1
Turkwel	Ndo	2C07	35.65	0.45	897	1133	1371	9.1	Extensive palaeo-floodplain & arid conditions	0	1	1981–1993	13	47.2
Mara	Amala	1LB02	35.44	-0.89	695	2100	1377	6.8	Low lying shrubs, semi-arid	0	0	1981–2017	37	25.6
Mara	Nyangores	1LA03	35.35	-0.79	692	2008	1262	11.8	Semi-arid savannah plains, low lying shrubs, semi-arid	0	0	1981–2017	37	15.5
Turkwel	Perkerra	2EE07A	35.97	0.46	371	1023	832	5.7	Extensive palaeo-floodplain and arid conditions	1	1	1985–2005	21	50.1

Table 2

Overview of the global reanalysis and the blended (Satellite and observation) Chirps precipitation dataset(s) used in the study.

Short Name	Full Name and details	Data sources (*)	Spatial resolution	Spatial coverage	Temporal coverage	Temporal resolution	Reference
CHIRPS V2.0	Climate Hazards group Infrared Precipitation (CHIRP) V2.0 (http://chg.ucsb.edu/data/chirps/)	G, S, R *	0.05°	Land, < 50	Daily	1981–NRT*	(Funk et al., 2015)
ERA5	European Centre for Medium-Range Weather Forecasts Reanalysis (https://www.ecmwf.int/en/research/climate-reanalysis/era-5)	R *	~ 0.31°	Global	Hourly	1979 – NRT*	(Hersbach, 2018)
ERA-Interim	European Centre for Medium-Range Weather Forecasts ReAnalysis Interim (https://www.ecmwf.int/en/research/climate-reanalysis/era-interim)	R *	~ 0.75°	Global	3-Hourly	1979–2019	(Dee et al., 2011)
JRA-55	Japanese 55-year Reanalysis (JRA-55) (https://jra.kishou.go.jp/JRA-55; https://data.diasjp.net/dl/storages/filelist/dataset:204)	R *	~ 0.56°	Global	3-Hourly	1951-NRT*	(Kobayashi et al., 2015)
NCEP-CFSR	National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR; http://cfs.ncep.noaa.gov/cfsr/)	R *	~ 0.38°	Global	Hourly	1979– 2010	(Saha et al., 2010)

(*) NRT= Near Real Time with a delay of several days, G = Gauge, S = Satellite, R = Reanalysis

Indian Ocean Dipole (IOD) (Blau et al., 2020; Owiti et al., 2008) being the most important modes. Other systems that influence rainfall variability include the high pressure systems (e.g. the Mascarene and the Arabian) (Ogwang et al., 2015), the Quasi-Biennial Oscillation (QBO) (Collier et al., 2016; Indeje and Semazzi, 2000), the Madden-Julian Oscillation (MJO) (Kilavi et al., 2018), Tropical cyclones (Finney et al., 2020; Wainwright et al., 2021) and jet streams, eg., the Turkana jet (Hartman, 2018; Kinuthia, 1992). The country has complex topography with the lowest altitudes along the coastline and Lake Victoria basin which are particularly prone to floods, while in the highlands, frequent thunderstorms and lightning threaten life.

3. Data and methodology

3.1. Datasets

3.1.1. Reanalysis and observational data

Four reanalysis products, namely ERA5, ERA-Interim (hereafter ERAI), Climate Forecast System Reanalysis (CFSR), and the Japanese 55-year Reanalysis (JRA55), and a gridded observational dataset, the Climate Hazards group Infrared Precipitation with Station (CHIRPS), were used in this study (see Table 2). We used the daily precipitation, maximum and minimum temperature variables from the reanalysis products for the study.

ERA5 is the latest global atmospheric reanalysis product from European Centre for Medium-Range Weather Forecasts (ECMWF) which spans the modern observing period from 1950 onward (Hersbach, 2018). In this study, 3-hourly ERA5 was obtained from ECMWF on a fixed grid of 0.31° × 0.31°. ERAI is the previous global reanalysis product created by ECMWF (Dee et al., 2011). Daily ERAI was obtained from ECMWF on a fixed grid of 0.75° × 0.75°. JRA55 is a global reanalysis dataset constructed by the Japan Meteorological Agency (JMA) (Kobayashi et al., 2015). Daily JRA55 was obtained from National Center for Atmospheric Research (NCAR) climate data guide at a fixed gird of 0.56° × 0.56°. CFSR is a global reanalysis dataset of atmosphere fields produced by the National Centers for Environmental Prediction and for Atmospheric Research (NCEP/NCAR) (Saha et al., 2010). The CHIRPS dataset was used as a benchmark observation dataset since it has been used in several studies showing good results compared to observations over eastern Africa (Dinku et al., 2018). CHIRPS is a quasi-global, high resolution, daily, pentad, and monthly precipitation dataset (Funk et al., 2015). Based on infrared Cold Cloud Duration (CCD) data, CHIRPS has a long enough history of precipitation data. The algorithm is based on (i) a 5 km climatology that uses satellite data to represent sparsely gauged locations, ii) includes daily, pentadal, and monthly 5 km CCD-based precipitation estimates from 1981 to the present, iii) combines station data to generate tentative information product with a latency of about 2 days and a final product with an average latency of about 3 weeks, and iv) interpolation weights are assigned based on a novel blending method which uses the spatial correlation structure of CCD estimates. This makes it comparatively an alternative in data scarce regions. We opted for the gridded observations as the daily observed gauge datasets were not available for the catchments of study and are known to be very sparse and present large data gaps (Dinku et. al, 2019; 2018; Le et. al, 2017).

3.1.2. Observed river discharge and potential evapotranspiration

River discharge datasets at daily time step for the period 1981–2016 were provided by the Kenya Water Resource Authority (WRA) for the selected catchments across the country, summarized in (Table 2). The potential evapotranspiration (PET) required for the catchment modelling was estimated from the average daily temperature of the four reanalysis products and CHIRTS_daily data from the Climate Hazard Centre (CHC). As temperature readings were the readily available meteorological data that relates to PET, for this study, temperature-based methods were used to estimate the PET (Hargreaves and Samani, 1985). For this study, the Hamon method

(Hamon, 1960) was used to estimate PET daily averages for different datasets.

3.2. Modelling experiment methodology

To obtain the monthly and annual totals for observations and reanalysis datasets, the daily values were accumulated. The seasonal total precipitation was calculated by summing monthly precipitation for three seasons: (i) the March-April-May, hereafter referred to as MAM, (ii) the June-July-August, hereafter JJA, and (iii) the October-November-December, hereafter OND. All datasets were converted to the same units for consistency (e.g., JRA55 and CFSR were converted from $kg/m^2/s$ to mm/d). ERA5, ERAI, JRA-55 and CFSR were regridded by first-order conservative interpolations to a horizontal grid of 0.5^o x 0.5^o (Schulzweida, 2019).

We first qualitatively evaluate the performance statistics of the reanalysis datasets in terms of temporal dynamics and biases with respect to precipitation observations (CHIRPS), considering the following metrics: Pearson Linear Correlation Coefficient (CC), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Error (ME), long-term relative bias (BIAS) and annual number of dry days calculated on monthly, annual, and seasonal scales. We produce spatial maps for the standardized precipitation anomalies, bias and annual number of dry days, to assess their consistency compared to observations, and tabulate the other statistics to show the aggregate performance across the different datasets.

Second, we calibrate the GR4J (Perrin et al., 2003) rainfall runoff model. In the study, we used five different inputs sources (for both precipitation and PET) into GR4J model from CHIRPS, ERA5, ERAI, CFSR and JRA55. We calibrate the model with each of the input source at a time and compute the KGE score and compare how this varies across the four different datasets relative to observations. The GR4J model is a simple daily lumped rainfall-runoff model belonging to the family of soil moisture accounting models. There are four main parameters (Fig. 2) to be calibrated in GR4J model, namely: (1) the maximum capacity of the production store (X1, mm), (2) the groundwater exchange coefficient (X2, mm), (3) the maximum capacity of the non-linear routing store (X3, mm), and (4) the time base of the unit hydrograph (X4, days). There are also a few fixed parameters, whose values were set by Perrin et al. (2003). All four free parameters are real numbers, X1 and X3 are positive, X4 is greater than 0.5 and X2 can be either positive, zero or negative. The typical inputs of GR4J are the areal precipitation depth (P, mm) and the potential evapotranspiration (PE, mm) estimate over the catchment. Most optimization algorithms used to calibrate the model parameters require knowledge of an initial parameter set. Given the small number of model parameters, simple optimization algorithms are generally capable of identifying parameter values yielding satisfactory results. The choice of an objective function depends on the objectives of model user. The choice and use of the GR4J model is mainly due to its simple and relatively quick to calibrate structure, ensuring high-levels of performance and robustness (Ficchì et al.,
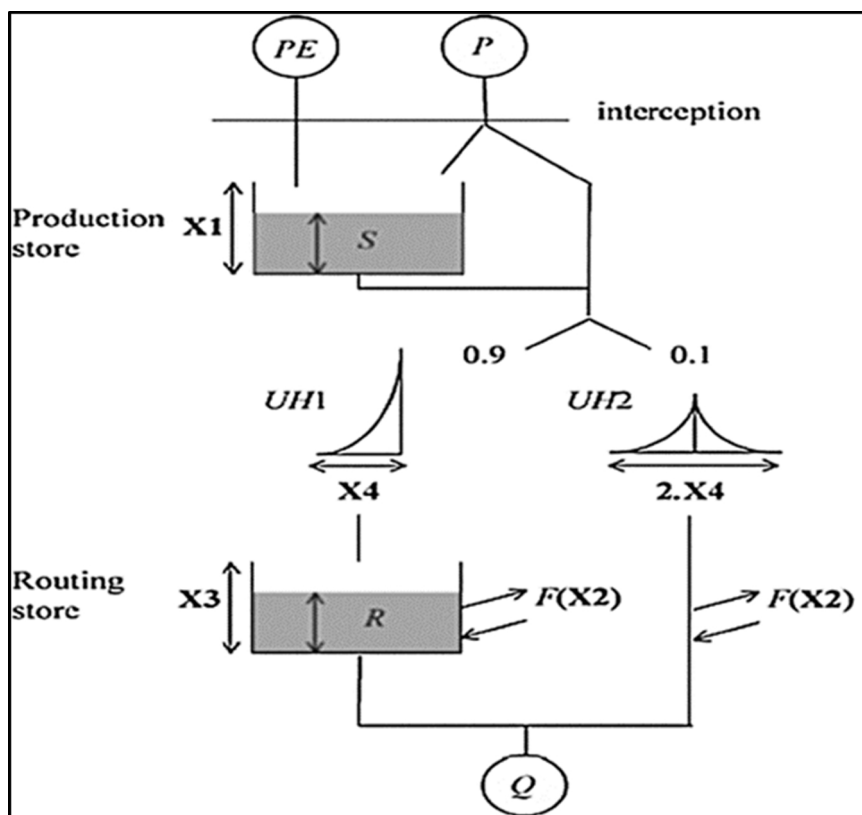

Fig. 2. The GR4J rainfall-runoff model (Source (Perrin et al., 2003)). P is the rainfall depth and E is the potential evapotranspiration (PE) averaged over the basin at a daily time step. X1, X2, X3 and X4 are the model parameters.

2019; Mostafaie et al., 2018; Oudin et al., 2004; Van Esse et al., 2013; Pushpalatha et al., 2011).

The four free parameters of the GR4J model were calibrated using the default optimisation algorithm provided in the airGR package (Coron et al., 2019; Delaigue et al., 2019). This simple optimization algorithm, mainly based on a local optimisation, proved to be equally efficient to locate a robust optimum compared to more complex global search algorithms (Coron et al., 2019) and proved effective in terms of the number of model runs required for convergence (Mathevet et al., 2006). The Michel method (Michel, 1983) is based on two steps:

I. A systematic inspection of the global parameter space is performed to determine the most likely zone of convergence. In our study, this is done by direct grid-screening.
II. A steepest descent local search procedure is carried out to find an estimate of the optimum parameter set starting from the best parameter set from step 1.

The four free model parameters were calibrated by applying the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009) as the objective function and the daily observed river discharge data of the selected catchments as reference. We use different inputs (precipitation datasets) from CHIRPS, ERA5, ERAI, CFSR and JRA55 to calibrate the GR4J model. The KGE was used also for evaluating the performance of the GR4J model when forced with different reanalysis data. The KGE objective function represents a weighting of three components that correspond to bias, correlation, and variability, ensuring that KGE is sensitive to errors in the overall distribution of streamflow (Adeyeri et al., 2020; Kling et al., 2012). We therefore calculated the hydrological model performance statistics for the calibration and validation periods and compared across the different reanalysis datasets to investigate the overall suitability of the different reanalysis as input data to simulate river flows. We adopted a threshold of model performance in the range $-0.41 < \text{KGE} \leq 1$ as reasonable, following Knoben et al., (2019) being -0.41 the KGE value corresponding to a mean flow benchmark.

A split-sample validation technique (Klemeš, 1986) was used to test model performance beyond the calibration period. For this study, 36 years (1981–2016) of streamflow data for each catchment were available, so we split into two equal 18-year Split-Sample Testing (SST) periods hereafter referred to as SST1 and SST2.

Third, we perform Sensitivity Analysis by applying the global Sobol sensitivity method for the GR4J model parameters using the KGE as our target function and the daily observed data of the 19 catchments as reference. We adopt the Sobol method because it estimates the relative contribution of individual model parameters and their interactions through the decomposition of model output variance (Nossent et al., 2011). A sensitivity analysis allows a reduction of the number of parameters incorporated in the optimization by determining the most influential parameters of a model and their identifiability (Saltelli et al., 2000). As no prior information is available on the parameters, the input parameter values for the Sensitivity Analysis are sampled from a uniform distribution (Nossent et al., 2011). The different parameter ranges are scaled between 0 and 1 with a linear transformation. Then, we obtain one value of the Sensitivity Indices (SI) per parameter, and we investigate the relative role of each parameter in explaining the output variance and assess possible over-parameterization issues by counting the number of sensitive parameters. The value of the objective function for the calibration of parameters can be used as the model performance statistics for sensitivity analysis, as such we adopted the KGE.

Last, we assess the overall suitability of the rainfall–runoff model when forced with different meteorological inputs by calculating the Model Suitability Index (MSI). We compare the performance of the four reanalysis datasets across the 19 catchments and investigate which of the input dataset is suitable and which require caution, because of low model performance and possible parameter identifiability or over-parameterization problems. The well-known problem of over-parameterisation due to insensitive parameters in models with large number of parameters (van Griensven et al., 2006) makes sensitivity and performance statistics important. This may result in uncertain model simulations arising from equifinality in model calibration but yielding unequifinal model simulations in
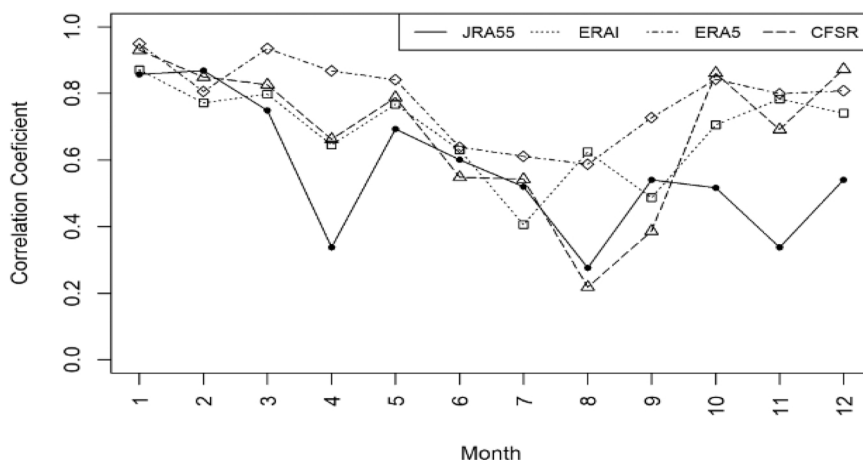

Fig. 3. Line graph of correlation coefficients (CC) between monthly observations and ERA5, ERAI, JRA55 and CFSR precipitation for the period 1981 – 2016 on average across the 19 study catchments.

validation (Beven, 2012). This mostly arises due to application of calibrated multiple optimal parameters sets with significantly variable parameter values (Shin et al., 2015; Shin and Kim, 2017). Therefore, in most cases the problem arising from prediction uncertainty may pose problems to modellers when it comes to decision making. By applying the quantitative method of Sobol's SA, this enabled us to couple the results with the performance statistics. The MSI aggregates both sensitivity indices & performance statistics (Shin and Kim, 2017), providing a clear index to judge the relative global performance of the reanalysis products with respect to observations. The computed MSI can be used in comparison studies with any catchment data. If all the model parameters are sensitive, this would yield a MSI of 1 and a perfectly matched hydrograph between the simulations and observations.

We adapt Shin' and Kim (2017) Model Suitability Index (MSI), which is a combined measure of performance statistics and Sensitivity Analysis results. The MSI can be expressed as:

$$MSI = 0.5 \times \left(\frac{1}{n} \sum_{i=1}^{n} SR_i \right) + 0.5 \times \left(\frac{1}{m} \sum_{j=1}^{m} PS_j \right) \qquad (1)$$

Where the SR is the sensitivity ratio (i.e., the ratio of the number of sensitive parameters out of the total number of model parameters) ranging from [0, 1] and PS is the performance statistics, n is the number of years over which the sensitivity analysis is run, and m is the number of split sample periods in model calibration. It is necessary to set a sensitivity threshold to ascertain the sensitive parameters, hence we adopted a minimum value of 0.2 for the TSI of a sensitive parameter. This value has been suggested and used in some past studies (e.g (van Werkhoven, 2009; Van Werkhoven et al., 2009 & Shin et al., 2013). It is worth noting that this is an arbitrary value, thus we acknowledge the need to practice caution when the parameters' TSI values are nearing the threshold. PS is computed by obtaining the average value of all the periods considered (i.e., two split sample periods). To calculate the average PS, we considered the calibration and validation performance statistics (KGE). As both measures are equally important, we gave the same averaged weight to PS and SR in calculating the MSI.

4. Results and discussion

4.1. Results

4.1.1. Overall performance evaluation using observations

The performance of ERA5, ERAI, JRA55 and CFSR on monthly, seasonal, and annual scales is presented in this section. We used the monthly scale as a base time scale and calculated CC, RMSE, MAE and ME for all the four reanalysis products.

4.1.1.1. Performance on monthly scale. ERA5, ERAI, JRA55 and CFSR were first evaluated on a monthly timescale with respect to observations at the country level. All the datasets passed the significance test of the correlation coefficient at the 99% confidence level and to eliminate the influence of the seasonal cycle on the values, each Correlation Coefficient was calculated per month as shown in Fig. 3. ERA5 shows the highest average correlation coefficient value of 0.71 on monthly timescale compared to observations and is consistently higher across all months (Fig. 3, and Table 3) than the other reanalysis products. ERAI and CFSR have good average correlation but show larger drops in some months (especially in the drier month of August). JRA55 obtained a poor correlation coefficient of 0.46 on average. In general, ERA5, ERAI and CFSR show higher correlations to observations in rainy months (March-April-May and October- November-December) and lower in the dry months (June-July- August), whereas JRA55 shows worst correlations during both rainy seasons.

The average twelve months evaluation indices for each of the reanalysis product is shown in Table 3. Overall, ERA5, ERAI and CFSR show a similar good ability to simulate the precipitation for all the indices under consideration. ERA5 has a better CC, BIAS and RMSE whereas JRA55 has the lowest CC and the largest BIAS and RMSE suggesting that JRA55 is worst performing reanalysis dataset over the Kenyan catchments.

4.1.1.2. Performance on seasonal and annual timescales. The overall performance of the four reanalyses (ERA5, ERAI, CFSR and JRA55) were evaluated on seasonal and annual timescales to explain the propagation of errors at these timescales. The results of the different performance indices are shown in Table 4.

The overall correlation coefficients on seasonal and annual timescale are shown in Fig. 4. Higher CC across all the datasets were obtained in the wet seasons of MAM and OND, whereas lower CC were obtained in the dry season of JJA, with the performance index

Table 3
Average CC, BIAS, RMSE, MAE and ME between the four reanalysis precipitation datasets and observations on monthly timescale for the period 1981 – 2016 over all the study catchments.

Index	ERA5	ERA-Interim	JRA55	CFSR
CC	0.71	0.63	0.46	0.68
BIAS (%)	49.72	-26.97	146.66	-76.68
RMSE (mm)	31.59	43.73	115.12	67.36
MAE (mm)	25.11	37.02	79.92	59.95
ME (mm)	1.64	-30.00	71.52	-59.91

Table 4

CC, BIAS, RMSE, MAE and ME between the reanalyses and observation precipitation data at a seasonal and annual timescale averaged over the 19 study catchments in Kenya.

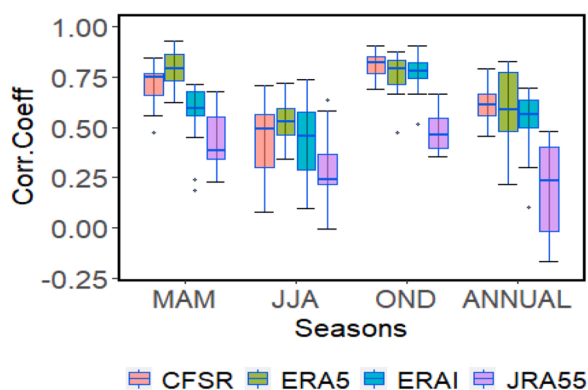| Season	Dataset	CC	BIAS (%)	RMSE (mm)	MAE (mm)	ME (mm)
MAM	JRA55	0.34	155	965.10	898.12	898.12
	ERAI	0.59	-28.1	193.87	168.09	-160.36
	ERA5	0.88	47.6	322.44	282.42	275.60
	CFSR	0.78	-84.6	481.03	475.14	-475.14
JJA	JRA55	0.52	67.4	361.72	290.77	267.76
	ERAI	0.24	-39.5	182.05	157.12	-154.93
	ERA5	0.25	-1.4	85.66	70.90	-5.446
	CFSR	0.22	-84.6	332.89	327.95	-327.95
OND	JRA55	0.44	271.7	1036.4	983.25	983.25
	ERAI	0.81	-27.9	123.34	111.09	-100.56
	ERA5	0.52	96.6	397.00	349.70	349.70
	CFSR	0.84	-84.3	307.76	296.47	-296.47
ANNUAL	JRA55	0.25	171	2902.21	2760.63	2760.63
	ERAI	0.46	-26.3	462.513	421.14	-421.14
	ERA5	0.52	44.7	801.75	728.57	720.70
	CFSR	0.60	-85.2	1354.36	1349.21	-1349.21

Fig. 4. Boxplots of the seasonal (MAM, JJA and OND) and annual Correlation Coefficients (CC) for four reanalysis CFSR (pink), ERA5(green), ERAI (blue) and JRA55 (purple) across the 19 catchments. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.
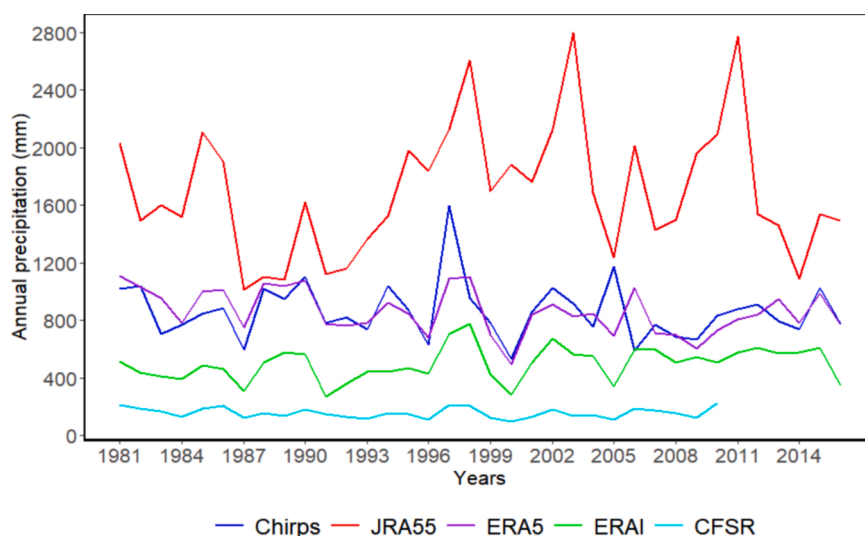

Fig. 5. Areal average annual precipitation from the observations and the reanalysis datasets, averaged across the 19 study catchments.
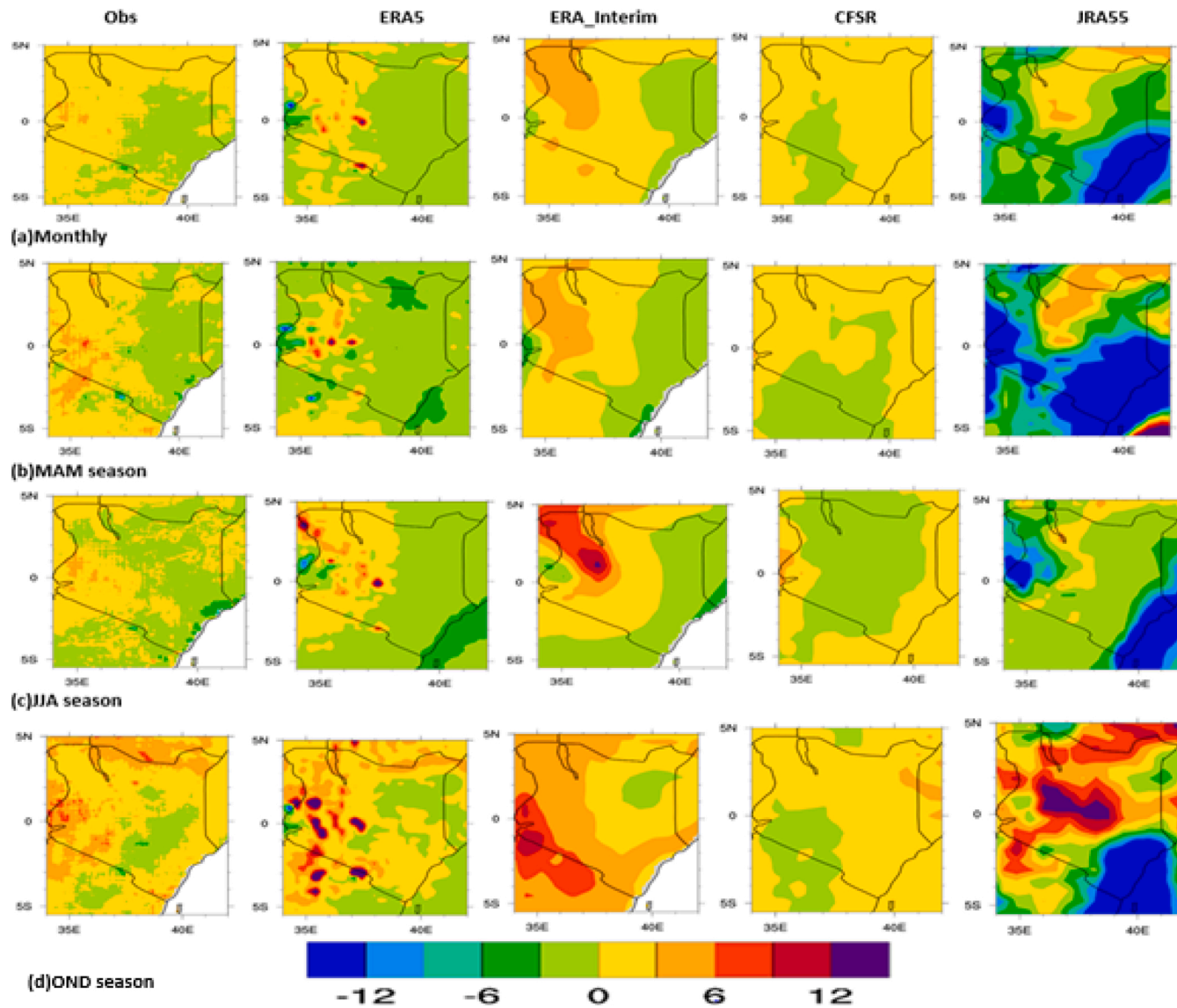
Fig. 6. Mean monthly and seasonal standardized precipitation anomalies in the four reanalysis products for 1981–2016 period: (a) Monthly anomalies, (b) MAM season, (c) JJA and (d) OND anomaly index in seasonal precipitation, for ERA5 (2nd column), ERA-I (3rd column), CFSR (4th column) and JRA-55 (5th column). Column 1 shows the observations (CHIRPS).
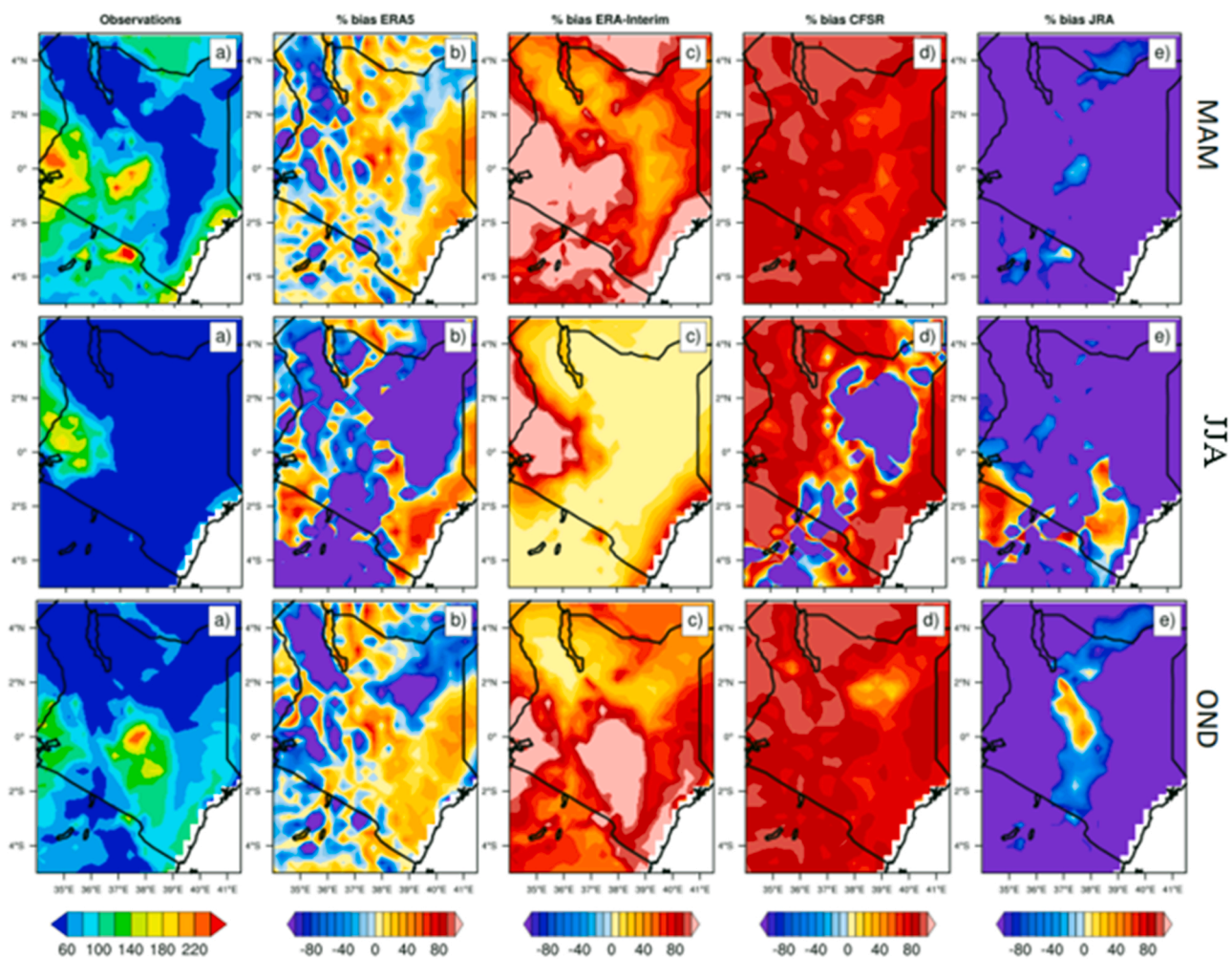
Fig. 7. Seasonal observed precipitation (mm) and mean bias (%) of the extreme rainy days at 95th percentile in the four reanalysis products for 1981–2016 period: (a) Long-term observed (OBS) seasonal average precipitation (mm) from CHIRPS, (b– e) mean relative bias (%), in seasonal precipitation in ERA5 (b), ERA-Interim (c), CFSR (d), JRA-55 (e), with respect to CHIRPS. MAM season (top), JJA (middle) and OND (bottom panel).
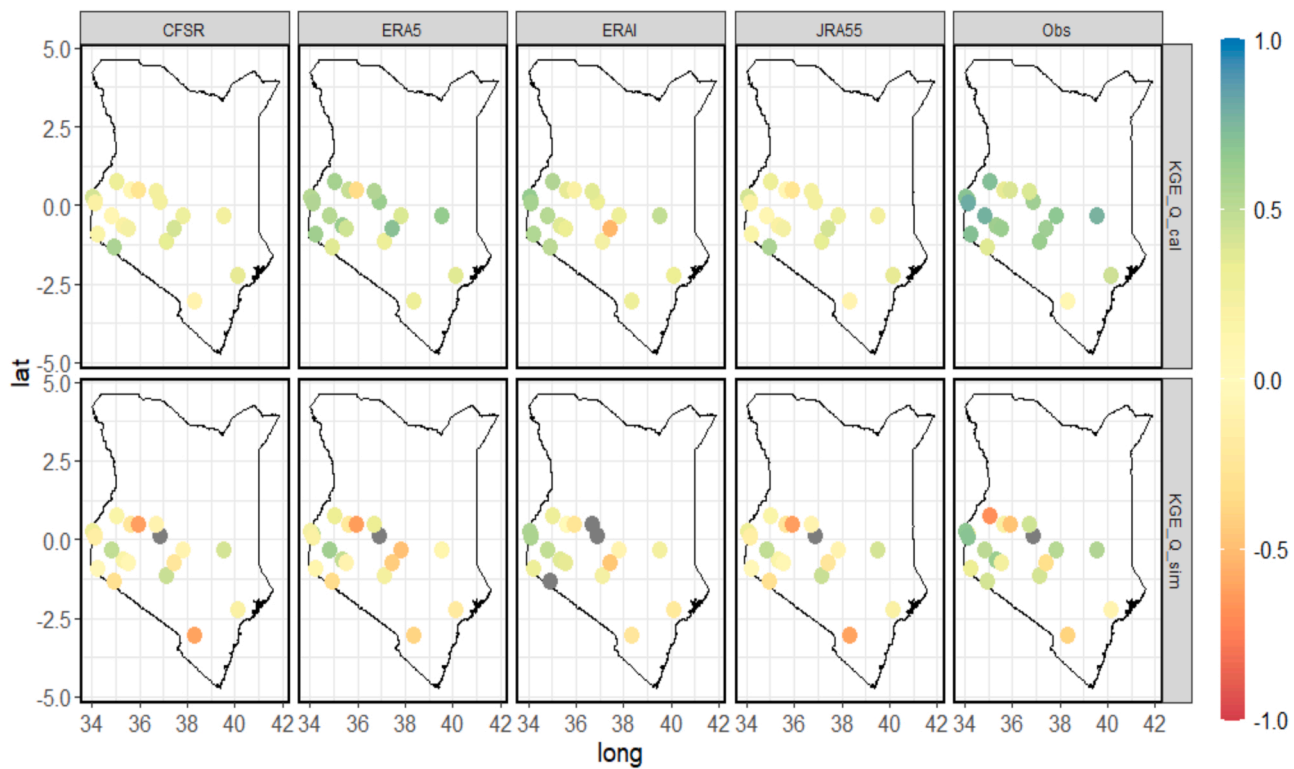
Fig. 8. GR4J model performance (KGE) in calibration (top panel) and validation (bottom panel) across the 19 catchments for different input datasets, Pan.1 (CFSR), Pan.2 (ERA5), Pan.3 (ERAI), Pan. 4 (JRA55), Pan.6 (CHIRPS).

higher in OND than in MAM. ERA5 obtained the highest CC (0.88) in MAM, whereas CFSR the highest (0.84) in OND. JRA55 showed lower CC of 0.34 and 0.44 in the two seasons respectively and a CC of 0.52 in the dry season, depicting a tendency of a wet bias over the dry months. On average, the variability in the CC index across the four datasets was relatively lower in the OND season and higher in the MAM season. The BIAS across the four datasets was lower in the dry season (JJA) and higher in the wet seasons (MAM &OND) with JRA55 showing a higher positive BIAS across all the seasons. There are large values in the RMSE and the MAE across the four datasets in the two wet seasons and this may be linked to the high precipitation concentrations during those seasons across most of the catchments. Generally, it can be noted that JRA55 shows the worst performance in comparison to observation especially in the wet seasons of MAM and OND but obtained relatively better scores in the dry season of JJA. ERA5 shows better agreement with observations across the three seasons thus may be an appropriate option for simulating precipitation over the Kenyan catchments.

On annual timescale, the average annual precipitation of CFSR, ERA5, ERAI and JRA55 was computed and compared with the observation (CHIRPS) (Fig. 5). ERA5, ERAI and CFSR show similar trend compared to observations across all the years, with CFSR and ERAI underestimating the precipitation. JRA55 shows a higher tendency of overestimating the annual precipitation over the study catchments. In terms of the performance indices, CFSR, ERAI and ERA5 showed better CC indices of 0.60, 0.46, 0.52 respectively whereas JRA obtained lower CC of 0.25 (Fig. 4). The variability in the CC was higher in JRA55 (Fig. 4). ERA5 and JRA55 show a positive bias of 45% and 171% respectively, whereas ERAI and CFSR show negative bias of − 26 and − 85%. ERA5 has a lower RMSE and ME whereas JRA55 has the highest. These results show that ERA5 is the best performing reanalysis dataset compared to observations on annual timescales whereas JRA55 is the worst performing.

The mean monthly and seasonal standardized precipitation anomalies in the four-reanalysis precipitation for a base climatological period 1981–2016 is shown in Fig. 6. On monthly timescale, the observations show a positive anomaly over the central highland and the western parts of Kenya (Fig. 6, pan1). The arid and semi-arid parts in the eastern and coastal lowlands show a negative anomaly (dry bias). This pattern is also captured in ERA5, ERAI and JRA55 although JRA55 has too high and widespread negative anomalies compared to the former two. On seasonal timescales, ERA5, ERAI and CFSR show positive anomalies in Western and central highland in the three seasons except for JRA55 which has a stronger negative and positive anomaly in MAM and OND seasons respectively.

An evaluation of the extreme precipitation in the four reanalysis was also performed (Fig. 7). For this case, we focused on the 95th percentile of rainy days for MAM, JJA and OND season during the period 1981–2016. A rainy day represents a day for which the recorded precipitation amount is greater than or equal to 1 mm (Gudoshava et al., 2020). The observed extreme precipitation varied between 60 mm to more than 240 mm across western, central highlands and coastal catchments for the rainy seasons (MAM and OND). The observed extreme precipitation during the dry season JJA varied between 100 mm and 160 mm across the western catchments
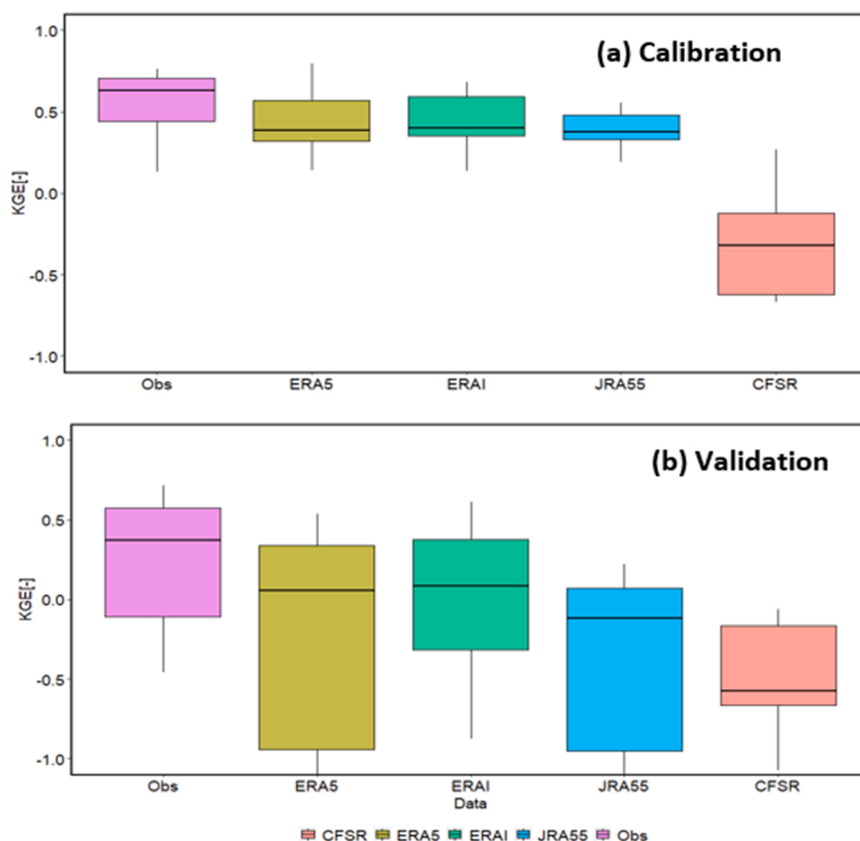

Fig. 9. Boxplots of the overall GR4J model performance (KGE [-]) in (a) calibration and (b) validation mode over the 19 catchments.
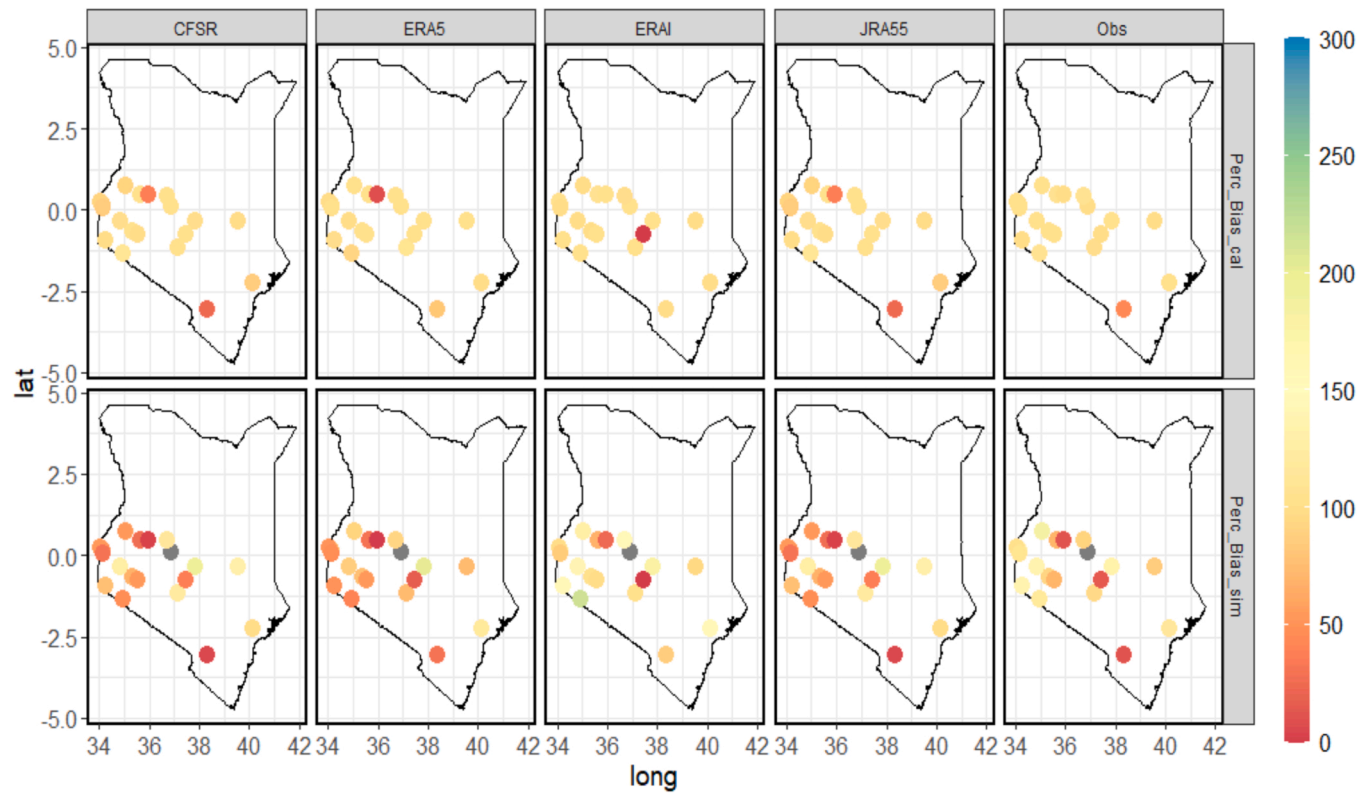
Fig. 10. GR4J model performance (Percentage KGE- Bias) in calibration (top panel) and validation (bottom panel) across the 19 catchments.
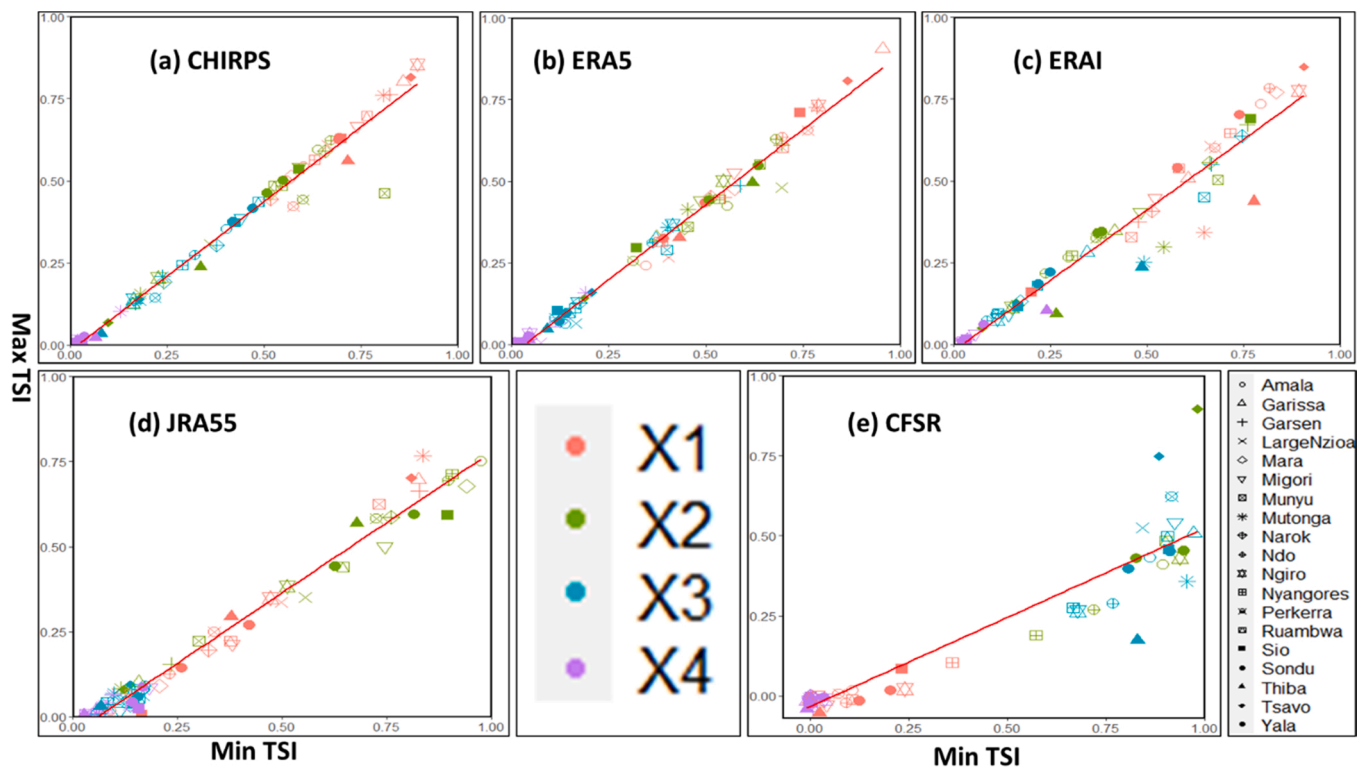
Fig. 11. Scatter plot of Sobol Total Sensitivity indices (TSI) for the different reanalysis datasets and the GR4J model parameters for the nineteen catchments. Minimum and maximum TSI were calculated for the whole data period. (a) Chirps, (b) ERA5, (c) ERAI, (d) JRA55 and (e) CFSR. Diagonal line is one-to-one line.

only whereas in the rest of the other regions the observed precipitation was less than 60 mm. Our results show that CFSR and ERAI show a positive bias for the extreme precipitation across most parts of the country in all the three seasons, like results in (Garibay et al., 2021). JRA55 has an enhanced negative bias of the extreme precipitation in most parts of the country except for isolated positive bias in the central highlands' region in JJA and OND season. ERA5 has a positive bias in MAM and OND in most parts of the country with some patches of negative bias in the western and central highlands catchments. It has an enhanced negative bias in JJA season with a positive bias in the western and coastal strip. We conclude that ERA5 outperforms other reanalysis products as it captures the wet extremes over the regions in which observations show enhanced precipitation in the respective seasons. The results are consistent with the findings in Gleixner et al. (2020) which showed both ERA5 and ERAI to have the capability to capture wet extremes in the dry seasons with ERA5 matching more closely to observations than the too wet ERA-interim. A promising performance in the ERA5 to simulate wet extremes can be attributed to improved bias correction method which incorporates aircraft measurements, satellite radiances, radiosonde measurements and surface pressure (Probst and Mauser, 2022). In addition, better performance in the central highlands can be attributed to the improved horizontal resolution in ERA5, which results in better estimates in orographic precipitation.

4.1.2. Evaluation of the reanalyses as inputs for hydrological modelling

4.1.2.1. Assessment of the overall model performance using different reanalysis. The performance of the four reanalysis datasets were evaluated using the GR4J model in the 19 catchments for the period spanning $1981 - 2016$. The KGE in calibration (top panel) and validation (bottom panel) scores obtained using different datasets for each of the catchments are represented in Fig. 8. Overall, wetland catchments in the western and highlands of Kenya obtained relatively better calibration scores compared to those in the semi-arid regions, with Yala, Sio, Nzioa and Gucha (wetland catchments) performing best and Perkerra, Ndo, Tsavo, Thiba and Tana (semi-arid catchments) performing worst. For each of the catchments, ERA5 showed better calibrated KGE scores compared to observations while CFSR and JRA55 obtained poorer KGE scores. However, we take caution in the interpretation of our results in terms of performance criteria because these catchments have a high influence of human activities such as irrigation schemes and dams. As such the low performance in some catchments may not be solely due to uncertainty in the input data.

The overall variability in GR4J model KGE scores across the four reanalyses are shown in Figure10. There are overall high-performance scores (KGE>0.5) in calibration mode in about half of the catchments for all datasets except CFSR, which suggests problems in using CFSR as hydrological model inputs in the region that cannot be solved or compensated by calibration. In Fig. 9a, ERA5, ERAI and JRA55 show similar overall performance compared to observation. The range for the performance statistic is narrower in the ERA5, ERAI and JRA55, indicating a more stable model performance in the region, while is wider in the CFSR data (Fig. 9a). In validation mode, the performance markedly decreases, as expected, for all datasets (Fig. 9b): ERAI, ERA5 and JRA55 have the highest median KGE value (just above or about 0) whereas CFSR has the lowest median values (KGE<−0.5). The range of KGE values is relatively larger compared to observations; thus, a relatively unstable prediction ability is expected for streamflow in reanalysis in the region. The range of performances is more variable in ERA5 and JRA55 and less variable in ERAI. Overall, the variability in KGE values is highest in validation than in calibration across all the reanalysis compared to observations, as expected. Fig. 10 shows the percentage
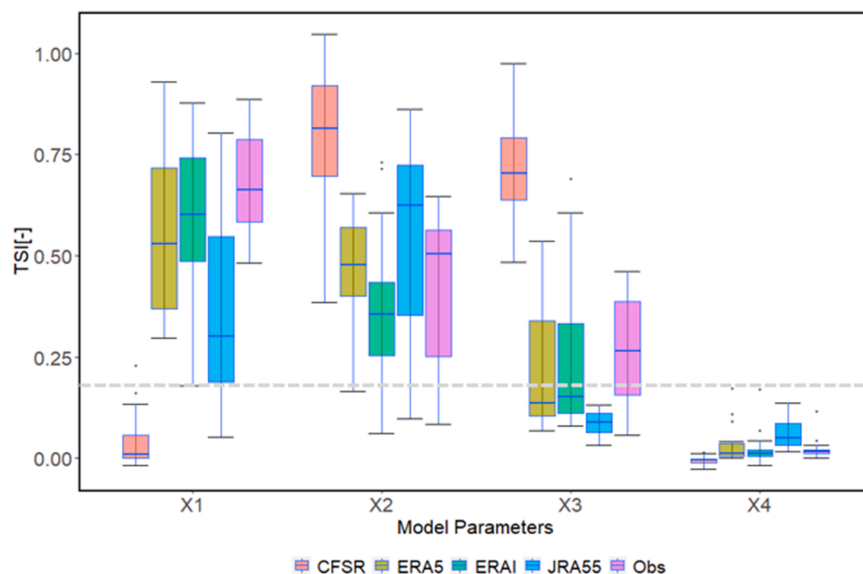

Fig. 12. Boxplots of the Sobol Total Sensitivity Indices (TSI) for the GR4J parameters for Obs. (pink), CFSR (orange), JRA55(blue), ERAI (green) and ERA5(forest-green) over the nineteen catchments. Dashed grey line represents the sensitivity threshold. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.

bias of the KGE component in each catchment in calibration (top panel) and in validation (bottom panel). The bias in all the four reanalysis is higher in calibration, whereas in validation most catchments exhibit lower biases except for Perkerra.

4.1.3. Sensitivity analysis results

4.1.3.1. Variability in sensitivity of model parameters. The GR4J's model maximum and minimum TSIs for the four reanalysis datasets is illustrated in Fig. 11. The maximum and minimum TSIs represent the variability of parameter sensitivity within the catchment with respect to KGE over the sampling periods and the variation across the four-reanalysis relative to observations. If the maximum and minimum TSIs for a parameter are equal (on the one-to-one line), that parameter has the same TSI for the sampling period, implying that the parameter is more stable across time, and would be expected to vary depending on the catchment characteristics and input data as well. In all the four datasets, the routing parameter (X4) related to the unit hydrograph is evidently the least sensitive as it is way below the threshold, followed by the capacity of the routing store (X3), whereas the two parameters governing the water-balance, i.e. the soil moisture accounting store (X1) and the groundwater exchange coefficient (X2), are the most sensitive across the datasets in most of the catchments, except for CFSR, where X1 is less sensitive.

Observations show more stability for the parameters for all catchments except six (Munyu, Thiba, Ndo, Ewaso Ngiro, Perkerra and Tsavo) with respect to reanalysis datasets (Fig. 11). In ERA5, most of the catchments showed stability in parameters except in Thiba, Tsavo, Large Nzioa and Ewaso Ngiro catchments (Fig. 11b). In ERAI, there is high variability in model parameter stability with less stability for some catchments such as Thiba, Munyu, Mutonga and Tsavo catchments (Fig. 11c). In JRA55 and CFSR (Figs. 11d and 11e respectively), the departure in sensitivity of model parameters from the diagonal is pronounced across most of the catchments. Overall, the variability in sensitivity of model parameters is high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, thus we can conclude that the reanalysis datasets are not suitable for model calibration is these catchments of Kenya characterized by arid and semi- arid conditions. However, the catchment's water balance may highly be affected by the dams constructed in the upstream areas and the massive irrigation schemes, which results in water attenuation.

4.1.3.2. Overall sensitivity of GR4J model parameters. The parameters related to water balance, i.e., the soil moisture accounting store (X1) and ground water exchange (X2), show higher sensitivity across all the four datasets except for CFSR, in which the production accounting store is less sensitive and falls below the threshold value of 0.2 from the model TSIs (Fig. 12). The first parameter responsible for water routing (X3) is less sensitive for most datasets (except CFSR), whereas the unit hydrograph parameter (X4) is the least sensitive across all the catchments in all datasets. In comparison to observations, ERA5, ERAI and JRA55 show similar parameter sensitivities of model parameters while CFSR show distinct higher variability and a difference in the parameters' sensitivity, which points to high uncertainty in the CFSR dataset. This result shows that the sensitivity of the model parameters can change with the input datasets, having very different hydrological characteristics.

4.1.4. Comparison of reanalysis datasets using Model Suitability Index

When the sensitivity indices and performance statistics are considered, it is difficult to determine which dataset is more
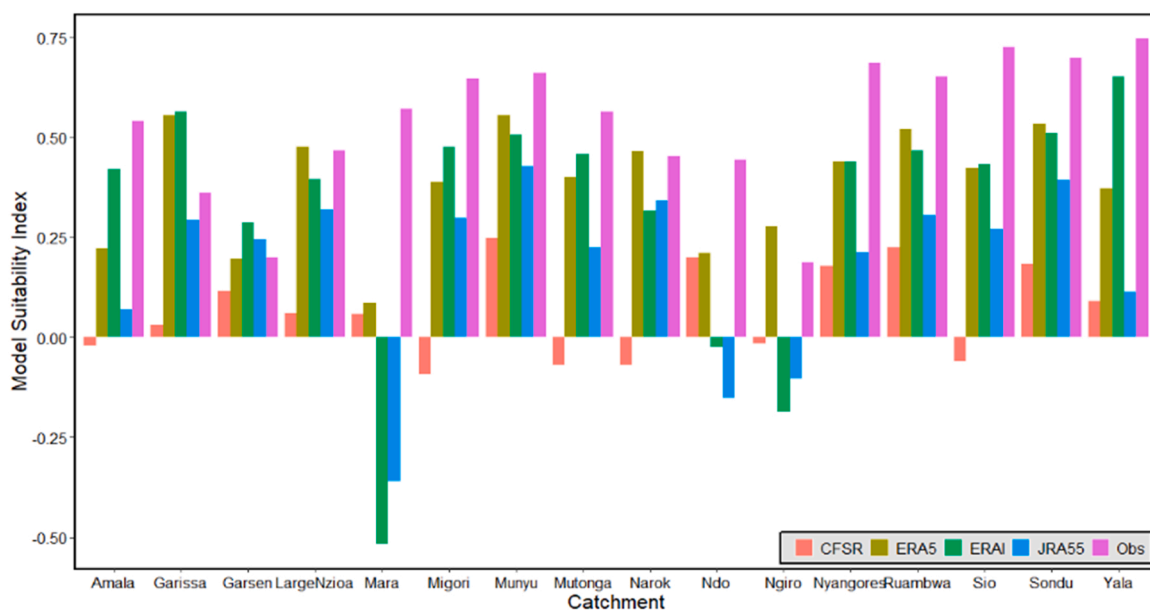

Fig. 13. Bar chart showing a comparison of model suitability in terms of performance and parameter sensitivity across different reanalysis using the Model Suitability Index (MSI).

appropriate. ERA5 and ERAI datasets, for example, had good and clear parameter sensitivities that captured their purposes, and the model performance score median values were higher than in CFSR and JRA55. However, the range of the performance statistics across the catchments was sometimes wider than in the other datasets, resulting in higher simulation uncertainty. When compared to the other two methods, the MSI, which considers both sensitivity indices and performance statistics, has the advantage of being easier and clearer to judge the superiority and inferiority of the datasets in terms of both model performance and parameter identifiability. We determined that a value of 0.5 for the MSI is a good MSI threshold, (Moriasi et al., 2007) We give the same weight to model performance and sensitivity, as described in the subsection 'MSI'; thus, the threshold value for good MSI is 0.5.

Combining the model performance and sensitivity indices discussed in the preceding subsections, Fig. 13 shows the MSIs for all the reanalysis datasets. The ERA5 has the highest MSI compared to observations across the nineteen catchments, followed by the ERAI reanalysis. As a result, the ERA5 and ERAI reanalysis are appropriate, at least for the selected sample of Kenyan catchments, whereas CFSR and JRA55 are least appropriate as they show lower MSI values across most of the catchments. CFSR shows negative MSI values for Amala, Migori, Mutonga, Narok, Ewaso Ngiro and Sio catchments meaning it is not appropriate for application in these Kenyan catchments. Overall, the four reanalysis datasets obtained relatively lower MSI values in Mara, Ndo, Ewasi Ngiro and Tana Garsen catchments. These catchments are mainly in arid and semi-arid areas of Kenya, like results in Section 5.2.

4.2. Discussion

4.2.1. Overall Performance of reanalysis precipitation products

In this study, we assessed four reanalysis precipitation products relative to observations for the period 1981–2016 on monthly, seasonal, and annual timescales. We also assessed how best they simulate streamflow using the GR4J model and sensitivity analysis for 19 catchments located in distinct geographical and climatic environments. Results show that the ERA5 reanalysis outperforms the other reanalysis products on monthly and seasonal scales, whereas CFSR outperforms ERA5 on annual and seasonal timescales. In general, ERA5 data were often closer to observations than other reanalysis data, which corresponds with earlier research on the datasets in different regions (e.g., Betts et al., 2019; Gleixner et al., 2020; Tarek et al., 2020), even though these studies considered different evaluation period, spatiotemporal resolution, hydrologic models, climates. However, the performance scores for the reanalysis products over the Kenyan catchments were lower which contrasts some of the studies carried out in other parts of the world with varying climates (e.g., Dhanya et al., 2017; Harada et al., 2016; Mahto et al., 2019; Wang et al., 2019), which obtained higher scores for their study areas. The low performance scores may be due to variations in the initial resolution of the datasets (Chen et al., 2018; Lemma et al., 2019) and the interpolation approach is likely to have some influence on the evaluation of various reanalysis data (Rapaić et al., 2015; Zhang et al., 2016). It is also worth noting that while the observed precipitation data are the best estimates available, they are likely to be subject to errors too (Beck at al., 2017a, Dinku et al., 2019). In addition, the seasonality of rainfall over Kenya is greatly influenced by weather phenomena such as El Niño -Southern Oscillation (ENSO) and Indian Ocean dipole (IOD) (Ayugi et al., 2020; Ojara et al., 2021; Onyutha, 2016) and play a major role in extreme rainfall events and inter-annual variability (Ongoma et al., 2015). For example, the warm phase of ENSO/El Niño results in unusually heavy rainfall, causing rare floods like as the 1997/1998 occurrence (Takaoka, 2005). ERA5, ERAI and JRA55 picked the enhanced annual precipitation totals of the strong El Niño years such as 1997/98 and 2015. However, relative to observations, ERA5 and ERAI underestimated the rainfall, and this may be attributed to incorrect configuration in the reanalysis products. For example, ERA5 precipitation is not customized to pick up the perturbations caused by the changes in the ocean – atmosphere interactions and the mountainous regions and so may miss picking up the extremes caused by events such as ENSO, thus the low performance scores.

Standardized precipitation anomalies in ERA5, ERAI and CFSR show a positive anomaly over the central highland and the western parts of Kenya (Fig. 6a, pan1) and a negative anomaly in arid and semi-arid parts in eastern and coastal lowlands in the three seasons (MAM, OND, JJA) except for JRA55 which has a stronger negative and positive anomaly in MAM and OND seasons respectively. This is consistent with study by Ongoma et al. (2018), which indicates a rise in the severity of severe precipitation events shown by a positive standardized rainfall anomaly over East Africa including the mentioned regions in Kenya. With the changing climate, temperatures in the region are projected to rise by the end of the twenty-first century, leading to an increase in rainfall extreme occurrences (Ongoma and Chen, 2017), thus, exacerbating flood risk.

Our analysis of the accuracy of precipitation reanalysis with respect to observations, over different timescales from monthly to annual, showed a positive but relatively small bias in CFSR, ERA5 and ERAI and a larger negative bias in JRA55 in MAM and OND seasons. Moreover, the first three reanalysis datasets showed a good average correlation at the monthly and seasonal scales. Therefore, the three reanalysis products have the potential to capture the rainfall seasonality and events in the study area. Recent worldwide research show that the frequency, severity, geographical range, length, and timing of climatic severe events are changing (Wainwright at al, 2021). A rise in rainfall severe events such as very wet days (R95p) and very wet days (R99p) anticipated for the future (2021–2100) (Gudoshava et al., 2020), is likely to cause the loss of life and property devastation owing to an increase in flood intensity (Finney, 2020). Therefore, further work should assess the capacity of the reanalysis datasets in capturing extreme rainfall event characteristics, such as timing and daily peaks. Performance of reanalysis as inputs into a hydrological model.

Using a hydrological model as integrator to compare simulated and observed streamflow, which can operate as an independent validation variable, is one approach to assess the quality of observation and reanalysis precipitation data. Each of the reanalysis precipitation and estimated potential evapotranspiration was supplied to the GR4J model, which were subsequently calibrated for each combination (using consistently precipitation and potential evapotranspiration from the same dataset), to independently analyze the quality of input data for each dataset relative to the observed streamflow gauge data. Streamflow gauges, of course, are subject to a variety of inaccuracies (Baldassarre and Montanari, 2009), but they represent the best available estimates for this study. Results of KGE

scores show that ERA5 is better than ERAI, JRA55 and CFSR but on overage, all the reanalyses are less skillful relative to observations across the catchments in this study and this is entirely due to the precipitation data quality. However, there is a marked improvement in the KGE scores for the catchments in the central highlands and western wet catchments which agrees with some studies on the datasets in other regions (e.g., Tarek et al., 2020; Essou et al., 2017; Lakew et al., 2020), pointing to the fact that reanalysis data can be used as a replacement for observations.

There are overall high-performance scores (KGE>0.5) in calibration mode in about half of the catchments for all datasets except CFSR, which suggests problems in using CFSR to reproduce the hydrological water balance in the region that cannot be solved or compensated by calibration (Diro et al., 2009).

4.2.2. Sensitivity analysis of model parameters

Sensitivity analysis is useful in identification of the parameters that have a strong impact on the model outputs, which in turn influences the effectiveness of the model. The greater the sensitivity of the model response to a parameter, the closer and sooner will that parameter be optimized so a high sensitivity is good. Such an in-depth analysis of a hydrological model may (i) help to identify any potential deficiencies in model structure and formulation; (ii) provide guidance for model parameterization; and (iii) provide the information content of available input data.

Based on provision of information content of the input data, different reanalyses show different sensitivities of model parameters and one that provides a higher sensitivity of model response means that it has less uncertainty and may be a lot easier to parameterize the values, but then this in practical sense does not reach the real value and the dataset with low sensitivity has got high uncertainty and model parameterization may be a lot difficult (Zeng et al., 2019). In comparison to observations, ERA5, ERAI and JRA55 show similar parameter sensitivities of model parameters while CFSR show distinct higher variability and a difference in the sensitive parameters, which points to high uncertainty in the CFSR dataset. This result shows that the sensitivity of the model parameters can change with the input datasets, having very different hydrological characteristics. Overall, the variability in sensitivity of model parameters is high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, thus we can conclude that the reanalysis datasets are not suitable for model calibration is these catchments of Kenya characterized by arid and semi- arid conditions. MSI considers both model performance and uncertainty quantitatively, therefore it can be used to compare any catchment. The ERA5 has the highest MSI compared to observations across the nineteen catchments, followed by the ERAI and JRA55, whereas CFSR has least MSI values. MSI's dependability may be increased by including more sensitivity indices and performance scores as well as assigning weights to the scores.

5. Summary and conclusion

This study addresses a notable gap that was found in the literature for evaluating the accuracy of multiple precipitation reanalysis datasets across data-scarce regions like Kenya, and for assessing their potential to supplement scarce rain gauge observations for hydrological modelling. Four different state-of-the-art reanalysis datasets were assessed. Precipitation data from ERA5 shows the highest average correlation coefficient value (0.71) on monthly timescale compared to observations and is consistently higher across all months than the other reanalysis. ERAI and CFSR have good average correlation but show larger drops in some months (especially in the drier month of August). JRA55 obtained a poor correlation coefficient of 0.46 on average. ERA5, ERAI and CFSR show higher correlations to observations in rainy months (March-April-May and October- November-December) and lower in the dry months (June-July- August), whereas JRA55 shows worst correlations during both rainy seasons. On annual timescales, CFSR, ERAI and ERA5 showed better CC indices of 0.60, 0.46, 0.52 respectively whereas JRA obtained lower CC of 0.25. ERA5 and JRA55 show a positive bias of 45% and 171% respectively, whereas ERAI and CFSR show negative bias of − 26 and − 85%.

Spatial rainfall patterns directly affect temporal distribution key in driving runoff and soil erosion processes, which is useful in management of hydrological risks and generation of sediments from rainwater. Monthly standardised anomaly maps in ERA5, ERAI and JRA55 showed a positive anomaly over the central highland and western parts. In the arid and semi-arid parts in the eastern and coastal lowlands parts of Kenya the three datasets showed enhanced negative anomaly. On seasonal timescales, ERA5, ERAI and CFSR show positive anomalies in Western and central highland in the three seasons except for JRA55 which has a stronger negative and positive anomaly in MAM and OND seasons respectively. Extreme precipitation showed a positive bias in CFSR, ERA5 and ERAI in MAM and OND seasons whereas JRA55 has enhanced negative bias in most parts of the country except for isolated positive bias in the central highlands' region in JJA and OND seasons.

The performance of GR4J model when forced with different reanalysis in the 19 catchments reveals a bigger role of localized catchment characteristics and process in model calibration. Wetland catchments in the western and highlands of Kenya obtained relatively better calibration scores compared to those in the semi-arid regions, with Yala, Sio, Nzioa and Gucha (wetland catchments) performing best and Perkerra, Ndo, Tsavo, Thiba and Tana (semi-arid catchments) performing worst. For each of the catchments, ERA5 showed better calibrated KGE scores compared to observations while CFSR and JRA55 obtained poorer KGE scores. The range of KGE values was relatively larger compared to observations; thus, a relatively unstable prediction ability is expected for streamflow in reanalysis for Kenyan catchments. The range of performances is more variable in ERA5 and JRA55 and less variable in ERAI. Overall, the variability in KGE values is highest in validation than in calibration across all the reanalysis compared to observations, as expected.

Sensitivity analysis allows the reduction of parameters incorporated in optimization by determining the convergence most influential model parameters. Sensitivity analysis revealed that in all the four datasets, the routing parameter (X4) related to the unit hydrograph was evidently the least sensitive, followed by the capacity of the routing store (X3), whereas the two parameters governing the water-balance, i.e. the soil moisture accounting store (X1) and the groundwater exchange coefficient (X2), are the most sensitive across the datasets in most of the catchments, except for CFSR, where X1 was less sensitive, with ERA5 showing a highest sensitivity in

the model parameters. However, the variability in sensitivity of model parameters was high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, thus we conclude that model calibration in arid and semi- arid catchments of Kenya does not yield skillful results using the reanalysis data. The MSI aggregates both sensitivity indices & performance statistics, providing a clear index to judge the superiority (or inferiority) of a reanalysis with respect to observations. On average ERA5, ERAI (& JRA55) have better MSI scores across most of the Kenyan catchments: ERAI & ERA5 perform better than JRA55 & CFSR, and lead to more robust model parameters. Using a catchment model and combined sensitivity - model performance analysis, allows an evaluation of the impact of the variability in the rainfall products throughout the catchment modelling process.

In conclusion, in this study we have demonstrated the usefulness of reanalysis rainfall products as potential alternatives for hydrological applications in Kenya. We assessed the suitability of reanalysis precipitation datasets for hydrological modelling across Kenyan catchments, but first assessed the propagation of errors when reanalysis is compared to observations. We performed the assessment on monthly, seasonal, and annual timescales. Then, using a lumped bucket-style hydrological model, we assessed the model performance via the KGE criterion and parameter uncertainty via Sobol's Sensitivity Analysis for four different reanalyses: - ERA5, ERAI, JRA55 and CFSR across 19 catchments. The parametric and model input uncertainty is investigated using the sensitivity indices and the comprehensive model performance analysis is used to examine the model's input strength, i.e., the amount to which the model captures the dynamics of rainfall–runoff processes with respect to different forcing. We also coupled the results of the performance scores and sensitivity indices to compute MSI for the 19 catchments.

We acknowledge the value and need of additional work, if reliable data at higher temporal frequency becomes available and can be used, as it contains more information. However, this is a big limitation for the current study due to high data gaps in the daily data (river discharge data used in the current study) and the lack of higher temporal resolution hydrological data. Future work should concentrate on assessing the sub-daily performance of hydrological modelling with reanalysis, testing its quality on other additional catchments in countries in the region with quality observed gauge data, but prior investments in data collection in Kenya seem to be needed. Our approach may be extended to various conceptual rainfall -runoff models as well as physically based distributed rainfall–runoff models. The MSI analysis is a practical method for weeding out the appropriate model input on a catchment scale basis, however a more robust analysis where weights are assigned would yield some improvements in the results. To fully ascertain the potential of alternative model forcing, catchments characteristics and human influence such as dams and reservoirs should be modeled.

Finally, it is essential to note that this work does not promote the use of products such as reanalysis to replace observed data from weather stations, nor can it be understood as giving reason to continue the present trend of retiring additional stations. Quality controlled ground observations still act as the best data for research. The ERA5 results demonstrate that atmospheric reanalysis has likely reached the stage where they can consistently supplement data from weather stations and offer trustworthy proxies in places with less dense station networks, at least across Kenya. Overall, reanalysis can be a viable alternative to observations in ungauged catchments, but the associated uncertainties need to be carefully communicated for informed choice of hydrological modelling tools.

CRediT authorship contribution statement

Maureen A. Wanzala: Conceptualization, Data curation- Lead, Methodology- Lead, Investigation, Software, Formal analysis - Lead, Visualization, Writing – original draft, Writing – review & editing. **Andrea Ficchì**: Conceptualization, Supervision, Data curation, Supporting, Methodology, Supporting, Investigation, Software, Analysis, Supporting, Visualization, Writing – review & editing, Supporting. **Hannah L. Cloke**: Conceptualization, Methodology, Investigation, Writing – review & editing, Project administration, Supervision. **Elisabeth M. Stephens**: Conceptualization, Methodology, Investigation, Writing – review & editing, Supporting, Project administration, Supervision, Funding acquisition. **Heou M. Badjana**: Methodology, Data curation, Supporting, Writing – review & editing, Supporting. **David A. Lavers**: Data curation, Investigation.

Declaration of Competing Interest

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

Acknowledgements

No. NE/R004668/1).

References

Acharya, Suwash Chandra, Nathan, Rory, Wang, Quan J., Su, Chun-Hsu, Eizenberg, Nathan, 2019. An evaluation of daily precipitation from a regional atmospheric reanalysis over Australia. Hydrol. Earth Syst. Sci. 23 (8), 3387–3403.

Adeyeri, O.E., Laux, P., Arnault, J., Lawin, A.E., Kunstmann, H., 2020. Conceptual hydrological model calibration using multi-objective optimization techniques over the transboundary Komadugu-Yobe basin, Lake Chad Area, West Africa. J. Hydrol. Reg. Stud. https://doi.org/10.1016/j.ejrh.2019.100655.

Alemayehu, Tadesse, Kilonzo, Fidelis, van Griensven, Ann, Bauwens, Willy, 2018. Evaluation and application of alternative rainfall data sources for forcing hydrologic models in the Mara Basin. Hydrol. Res. 49 (4), 1271–1282.

Arshad, Muhammad, Ma, Xieyao, Yin, Jun, Ullah, Waheed, Liu, Mengyang, Ullah, Irfan, 2021. Performance evaluation of ERA-5, JRA-55, MERRA-2, and CFS-2 reanalysis datasets, over diverse climate regions of Pakistan. Weather Clim. Extrem. 33, 100373.

Ayugi, Brian, Tan, Guirong, Niu, Rouyun, Dong, Zeyao, Ojara, Moses, Mumo, Lucia, Ongoma, Victor, 2020. Evaluation of meteorological drought and flood scenarios over Kenya. East Afr. Atmosph. 11 (3), 307.

Ayugi, Brian Odhiambo, Wen, Wang, Chepkemoi, Daisy, 2016. Analysis of spatial and temporal patterns of rainfall variations over Kenya. J. Environ. Earth Sci. 6 (11).

Baldassarre, G.Di, Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrol. Earth Syst. Sci. 13 (6), 913–921.

Beck, Hylke E., van Dijk, Albert, I.J.M., de Roo, Ad, Dutra, Emanuel, Fink, Gabriel, Orth, Rene, Schellekens, Jaap, 2017a. Global evaluation of runoff from ten state-of-the-art hydrological models. Hydrol. Earth Syst. Sci. 21 (6), 2881–2903.

Beck, Hylke E., Vergopolan, Noemi, Pan, Ming, Levizzani, Vincenzo, van Dijk, Albert I.J.M., Weedon, Graham P., Wood, Eric F., 2017b. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrol. Earth Syst. Sci. 21 (12), 6201–6217.

Beck, Hylke E., Pan, Ming, Miralles, Diego, G., Reichle, Rolf, H., Dorigo, Wouter, A., Hahn, Sebastian, Parinussa, Robert M., 2021. Evaluation of 18 satellite-and model-based soil moisture products using in situ measurements from 826 sensors. Hydrol. Earth Syst. Sci. 25 (1), 17–40.

Betts, Alan K., Chan, Darren Z., Desjardins, Raymond L., 2019. Near-surface biases in ERA5 over the Canadian prairies. Front. Environ. Sci. 7, 129.

Beven, Keith, 2012. Rainfall-Runoff Modelling: The Primer: Second Edition. Rainfall-Runoff Modelling: The Primer, second ed. John Wiley and Sons. https://doi.org/10.1002/9781119951001.

Bitew, Menberu M., Gebremichael, Mekonnen, Ghebremichael, Lula T., Bayissa, Yared A., 2012. Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia. J. Hydrometeorol. 13 (1), 338–350.

Black, Emily, Slingo, Julia, Sperber, Kenneth R., 2003. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. Mon. Weather Rev. 131 (1), 74–94.

Blau, M.T., Ha, K.-J., 2020. The Indian Ocean dipole and its impact on East African short rains in two CMIP5 historical scenarios with and without anthropogenic influence. J. Geophys. Res. Atmosph. 125 (16), e2020JD033121.

Blöschl, G.ünter, Sivapalan, Murugesu, 1995. Scale issues in hydrological modelling: a review. Hydrol. Process. 9 (3–4), 251–290.

Chen, Jie, Brissette, François P., Chen, Hua, 2018. Using reanalysis-driven regional climate model outputs for hydrology modelling. Hydrol. Process. 32 (19), 3019–3031.

Collier, Emily, & Mölg, Thomas. (2016). Analysis and simulation of recent climate variability in the high-mountain regions of East Africa. In EGU General Assembly Conference Abstracts (pp. EPSC2016–6136).

Coron, Laurent, Delaigue, Olivier, Thirel, Guillaume, Perrin, Charles, Michel, Claude, Andréassian, Vazken, … Mathevet, Thibaut. (2019). airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling (v. 1.2. 13.16).

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Vitart, F., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137 (656), 553–597. https://doi.org/10.1002/qj.828.

Delaigue, Olivier, Thirel, Guillaume, Coron, Laurent, & Brigode, Pierre. (2019). airGR and airGRteaching: two packages for rainfall-runoff modeling and teaching hydrology. In 15th edition of the International R User Conference (p. 1).

Dhanya, C.T., Villarini, Gabriele, 2017. An investigation of predictability dynamics of temperature and precipitation in reanalysis datasets over the continental United States. Atmos. Res. 183, 341–350.

Dile, Yihun Taddele, Srinivasan, Raghavan, 2014. Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. JAWRA J. Am. Water Resour. Assoc. 50 (5), 1226–1241.

Dinku, T., 2019. Challenges with availability and quality of climate data in Africa. Extreme Hydrology and Climate Variability. Elsevier, pp. 71–80.

Dinku, Tufa, Funk, Chris, Peterson, Pete, Maidment, Ross, Tadesse, Tsegaye, Gadain, Hussein, Ceccato, Pietro, 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Q. J. R. Meteorol. Soc. 144, 292–312.

Diro, G.T., Grimes, D.I.F., Black, E., O'Neill, A., Pardo-Iguzquiza, E., 2009. Evaluation of reanalysis rainfall estimates over Ethiopia. Int. J. Climatol. A J. R. Meteorol. Soc. 29 (1), 67–78.

Emerton, Rebecca, Cloke, H.L., Stephens, E.M., Zsoter, E., Woolnough, S.J., Pappenberger, F., 2017. Complex picture for likelihood of ENSO-driven flood hazard. Nat. Commun. 8 (1), 1–9.

Essou, Gilles R.C., Brissette, François, Lucas-Picher, Philippe, 2017. The use of reanalyses and gridded observations as weather input data for a hydrological model: comparison of performances of simulated river flows based on the density of weather stations. J. Hydrometeorol. 18 (2), 497–513.

Ficchì, Andrea, Perrin, Charles, Andréassian, Vazken, 2016. Impact of temporal resolution of inputs on hydrological model performance: an analysis based on 2400 flood events. J. Hydrol. 538, 454–470.

Ficchì, Andrea, Perrin, Charles, Andréassian, Vazken, 2019. Hydrological modelling at multiple sub-daily time steps: model improvement via flux-matching. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2019.05.084.

Finney, Declan L., Marsham, John H., Walker, Dean P., Birch, Cathryn E., Woodhams, Beth J., Jackson, Lawrence S., Hardy, Sam, 2020. The effect of westerlies on East African rainfall and the associated role of tropical cyclones and the Madden–Julian Oscillation. Q. J. R. Meteorol. Soc. 146 (727), 647–664.

Fortin, Vincent, Roy, Guy, Donaldson, Norman, Mahidjiba, Ahmed, 2015. Assimilation of radar quantitative precipitation estimations in the Canadian Precipitation Analysis (CaPA). J. Hydrol. 531, 296–307.

Funk, Chris, Peterson, Pete, Landsfeld, Martin, Pedreros, Diego, Verdin, James, Shukla, Shraddhanand, Michaelsen, Joel, 2015. The climate hazards infrared precipitation with stations - a new environmental record for monitoring extremes. Sci. Data 2, 1–21. https://doi.org/10.1038/sdata.2015.66.

Garibay, Victoria M., Joseph, H.A.Guillaume, Crokeab, Barry F.W., Jakemanbc, Anthony J., 2021. Evaluation of Reanalysis Precipitation Data and Potential Bias Correction Methods for Use in Data-Scarce Areas. Water Resources Management 35, 1587–1602. https://doi.org/10.1007/s11269-021-02804-8.

Gabellani, S., Boni, G., Ferraris, L., Von Hardenberg, J., Provenzale, A., 2007. Propagation of uncertainty from rainfall to runoff: a case study with a stochastic rainfall generator. Adv. Water Resour. 30 (10), 2061–2071.

Gleixner, Stephanie, Demissie, Teferi, Diro, Gulilat Tefera, 2020. Did ERA5 improve temperature and precipitation reanalysis over East Africa? Atmosphere 11 (9), 996.

Golian, Saeed, Murphy, Conor, 2021. Evaluation of sub-selection methods for assessing climate change impacts on low-flow and hydrological drought conditions. Water Resour. Manag. 35 (1), 113–133.

Gosset, Marielle, Viarre, Julien, Quantin, Guillaume, Alcoba, Matias, 2013. Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. Q. J. R. Meteorol. Soc. 139 (673), 923–940. https://doi.org/10.1002/qj.2130.

Gudoshava, Masilin, Misiani, Herbert O., Segele, Zewdu T., Jain, Suman, Ouma, Jully O., Otieno, George, Osima, Sarah, 2020. Projected effects of 1.5C and 2C global warming levels on the intra-seasonal rainfall characteristics over the Greater Horn of Africa. Environ. Res. Lett. 15 (3), 34037.

Guo, Binbin, Zhang, Jing, Xu, Tingbao, Croke, Barry, Jakeman, Anthony, Song, Yongyu, Liao, Weihong, 2018. Applicability assessment and uncertainty analysis of multi-precipitation datasets for the simulation of hydrologic models. Water 10 (11), 1611.

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

Hamon, W.Russell, 1960. Estimating Potential Evapotranspiration. Massachusetts Institute of Technology.

Harada, Yayoi, Kamahori, Hirotaka, Kobayashi, Chiaki, Endo, Hirokazu, Kobayashi, Shinya, Ota, Yukinari, Takahashi, Kiyotoshi, 2016. The JRA-55 Reanalysis: Representation of atmospheric circulation and climate variability. J. Meteorol. Soc. Jpn. Ser. II 94 (3), 269–302.

Hargreaves, George H., Samani, Zohrab A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1 (2), 96–99.

Hartman, Adam T., 2018. An analysis of the effects of temperatures and circulations on the strength of the low-level jet in the Turkana Channel in East Africa. Theor. Appl. Climatol. 132 (3), 1003–1017.

Hersbach, H. (2018). Operational global reanalysis: progress, future directions and synergies with NWP. European Centre for Medium Range Weather Forecasts.

Indeje, M., Semazzi, F.H.M., 2000. Relationships between QBO in the lower equatorial stratospheric zonal winds and East African seasonal rainfall. Meteor. Atmos. Phys. vol. 73, 227–244. https://doi.org/10.1007/s007030050075.

Khan, Sadiq I., Adhikari, Pradeep, Hong, Yang, Vergara, H., F Adler, R., Policelli, F., Okello, L., 2011. Hydroclimatology of Lake Victoria region using hydrologic model and satellite remote sensing data. Hydrol. Earth Syst. Sci. 15 (1), 107–117.

Kilavi, Mary, MacLeod, Dave, Ambani, Maurine, Robbins, Joanne, Dankers, Rutger, Graham, Richard, Todd, Martin C., 2018. Extreme rainfall and flooding over central Kenya including Nairobi city during the long-rains season 2018: causes, predictability, and potential for early warning and actions. Atmosphere 9 (12), 472.

Kinuthia, Joseph Hiri, 1992. Horizontal and vertical structure of the Lake Turkana jet. J. Appl. Meteorol. Climatol. 31 (11), 1248–1274.

Klemeš, Vit, 1986. Operational testing of hydrological simulation models. Hydrol. Sci. J. 31 (1), 13–24.

Kling, Harald, Fuchs, Martin, Paulin, Maria, 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. J. Hydrol. 424, 264–277.

Knoben, Wouter J.M., Freer, Jim E., Woods, Ross A., 2019. Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23 (10), 4323–4331.

Kobayashi, Shinya, Yukinari, Ota, Yayoi, Harada, Ebita, Ayataka, Moriya, Masami, Onoda, Hirokatsu, Takahashi, Kiyotoshi, 2015. The JRA-55 reanalysis: general specifications and basic characteristics. J. Meteorol. Soc. Jpn. Ser. II 93 (1), 5–48. https://doi.org/10.2151/jmsj.2015-001.

Koukoula, Marika, Nikolopoulos, Efthymios I., Dokou, Zoi, Anagnostou, Emmanouil N., 2020. Evaluation of global water resources reanalysis products in the Upper Blue Nile River Basin. J. Hydrometeorol. 21 (5), 935–952.

Lakew, Haileyesus Belay, Moges, Semu Ayalew, Asfaw, Dereje Hailu, 2020. Hydrological performance evaluation of multiple satellite precipitation products in the upper Blue Nile basin, Ethiopia. J. Hydrol. Reg. Stud. 27, 100664.

Langat, Philip Kibet, Kumar, Lalit, Koech, Richard, 2017. Temporal variability and trends of rainfall and streamflow in Tana River Basin, Kenya. Sustainability 9 (11), 1963.

Le, Alyssa M., Pricope, Narcisa G., 2017. Increasing the accuracy of runoffand streamflow simulation in the Nzoia Basin, Western Kenya, through the incorporation of satellite-derived CHIRPS data. Water 9 (2). https://doi.org/10.3390/w9020114.

Lemma, Estifanos, Upadhyaya, Shruti, Ramsankaran, R.A.A.J., 2019. Investigating the performance of satellite and reanalysis rainfall products at monthly timescales across different rainfall regimes of Ethiopia. Int. J. Remote Sens. 40 (10), 4019–4042.

MacLeod, D.A., Dankers, R., Graham, R., Guigma, K., Jenkins, L., Todd, M.C., Mwangi, E., 2021. Drivers and subseasonal predictability of heavy rainfall in equatorial East Africa and relationship with flood risk. J. Hydrometeorol. 22 (4), 887–903.

Mahto, Shanti Shwarup, Mishra, Vimal, 2019. Does ERA-5 outperform other reanalysis products for hydrologic applications in India? J. Geophys. Res. Atmosph. 124 (16), 9423–9441.

Mathevet, T., Michel, C., Andreassian, V., Perrin, C., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. IAHS Publ. 307, 211.

Menne, Matthew J., Williams, Claude N., Gleason, Byron E., Rennie, J.Jared, Lawrimore, Jay H., 2018. The global historical climatology network monthly temperature dataset, version 4. J. Clim. 31 (24), 9835–9854.

Meresa, Hadush, Murphy, Conor, Fealy, Rowan, Golian, Saeed, 2021. Uncertainties and their interaction in flood hazard assessment with climate change. Hydrol. Earth Syst. Sci. 25 (9), 5237–5257.

Michel, Claude, 1983. Que peut-on faire en hydrologie avec modèle conceptuel à un seul paramètre? La Houille Blanc (1), 39–44.

Moriasi, Daniel N., Arnold, Jeffrey G., Van Liew, Michael W., Bingner, Ronald L., Harmel, R.Daren, Veith, Tamie L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Mostafaie, A., Forootan, E., Safari, A., Schumacher, M., 2018. Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data. Comput. Geosci. 22 (3), 789–814. https://doi.org/10.1007/s10596-018-9726-8.

Nicholson, S.E., Klotter, D., Zhou, L., Hua, W., 2019. Validation of satellite precipitation estimates over the Congo Basin. J. Hydrometeorol. 20 (4), 631–656.

Njogu, Humphrey Waita, 2021. Effects of floods on infrastructure users in Kenya. J. Flood Risk Manag. 14 (4), e12746.

Nkiaka, E., Nawaz, N.R., Lovett, J.C., 2017. Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: case study of the Logone catchment, Lake Chad Basin. Meteorol. Appl. 24 (1), 9–18.

Nossent, Jiri, Elsen, Pieter, Bauwens, Willy, 2011. Sobol'sensitivity analysis of a complex environmental model. Environ. Model. Softw. 26 (12), 1515–1525.

Nyenzi, B.S. (1988). Mechanisms of East African rainfall variability.

Ogallo, L.A., 1993. Dynamics of the East African climate. Proc. Indian Acad. Sci. -Earth Planet. Sci. 102 (1), 203–217.

Ogwang, Bob Alex, Ongoma, Victor, Xing, Li, Ogou, Katchele Faustin, 2015. Influence of Mascarene high and Indian Ocean dipole on East African extreme weather events. Geogr. Pannonica 19 (2), 64–72.

Ojara, Moses A., Yunsheng, Lou, Ongoma, Victor, Mumo, Lucia, Akodi, David, Ayugi, Brian, Ogwang, Bob Alex, 2021. Projected changes in East African climate and its impacts on climatic suitability of maize production areas by the mid-twenty-first century. Environ. Monit. Assess. 193 (12), 1–24.

Ongoma, V., Guirong, T., Ogwang, B.A., Ngarukiyimana, J.P., 2015. Diagnosis of seasonal rainfall variability over east Africa: a case study of 2010-2011 drought over Kenya. Pak. J. Meteorol. Vol. 11.

Ongoma, Victor, Chen, Haishan, 2017. Temporal and spatial variability of temperature and precipitation over East Africa from 1951 to 2010. Meteorol. Atmos. Phys. 129 (2), 131–144.

Ongoma, Victor, Chen, Haishan, Omony, George William, 2018. Variability of extreme weather events over the equatorial East Africa, a case study of rainfall in Kenya and Uganda. Theor. Appl. Climatol. 131 (1–2), 295–308.

Onyutha, Charles, 2016. Variability of seasonal and annual rainfall in the River Nile riparian countries and possible linkages to ocean–atmosphere interactions. Hydrol. Res. 47 (1), 171–184.

Opere, A., 2013. Floods in Kenya. In: Developments in Earth Surface Processes, Vol. 16. Elsevier.

Oudin, Ludovic, Andréassian, Vazken, Perrin, Charles, Anctil, François, 2004. Locating the sources of low-pass behavior within rainfall-runoff models. Water Resour. Res. 40, 11.

Owiti, Z.O., Ogallo, L.A., Mutemi, J., 2008. Linkages between the Indian Ocean Dipole and East African seasonal rainfall anomalies. J. Kenya Meteor. Soc. vol. 2, 3–17.

Perrin, Charles, Michel, Claude, Andréassian, Vazken, 2003. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 279 (1–4), 275–289.

Probst, E., Mauser, W., 2022. Evaluation of ERA5 and WFDE5 forcing data for hydrological modelling and the impact of bias correction with regional climatologies: a case study in the Danube River Basin. J. Hydrol.: Reg. Stud. 40, 101023.

Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., Andréassian, V., 2011. A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. J. Hydrol. 411 (1–2), 66–76.

Rapaić, Maja, Brown, Ross, Markovic, Marko, Chaumont, Diane, 2015. An evaluation of temperature and precipitation surface-based and reanalysis datasets for the Canadian Arctic, 1950–2010. Atmosph. Ocean 53 (3), 283–303.

Saha, Suranjana, Moorthi, Shrinivas, Pan, Hua-Lu, Wu, Xingren, Wang, Jiande, Nadiga, Sudhir, Goldberg, Mitch, 2010. The NCEP climate forecast system reanalysis. Bull. Am. Meteorol. Soc. 91 (8), 1015–1058. https://doi.org/10.1175/2010bams3001.1.

Saltelli, Andrea, Tarantola, Stefano, Campolongo, Francesca, 2000. Sensitivity analysis as an ingredient of modeling. Stat. Sci. 377–395.

Schulzweida, Uwe. (2019). CDO user guide. Clim Data Oper.

Shayeghi, Afshin, Azizian, Asghar, Brocca, Luca, 2020. Reliability of reanalysis and remotely sensed precipitation products for hydrological simulation over the Sefidrood River Basin, Iran. Hydrol. Sci. J. 65 (2), 296–310.

Shin, Mun-Ju, Kim, Chung-Soo, 2017. Assessment of the suitability of rainfall–runoff models by coupling performance statistics and sensitivity analysis. Hydrol. Res. 48 (5), 1192–1213.

Shin, Mun-Ju, Gitau, Margaret W., Kiggundu, Nicholas, Moriasi, Daniel, Mishili, Fulgence, 2013. Addressing ten questions about conceptual rainfall–runoff models with global sensitivity analyses in R. Journal of Hydrology. https://doi.org/10.1016/j.jhydrol.2013.08.047.

Shin, Mun-Ju, Guillaume, Joseph, H.A., Croke, Barry, F.W., Jakeman, Anthony J., 2015. A review of foundational methods for checking the structural identifiability of models: results for rainfall-runoff. J. Hydrol. 520, 1–16.

Sun, Qiaohong, Miao, Chiyuan, Duan, Qingyun, Ashouri, Hamed, Sorooshian, Soroosh, Hsu, Kuo-Lin, 2018. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Rev. Geophys. 56 (1), 79–107.

ReliefWeb, 2019b: UNICEF sends supplies to flood-affected children and families in 12counties in Kenya - Kenya | ReliefWeb. (n.d.). Retrieved March 3, 2020, from ⟨https://reliefweb.int/report/kenya/unicef-sends-supplies-flood-affected-children-and-families-12-counties-kenya⟩.

ReliefWeb, 2019a - Kenya: Floods and Landslides - Oct 2019 | ReliefWeb. (n.d.). Retrieved March 3, 2020, from ⟨https://reliefweb.int/disaster/fl-2019–000138-ken⟩.

Takaoka, S., 2005. Impact of the 1997–1998 El Niño rains on farms in the Mount Kenya region. Mt. Res. Dev. 25 (4), 326–331.

Tang, Guoqiang, Clark, Martyn P., Papalexiou, Simon Michael, Ma, Ziqiang, Hong, Yang, 2020. Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. Remote Sens. Environ. 240, 111697.

Tapiador, Francisco J., Turk, Francis J., Petersen, Walt, Hou, Arthur Y., García-Ortega, Eduardo, Machado, Luiz A.T., Huffman, George J., 2012. Global precipitation measurement: methods, datasets and applications. Atmos. Res. 104, 70–97.

Tarek, Mostafa, Brissette, François P., Arsenault, Richard, 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrol. Earth Syst. Sci. 24 (5), 2527–2544.

Tarek, Mostafa, Brissette, François, Arsenault, Richard, 2021. Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies. Hydrol. Earth Syst. Sci. 25 (6), 3331–3350.

Tesfaye, T.W., Dhanya, C., Gosain, A., 2017. Evaluation of ERA-interim, MERRA, NCEP-DOE R2 and CFSR reanalysis precipitation data using gauge observation over ethiopia for a period of 33 years. AIMS Environ. Sci. 4, 596–620.

van Werkhoven, Kathryn, Wagener, Thorsten, Reed, Patrick, Tang, Yong, 2009. Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. Advances in Water Resources 32, 1154–1169. https://doi.org/10.1016/j.advwatres.2009.03.002.

Van Esse, W.R., Perrin, C., Booij, Martijn J., Augustijn, Dionysius C.M., Fenicia, F., Kavetski, D., Lobligeois, F., 2013. The influence of conceptual model structure on model performance: a comparative study for 237 French catchments. Hydrol. Earth Syst. Sci. 17 (10), 4227–4239.

van Griensven, A. van, Meixner, Thomas, Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. J. Hydrol. 324 (1–4), 10–23.

Viglione, Alberto, Chirico, Giovanni Battista, Komma, J.ürgen, Woods, Ross, Borga, Marco, Blöschl, G.ünter, 2010. Quantifying space-time dynamics of flood event types. J. Hydrol. 394 (1–2), 213–229.

Vischel, T., Quantin, G., Lebel, Thierry, Viarre, J., Gosset, Marielle, Cazenave, Frédéric, Panthou, G., 2011. Generation of high-resolution rain fields in West Africa: evaluation of dynamic interpolation methods. J. Hydrometeorol. 12 (6), 1465–1482.

Wainwright, Caroline M., Finney, Declan L., Kilavi, Mary, Black, Emily, Marsham, John H., 2021. Extreme rainfall in East Africa, October 2019–January 2020 and context under future climate change. Weather 76 (1), 26–31.

Wang, Gefei, Zhang, Xiaowen, Zhang, Shiqiang, 2019. Performance of three reanalysis precipitation datasets over the qinling-daba mountains, eastern fringe of Tibetan Plateau, China. Adv. Meteorol. 2019.

Worqlul, Abeyou W., Yen, Haw, Collick, Amy S., Tilahun, Seifu A., Langan, Simon, Steenhuis, Tammo S., 2017. Evaluation of CFSR, TMPA 3B42 and ground-based rainfall data as input for hydrological models, in data-scarce regions: the upper Blue Nile Basin, Ethiopia. Catena 152, 242–251.

Xu, Lei, Chen, Nengcheng, Moradkhani, Hamid, Zhang, Xiang, Hu, Chuli, 2020. Improving global monthly and daily precipitation estimation by fusing gauge observations, remote sensing, and reanalysis data sets. Water Resour. Res. 56 (3), e2019WR026444.

Yang, Wenchang, Seager, Richard, Cane, Mark A., Lyon, Bradfield, 2015. The annual cycle of East African precipitation. J. Clim. 28 (6), 2385–2404.

Zaitchik, Benjamin F., Montanari, Alberto, Beven, Keith J., Kirkby, Michael J., Mike, J., Pappenberger, Floria N., Smith, Tyler Jon, Obled, Charles, 2011. Rainfall estimation in the Sahel. Part II: evaluation of rain gauge networks in the CILSS countries and objective intercomparison of rainfall products. J. Hydrol. 9 (3), 2. https://doi.org/10.1016/j.jhydrol.2019.05.084.

Zeng, Ling, Xiong, Lihua, Liu, Dedi, Chen, Jie, Kim, Jong Suk, 2019. Improving parameter transferability of GR4J model under changing environments considering nonstationarity. Water 11 (10). https://doi.org/10.3390/w11102029.

Zhang, Chi, Academy, Chinese, 2016. Hu et al-2016-JGR. J. Geophys. Res. Atmosph. 121 (June), 5648–5663. https://doi.org/10.1002/2016JD024781.Received.