Global scale evaluation of precipitation datasets for hydrological modelling

Solomon H. Gebrechorkos[1,2], Julian Leyland[2], Simon J. Dadson[1], Sagy Cohen[3], Louise Slater[1], Michel Wortmann[1], Philip J. Ashworth[4], Georgina L. Bennett[5], Richard Boothroyd[6], Hannah Cloke[7,8], Pauline Delorme[9], Helen Griffith[7], Richard Hardy[10], Laurence Hawker[11], Stuart McLelland[9], Jeffrey Neal[11], Andrew Nicholas[5], Andrew J. Tatem[2], Ellie Vahidi[5], Yinxue Liu[1], Justin Sheffield[2], Daniel R. Parsons[10], Stephen E. Darby[2]

[1]School of Geography and the Environment, University of Oxford, Oxford, UK

[2]School of Geography and Environmental Science, University of Southampton, Southampton, SO17 1BJ, United Kingdom

[3]Department of Geography and the Environment, University of Alabama, Tuscaloosa, AL, USA

[4]School of Applied Sciences, University of Brighton, Sussex, BN2 4AT

[5]Department of Geography, Faculty of Environment, Science and Economy, University of Exeter, Exeter, EX4 4RJ, United Kingdom

[6]School of Geographical & Earth Sciences, University of Glasgow, UK

[7]Department of Geography and Environmental Science, University of Reading, UK

[8]Department of Meteorology, University of Reading, UK

[9]Energy and Environment Institute, University of Hull, Hull, United Kingdom

[10]Department of Geography, Durham University, Lower Mountjoy, South Road, Durham, DH1 3LE

[11]School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, UK

Correspondence to: Solomon H. Gebrechorkos (solomon.gebrechorkos@ouce.ox.ac.uk)

Abstract. Precipitation is the ~~most important~~dominant driver of the hydrological cycle but it is challenging to ~~estimate~~ accurately measure at suitably high resolution over ~~large~~ global scales ~~using~~from satellites and models. Here, we assessed the performance of six global and quasi-global high-resolution precipitation datasets (European Center for Medium-range Weather Forecast (ECMWF) Reanalysis version 5 ~~ERA5 global reanalysis~~ (ERA5), Climate Hazards group Infrared Precipitation with Stations version 2.0 (CHIRPS), Multi-Source Weighted-Ensemble Precipitation version 2.80 (MSWEP), TerraClimate (TERRA), Climate Prediction Centre Unified version 1.0 (CPCU) and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System-Climate Data Record (PERCCDR)) for hydrological modelling globally and quasi-globally. We forced the WBMsed global hydrological model with the precipitation datasets to simulate river discharge from 1983 to 2019 and evaluated the predicted discharge against ~~more than 1800~~1825 hydrological stations worldwide, using a range of statistical methods. The results show large differences in the accuracy of discharge predictions when using different precipitation input datasets. Based on evaluation at annual, monthly and daily time scales, MSWEP followed by ERA5 demonstrated a higher CC and KGE than other datasets for more than 50% of the stations. Whilst~~,~~ ERA5 was the second-highest performing dataset,~~, and~~ it ~~showed~~exhibited the highest error and bias in about 20% of the stations. The PERCCDR is the least well-performing dataset with ~~large bias (percentage of~~ bias up to 99%~~)~~ and e~~rrors (a~~ normalised root mean square error up to 247%~~).~~ PERCCDR ~~only~~ ~~revealed~~show with a higher KGE and CC than the other products in less than 10% of the stations. Even though MSWEP provided the ~~highest~~ best performance overall, our analysis reveals high spatial variability, meaning that it is important to consider other datasets in areas where MSWEP showed a lower performance. The results of this study highlight the importance of, and provide guidance on, the selection of precipitation datasets for modelling river discharge for a basin, region or climatic zone, as there is no single best precipitation dataset globally. Finally, the large discrepancy in the performance of the datasets in different parts of the world highlights the need to improve global precipitation data products.

1. Introduction

Whilst precipitation is one the most important components of the global hydrological cycle and regulates the climate system (Miao et al., 2019; Sadeghi et al., 2021), it remains one of the most challenging variables to estimate at a global scale using satellite data and modelling approaches (Michaelides et al., 2009; Kidd and Levizzani, 2011; Beck et al., 2017a; Ursulak and Coulibaly, 2021). Reliable precipitation data with sufficient spatial and temporal coverage and accurate representation of extreme events is crucial for various applications. These include the development ofing water resource management and planning strategies, hydrological applications including forecasting hydrological extremes, and climate change analysis (Mehran and AghaKouchak, 2014; Nguyen et al., 2018; Sadeghi et al., 2021; Acharya et al., 2019). Observed precipitation from meteorological stations is typically used at local to river basin scale with gauge-based gridded precipitation datasets, such as from the Global Historical Climatology Network (Menne et al., 2012), developed to study climate and hydrology over larger scales. However, precipitation from gauges and gauge-based gridded datasets have several drawbacks such as limited spatial and temporal coverage, prevalence of missing values, and limited accuracy in sparsely populated and remote areas (Kidd and Levizzani, 2011; Reichle et al., 2011; Kidd et al., 2017; Sun et al., 2018; Gebrechorkos et al., 2018; Hafizi and Sorman, 2022). In addition, data-sharing policies have caused significant challenges in obtaining data, particularly in developing countries (Gebrechorkos et al., 2018; Hafizi and Sorman, 2022).

Given the challenges in representing precipitation at global scales, satellite, climate model, and reanalysis-based precipitation datasets can form the basis for monitoring and prediction of water resources and hydrological extremes, particularly in data-scarce regions of the world (Sheffield et al., 2018; Dembélé et al., 2020). Nevertheless, uncertainties and errors in these datasets require careful analysis to assess their suitability for a specific use. Error in satellite-based precipitation estimates can be due to errors in the sensor measurements, the frequency of sampling, and the retrieval algorithms, including the representation of cloud physics (Dembélé et al., 2020; Laiti et al., 2018; Alazzy et al., 2017). Climate model-based datasets, including reanalyses, have large uncertainty due to their coarse spatial resolution and ambiguity associated with model parameters (Gebrechorkos et al., 2018; AL-Falahi et al., 2020; Dembélé et al., 2020; Her et al., 2019). Reanalysis datasets may correct for some of these errors via the assimilation of observational data, but this comes with its own uncertainties due to the error characteristics of the assimilated observations and the assimilation scheme (Sheffield et al., 2006; Parker, 2016). In hydrological modelling, errors and biases in precipitation data result in poor representation of the hydrological responses and affect applications (Maggioni and Massari, 2018; Zambrano-Bigiarini et al., 2016). For example, according to Bárdossy et al. (2022), uncertainty in precipitation can lead to hydrological model errors of up to 50%. Hence, it is important to assess the quality and accuracy of the precipitation products before using them in global or basin-scale hydrological models. In data-limited regions, hydrological models driven by precipitation datasets developed from satellite sources, reanalysis or climate models are the only plausible way to represent the terrestrial water cycle (van Huijgevoort et al., 2013).

Over the last few decades, several global and quasi-global precipitation datasets have been developed that address some of these challenges and can be used to drive hydrological models at regional and global scales. These precipitation datasets differ in terms of their spatial resolution, spatial coverage (e.g., global or regional), data

96 sources (e.g., gauge, satellite, reanalysis, and radar), temporal resolution (e.g., sub-daily and daily), and length of

97 record. It is therefore important to evaluate the accuracy of the datasets before they are used to drive global or

98 regional scale hydrological models. Most studies have evaluated precipitation datasets using observed data from

99 field-based meteorological stations at a range of scales (e.g., Beck et al., 2017a; Gebrechorkos et al., 2018; Xiang

100 et al., 2021; Sun et al., 2018; Hong et al., 2022; Wati et al., 2022; AL-Falahi et al., 2020; Ahmed et al., 2019;

101 Fallah et al., 2020). Hydrological models have also been used to assess the quality of the precipitation dataset by

102 comparing simulated and observed discharge across different spatial scales (e.g., Mazzoleni et al., 2019; Beck et

103 al., 2017a; Zhu et al., 2018; Raimonet et al., 2017; Guo et al., 2018; Wang et al., 2020; Salehi et al., 2022; Zhu et

104 al., 2018; Seyyedi et al., 2015). In principle, this latter approach is able to identify the precipitation datasets which

105 best represent hydrological variability including extremes, even in catchments where there have been multiple

106 drivers of change.

107 There are a limited number of studies assessing multiple precipitation datasets for global hydrological model

108 applications (Voisin et al., 2008; Beck et al., 2017a; Mazzoleni et al., 2019). _Voisin et al. (2008) conducted a_

109 _global-scale evaluation of two precipitation products for hydrological modelling._ Beck et al., (2017a) compared

110 the performance of ~~multiple~~ 22 precipitation datasets ~~(e.g., the Climate Hazards group Infrared Precipitation with~~

111 ~~Stations (CHIRPS, version 2.0), Multi Source Weighted Ensemble Precipitation (MSWEP, version 2.0),~~

112 ~~European Centre for Medium-range Weather Forecasts ReAnalysis Interim (ERA-Interim), and National Centers~~

113 ~~for Environmental Prediction Climate Forecast System Reanalysis (NCEP-CFSR))~~ for global hydrological

114 modelling. Mazzoleni et al. (2019) evaluated ~~multiple~~ 18 different precipitation datasets ~~including MSWEP~~

115 ~~(Version 2.1) and CHIRPS~~ in eight river basins on different continents. Both Beck et al. (2017a) and Mazzoleni

116 et al. (2019) found that merged satellite-observation precipitation products showed the best performance compared

117 to satellite-only products. These studies exclusively concentrate on a daily time scale, evaluating performance

118 solely through the Nash-Sutcliffe Efficiency (NSE). Neither study extends the~~ir~~_ir~~is~~_ assessment to monthly and

119 annual time scales, and notably, they do not assess the hydrological extremes_,_ which _in hydrological terms_ are

120 ~~often~~ considered important to capture. Here, we build upon the work by Beck et al., (2017a) by adding recently

121 developed high-resolution precipitation datasets_. These include_ ~~such as~~ the European ~~Center~~Centre for Medium-

122 range Weather Forecast (ECMWF) Reanalysis version 5 (ERA5_)_ (Hersbach et al., 2020), TerraClimate

123 (Abatzoglou et al., 2018)_,_ ~~and~~ Precipitation Estimation from Remotely Sensed Information using Artificial Neural

124 Networks-Cloud Classification System-Climate Data Record (_PERCCDR_~~PERSIANN CCS CDR~~, Sadeghi et al.,

125 2021) and the latest _Multi-Source Weighted-Ensemble Precipitation_ ~~MSWEP~~ version (~~2.80~~_)_ _(MSWEP)_. These

126 additions significantly broaden the scope of ~~our~~ _this_ study _when compared to earlier efforts_, offering a diverse

127 range of products with distinct methodologies. In addition, _herein_ we use multiple statistical metrics to evaluate

128 the performance of the precipitation products for hydrological modelling at daily, monthly and annual time scales

129 and for daily extremes, which represents a current gap in the modelling literature.

130 The aim of this study is to undertake a comprehensive evaluation, spanning various temporal and spatial scales,

131 to examine how different input precipitation datasets impact the predictions of a global hydrological model. We

132 assess six high-resolution precipitation datasets, each with records spanning over 30 years. A comprehensive and

133 physically based gridded global hydrological model (WBMsed; Cohen et al., (2013)) is used to simulate river

134 discharge globally. _The objective is not to evaluate the absolute performance of the hydrological model, which_

135 can be influenced by local factors, rather our focus is on comparing the relative performance of these six

136 precipitation datasets at individual locations. The modelled discharge, derived from the six precipitation datasets,

137 is assessed across the various time scales by comparing it with observed discharge data collected from 1825 river

138 gauge stations worldwide. Furthermore, we assess the performance of the precipitation products by examining

139 their accuracy in representing a range of daily extreme precipitation events across various percentiles. In summary,

140 this research offers a thorough evaluation of a this set of diverse set precipitation products, spanning from daily

141 extreme events to annual time scales, providing an invaluable resource for selecting appropriate basin-to-regional-

142 to-global scale inputs for hydrological modelling applications.

143   2.   Data and methods

144 In the following sections, we outline the various input and evaluation datasets which were used within the

145 WBMsed hydrological modelling framework. The statistical evaluation methods used to assess the results are also

146 outlined.

147   2.1. Input global and quasi-global pPrecipitation datasets

148 The precipitation datasets used herein are selected based on their length of record (>30 years period), and spatial

149 coverage (global and quasi-global) and recommendations from previous research (Beck et al., 2017a) (Table 1).

150 Based on the findings of Beck et al. (2017a), datasets with low performance were excluded, while those

151 demonstrating the highest performance, such as MSWEP and Climate Hazards group Infrared Precipitation with

152 Stations version 2.0 (CHIRPS), were retained, and new datasets were also incorporated. The selected precipitation

153 datasets are the ERA5 global reanalysis (ERA5), Climate Hazards group Infrared Precipitation with Stations

154 version 2.0 (CHIRPS), Multi Source Weighted Ensemble Precipitation version 2.80 (MSWEP), TerraClimate

155 (TERRA), Climate Prediction Centre Unified version 1.0 (CPCU), and Precipitation Estimation from Remotely

156 Sensed Information using Artificial Neural Networks Cloud Classification System Climate Data Record

157 (PERCCDR). Due to their spatial coverage, CHIRPS and PERCCDR are evaluated only up to latitudes of 50°N

158 and 60°N, respectively (Table 1). Each dataset was subsequently used to force the WBMsed hydrological model,

159 to generate streamflow estimates. The availability of these datasets with longer records enables the assessment of

160 long-term hydrological changes at global, regional, and catchment scales.

161 ERA5 is the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data

162 available globally from 1940 to present (Hersbach et al., 2020). ERA5 combines modelled data and observations

163 to create a complete and consistent global climate dataset using advanced data assimilation methods. ERA5

164 provides improved precipitation representation such as the inclusion of tropical cyclones when compared to the

165 ERA-Interim (He et al., 2020; Jiao et al., 2021). In addition, ERA5-Land, a subset of ERA5 focusing on land

166 areas, delivers more detailed climate information is available at higher spatial resolution (0.1°) from 1950 to the

167 present compared to ERA5 (Hersbach et al., 2020). Here, ERA5-Land (referred to as ERA5) is used to evaluate

168 its performance for global hydrological modelling. The data is freely available from Copernicus Climate Data

169 Store (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview).

170 CHIRPS is a high-resolution (0.05°) quasi-global rainfall product primarily developed for monitoring droughts

171 and global environmental changes (Funk et al., 2015). CHIRPS provides coupled gauge-satellite precipitation

172 estimates with a 0.05° spatial resolution and long-period records. The product is developed by combining satellite-

173 only Climate Hazards group Infrared Precipitation (CHIRP), Climate Hazards group Precipitation climatology

174 (CHPclim), and data from ground stations. CHIRP and CHPclim were developed based on calibrated infrared

175 cold cloud duration (CCD) precipitation estimates and ground station data from the Global Historical Climate

176 Network (GHCN). The product is available at the Climate Hazards Group (https://www.chc.ucsb.edu/data/chirps/)

177 on daily, 10-day, and monthly timescales from the 1981-near present. Due to its availability at high spatial and

178 temporal resolution, CHIRPS is widely used in hydrological studies (Luo et al., 2019; Gebrechorkos et al., 2020;

179 Geleta and Deressa, 2021; Wang et al., 2021; Opere et al., 2022; Day and Howarth, 2019; Gebrechorkos et al.,

180 2019) and modelling of hydrological extremes such as droughts and floods (Chen et al., 2020; Mianabadi et al.,

181 2022; Peng et al., 2020).

182 MSWEP is a global high-resolution (0.1°) precipitation product developed by merging multiple datasets such as

183 ground stations (~77,000), satellite-based rainfall estimates, and reanalysis data (Beck et al., 2019b). MSWEP

184 includes was developed by merging station data, from the Global Historical Climatology Network Daily (GHCN-

185 D), Global Summary of the Day (GSOD), Global Precipitation Climatology Centre (GPCC), and WorldClim;

186 satellite datasets from the Global Satellite Mapping of Precipitation (GSMaP), Tropical Rainfall Measuring

187 Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA 3B42RT), Climate Prediction Center morphing

188 technique (CMORPH), and Gridded Satellite (GridSat); and reanalysis datasets such as the Japanese 55-year

189 Reanalysis (JRA-55) and European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis

190 (ERA-Interim) (Beck et al., 2017b, 2019b). MSWEP has been widely used in regional and global scale

191 hydrological studies such as for floods and droughts (Gu et al., 2023; Gebrechorkos et al., 2022b; Reis et al., 2022;

192 Wu et al., 2018; Sun et al., 2022; Gebrechorkos et al., 2022c; Xiang et al., 2021; López López et al., 2017) and

193 for developing high-resolution global scale hydrological extreme and climate datasets and regional drought

194 monitoring (Gebrechorkos et al., 2023, 2022a; Li et al., 2022b). MSWEP is available from 1979-present at

195 multiple timescales (e.g., 3 hourly) and can be accessed from the GloH2O website

196 (https://www.gloh2o.org/mswep/).

197 TerraClimate (TERRA) is a high-resolution (0.04°) terrestrial monthly climate (e.g., precipitation and

198 temperature) and climatic water-balance dataset available from 1958-2020 (Abatzoglou et al., 2018). TERRA was

199 developed by combining high and coarse spatial resolution datasets such as WorldClim climatological normals

200 and Climatic Research Unit gridded Time Series (CRU TS) and JRA-55, respectively. The data was evaluated

201 against ground observation from the Historical Climate Network and exhibited better performance than the CRU-

202 TS (Abatzoglou et al., 2018). The monthly climate and climatic water balance is available from the Climatology

203 Lab website (https://www.climatologylab.org/).

204 CPCU is a gauge-based analysis of daily precipitation datasets available globally from 1979 to present at a spatial

205 resolution of 0.5° (Chen et al., 2008). CPCU is the product of the CPC Unified Precipitation project at NOAA

206 Climate Prediction Center. The product uses data from more than 30,000 (1979-2005) and 17,000 (2006-present)

207 stations. The CPCU data is publicly available at the NOAA Physical Sciences Laboratory (PSL,

208 https://downloads.psl.noaa.gov/Datasets/cpc_global_precip/) and has been used for hydrological and climate
209 studies (Beck et al., 2017a; Zhu et al., 2021; Hou et al., 2014).

210 The PERCCDR is a quasi-global (latitude from 60°S to 60°N) dataset developed at the University of California
211 (Sadeghi et al., 2021). PERCCDR provides precipitation estimates at high spatial (0.04°) and temporal (3-hourly)
212 resolutions from 1983 to present. The dataset is developed using the rain rate output from the PERSIANN-CCS
213 model, which uses GridSat-B1 IR and NOAA Climate Prediction Center (CPC-4km) IR data. Compared to other
214 PERSIANN precipitation datasets, PERCCDR provides a realistic representation of precipitation extremes
215 globally and shows better agreement with CPCU precipitation (Sadeghi et al., 2021). The PERCCDR has been
216 used in hydrological studies (Salehi et al., 2022; Eini et al., 2022) and is freely available from the Center for
217 Hydrometeorology and Remote Sensing (CHRS) Data Portal (https://chrsdata.eng.uci.edu/).

218 Table 1. The six precipitation datasets used in this study, their spatial and temporal resolution, spatial coverage
219 and data sources.

Abbreviation	Full name	Spatial resolution and coverage	Temporal resolution	Temporal coverage	Data source	Reference
ERA5	ECMWF (European Centre for Medium-Range Weather Forecasts) Reanalysis V5	0.1°, global	Sub-daily	1979-present	Gauge and reanalysis	(Hersbach et al., 2020)
CHIRPS	Climate Hazards group Infrared Precipitation with Stations (CHIRPS) version 2.0	0.05°, quasi global (50°S-50°N)	Daily	1981-present	Gauge, satellite, and reanalysis	(Funk et al., 2015)
MSWEP	Multi-Source Weighted-Ensemble Precipitation (MSWEP) version 2.80	0.1°, global	Daily	1979-present	Gauge, satellite, and reanalysis	(Beck et al., 2019b)
TERRA	TerraClimate	0.042°, global	Monthly	1958-present	Gauge and reanalysis	(Abatzoglou et al., 2018)
CPCU	Climate Prediction Centre (CPC) Unified V1.0	0.5°, global	Daily	1979-present	Gauge only	(Chen et al., 2008)
PERCCDR	Precipitation Estimation from Remotely Sensed Information using	0.04°, Quasi global	Sub-daily	1983-present	Gauge and satellite	(Sadeghi et al., 2021)

	Artificial Neural Networks-Cloud Classification System-Climate Data Record (PERSIANN-CCS-CDR)	(60°S-60°N)				

2.2. WBMsed hydrological model

The WBMsed (Cohen et al., 2013, 2014) hydrological model is used to assess the performance of the different precipitation datasets for hydrological modelling globally. WBMsed is a global-scale hydrogeomorphic model, an extension of the WBMplus global hydrology model (Wisser et al., 2010), which is part of the FrAMES biogeochemical modelling framework (Wollheim et al., 2008). The WBMplus model is one of the first Global Hydrological Models (GHMs) applied to a global domain (Cohen et al., 2013; Grogan et al., 2022). The model represents the major hydrological cycle components of the land surface and tracks the balances and fluxes between the atmosphere, surface water storages, vegetation, runoff, and groundwater (Grogan et al., 2022). The model includes hydrological infrastructure (e.g., dams), agricultural water requirements, and domestic and industrial water uses. A ~~high resolution~~ gridded river network connects grid cells, which allows the routing of fluxes downstream (e.g., streamflow). The model requires several climate datasets as input in addition to precipitation, including temperature, humidity, air pressure, and wind speed (Table S1). Additional parameters such as field capacity, rooting depth, and riverbed slope are used to drive the model.

We use an identical model setup to that used by Cohen et al., (2022) with all input datasets as detailed in Cohen et al. (2013). ~~with U~~updates include ~~to air temperature which used the~~ daily ERA5 air temperature (Hersbach et al., 2020) ~~dataset~~ re-gridded at 10 arc-minutes resolution~~.;~~ reservoir capacity from ~~—~~global reservoir and dam database (GRanD v1.3; Lehner et al., (2011))~~.; and flow network~~ and a 6 arc-minute HydroSTN30 network ~~which is a derivative~~derived from HydroSHEDS (Lehner et al., 2008) ~~high resolution gridded network~~for flow ~~network (Lehner et al., 2008)~~. In addition, we used each of our six input precipitation datasets, ERA5, CHIRPS, MSWEP, TERRA, CPCU, and PERCCDR in turn, keeping all other parameters and inputs the same. All the input precipitation datasets are bilinearly interpolated to the same spatial resolution of 0.1°. Even though WBMsed can disaggregate monthly time series into daily, TERRA (only available at monthly resolution, see ~~T~~table 1) is evaluated on monthly and annual time scales, whilst all other datasets are evaluated at daily, monthly and annual time scales.~~ in addition.~~ WBMsed simulations were run at 0.1° (~11km at the equator) spatial and daily and monthly temporal resolutions. Several WBMsed streamflow validation analyses have been reported previously (e.g., Cohen et al., 2022; Dunn et al., 2019; Cohen et al., 2014, 2013; Moragoda and Cohen, 2020), which indicate that the model represents the long-term average observed streamflow globally. It is important to note that this study assesses the precipitation datasets without calibration of the WBMsed model for each precipitation dataset, which could theoretically improve their performance in replicating observed river discharge. ~~Cohen et al. (2022) report R²=0.99 in 30-year average prediction against USGS gauge data and a global river dataset.~~

2.3. Observed river discharge from ground stations

Observed daily and monthly river discharge used to evaluate the hydrological model were obtained from the Global Runoff Data Centre (GRDC, 2023). The GRDC is an international data archive (https://www.bafg.de/GRDC/), which hosts data for over 10,000 hydrological stations. The number of stations with a length of record greater than 10 years during the evaluation period (1981-2019) are limited. Here, we consider stations with a minimum record length of 10 years, allowing for missing values within this period. Due to the spatial resolution of the input datasets and the model simulations (~11x11 km), we only consider stations with a catchment area of greater than 100 km^2. Overall, 1825 suitable stations were identified with daily and monthly records, largely in North and South America, Europe and Australia, with very few stations in Africa and Asia (Figure 1).

2.4. Evaluation metrics

Several methods are used to assess the modelled discharge using the streamflow observations: the Pearson correlation coefficient (CC, Eq. 1), Kling-Gupta Efficiency (KGE, Eq. 2) (Gupta et al., 2009), Root-Mean-Square Error (RMSE, Eq.3) and Percentage of bias (Pbias, Eq.4). CC measures the linear relationship between observed discharge and simulated discharge, focusing primarily on the degree of association between the two datasets. It is particularly useful for assessing the strength and direction of this relationship, highlighting how well the model captures the variability in discharge (Moazami et al., 2013). KGE is a comprehensive metric that evaluates the overall agreement between observed and simulated streamflow, considering similarities in variability, amplitude, and timing. It provides an assessment of the model's ability to capture both the magnitude and temporal dynamics of the observed discharge (Gupta et al., 2009). RMSE measures the average magnitude of the differences between observed and simulated discharge, providing a measure of the overall goodness of fit. Moreover, the percentage of bias is used to quantify the systematic overestimation or underestimation of discharge by the model compared to observations (Moazami et al., 2013). A KGE value of 1.0 indicates a perfect match between the observed and simulated discharge, whereas values lower than -0.41 show that the model is worse than using the mean of the observed discharge as a predictor (Knoben et al., 2019). For spatial comparison, the RMSE is normalised by the standard deviation of the observed data (NRMSE; Eq. 5).

$$CC = \frac{\sum_{i=1}^{N}(M_i - \bar{M}) * (O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N}(M_i - \bar{M})^2} * \sqrt{\sum_{i=1}^{N}(O_i - \bar{O})^2}} \tag{1}$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(O_i - M_i)^2}{N}} \tag{3}$$

$$Pbias = \frac{\sum_{i=1}^{N}(M_i - O_i)}{\sum_{i=1}^{N} O_i} * 100 \tag{4}$$

$$NRMSE = \frac{RMSE}{SD} * 100 \tag{5}$$

281 where r is the linear correlation between observed (O) and modelled (M) discharge and α and β are the variability

282 and bias ratios, respectively. The NRMSE and SD are the normalised RMSE and standard deviation, respectively.

283 To assess the performance of the precipitation datasets for representing daily hydrological extremes, the 90th <s>(Q10)</s>

284 and 10th <s>(Q90)</s> percentile are used, which indicates high and low flows, respectively. To derive high and low flow

285 thresholds from a daily flow time series, the data is first arranged in ascending order. The 90th percentile (Q10) is

286 <s>then</s> determined as the flow value above which just <s>90</s>10% of the daily flows lie, representing high-flow

287 conditions. <s>Similarly, t</s>The 10th percentile (Q90) represents the flow value <s>below</s>above which 90% of the daily

288 flows occur, indicating low-flow conditions. <s>The Q10 and Q90 represent the streamflow value that is equalled or</s>

289 <s>exceeded 10% and 90% of the time, respectively.</s>

3.   Results

291 <s>The performance of the six different precipitation datasets in simulating discharge is evaluated at annual, monthly</s>

292 <s>and daily time steps and for extremes during the period 1983-2019. The WBMsed output discharge forced by the</s>

293 <s>six precipitation datasets is referred to as ERA5, CHIRPS, MSWEP, TERRA, CPCU, and PERCCDR below.</s>

3.1. Performance of the six precipitation datasets for annual discharge prediction

295 The temporal correlation coefficient (CC) between the observed and simulated annual discharge based on the six

296 precipitation datasets is summarised in Figure 1. Most of the datasets, particularly ERA5, MSWEP, and CHIRPS,

297 showed a high CC in basins of Europe (e.g., Danube basin), South America (e.g., Rio de la Plata-Parana), North

298 America and Australia (e.g., Murray-Darling). MSWEP and ERA5 showed the highest CC for 34% and 32% of

299 the stations, respectively, followed by CPCU and CHIRPS. The TERRA and PERCCDR were the least well-

300 performing datasets with lower CC overall, and a higher CC than other datasets for less than 9% of stations. The

301 median CC of MSWEP and ERA5 is 0.82 and 0.8, respectively. MSWEP and TERRA showed lower Pbias and

302 NRMSE compared to the other datasets (Figures S1 and S2). ERA5 and PERCCDR showed a high NRMSE (up

303 to 247%) and Pbias (up to 99%) for more than 46% of stations. Similar to the CC, ERA5 and MSWEP

304 outperformed the other datasets for KGE, with higher values for 32% and 27% of stations, respectively. The

305 performance of MSWEP and ERA5 is higher in basins of Europe, South America, and Australia compared to Asia

306 and Africa. The median KGE values of ERA5 and MSWEP are 0.33 and 0.32, respectively (Figure 2). The

307 PERCCDR and CPU demonstrate high KGE only in about 9% of the stations, with median values of 0.10 and

308 0.13, respectively. Based on the annual CC and KGE, there is no single precipitation dataset that is best

309 everywhere, and even the least well-performing dataset overall shows better performance in some stations (Figure

310 3). Figure 3 summarizes the spatial representation of precipitation dataset performance, highlighting the individual

311 datasets exhibiting the highest CC and KGE values at each observation point.

312

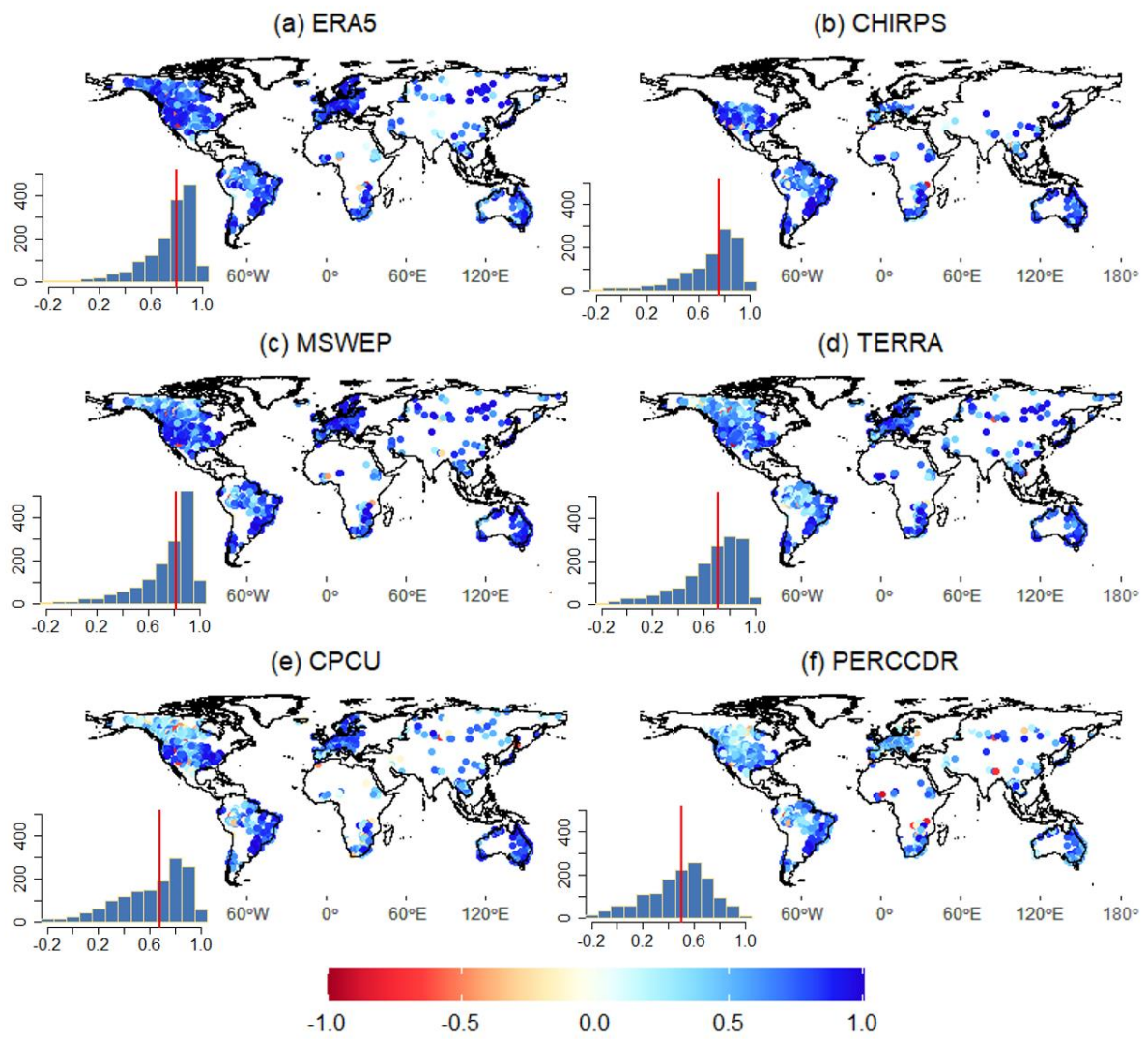
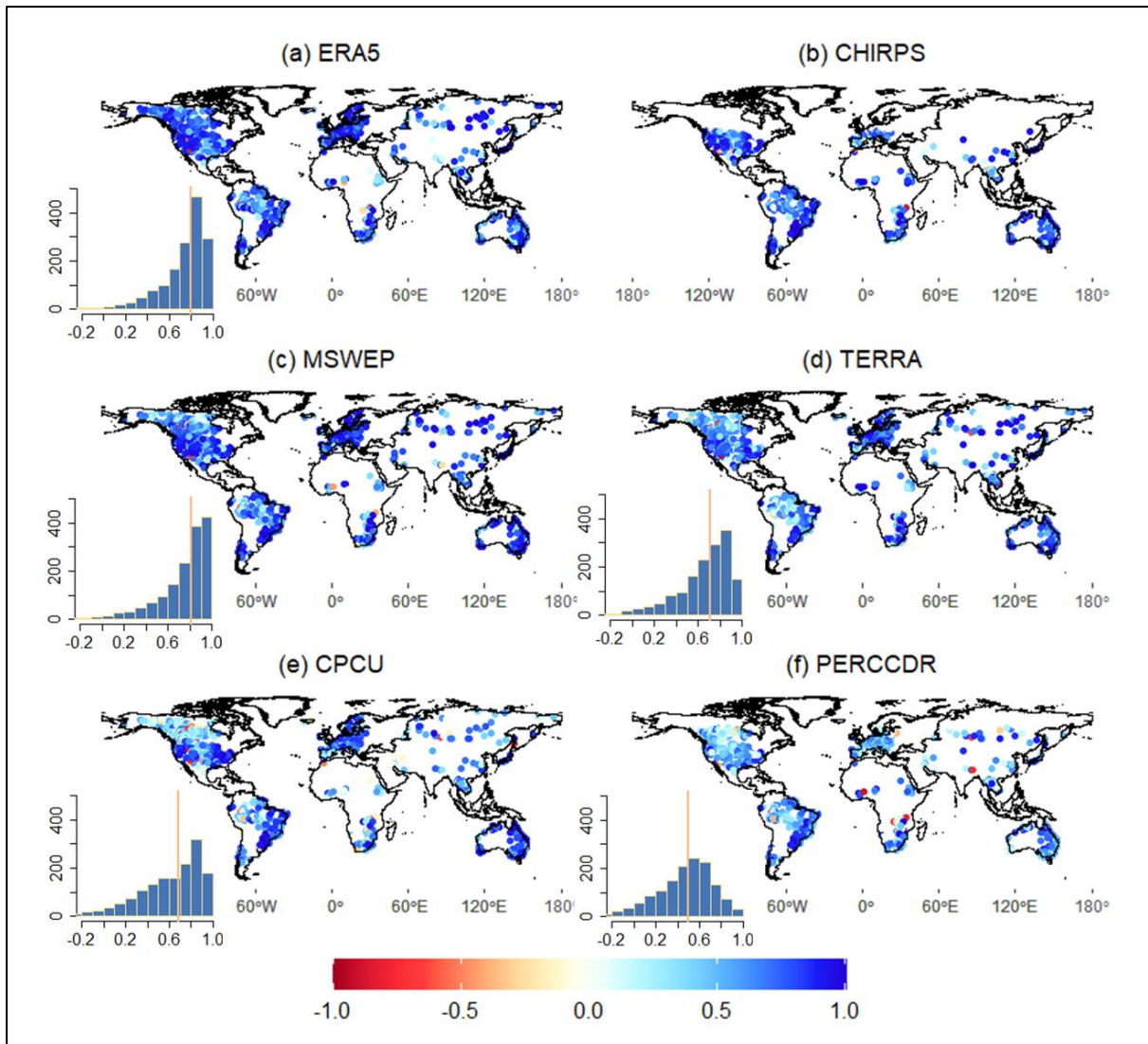



Figure 1: Correlation (CC) between annual observed and modelled streamflow data using a) ERA5, b) CHIRPS, c) MSWEP, d) TERRA, e) CPCU and f) PERCCDR precipitation datasets. The inset histograms show the frequency distribution (y-axis) of the ~~monthly~~ annual CC (x-axis), with the ~~yellow~~ red vertical line indicating the median value.

319

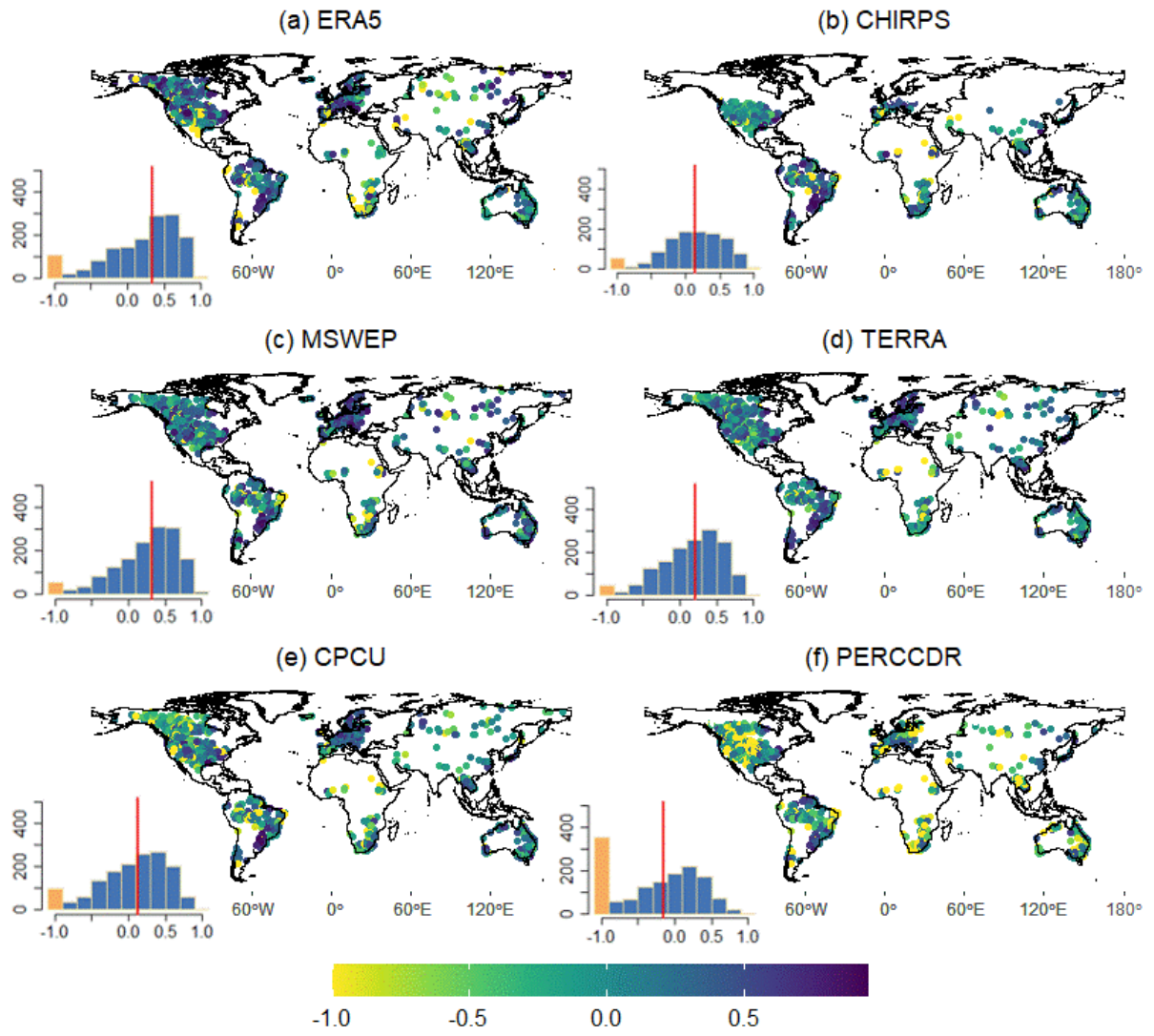
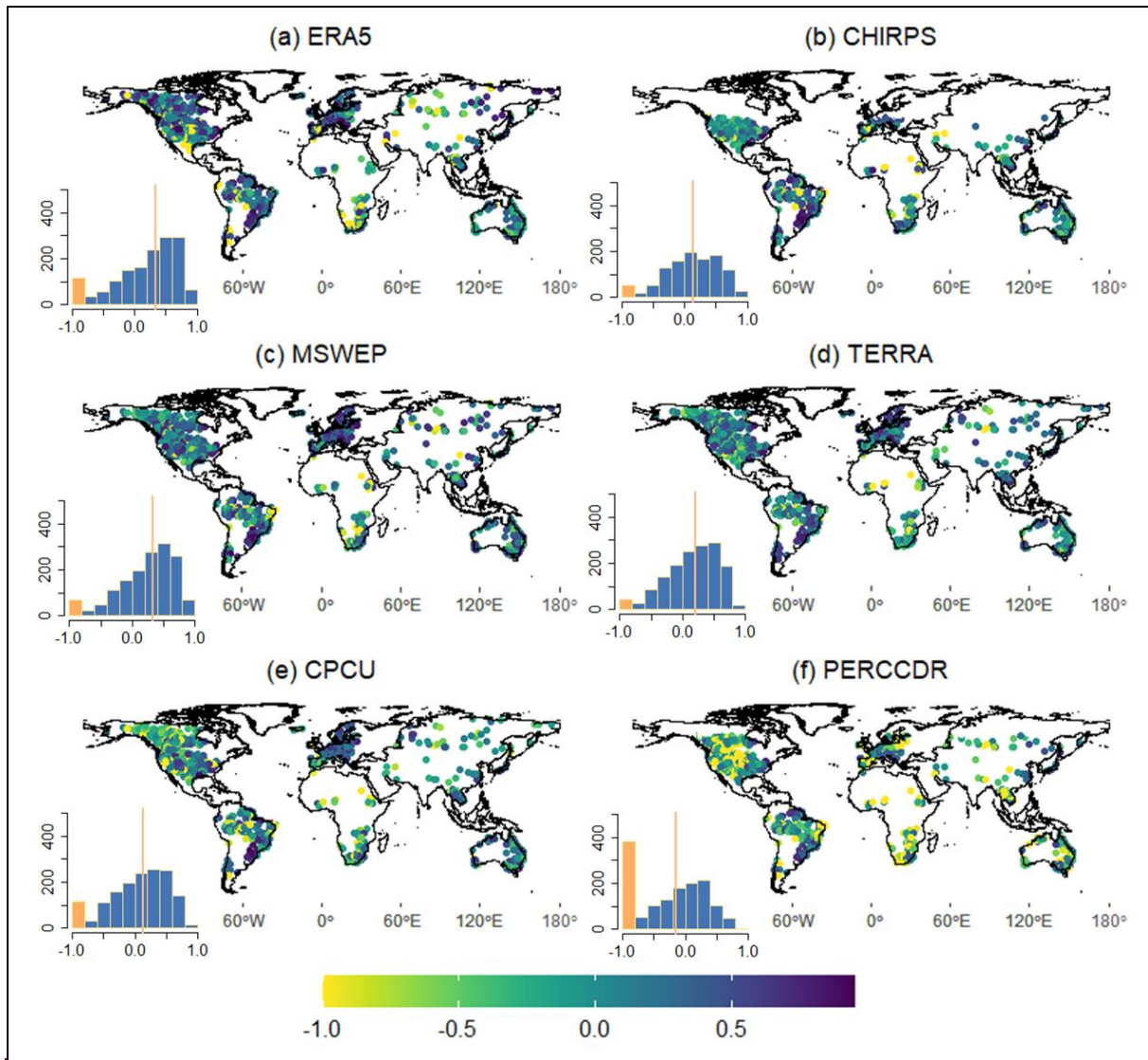



Figure 2: KGE between observed and modelled annual streamflow based on a) ERA5, b) CHIRPS, c) MSWEP, d) TERRA, e) CPCU, and f) PERCCDR precipitation datasets. KGE values below -0.41 indicate bad model performance than using observed discharge mean as a predictor. The inset histograms show the frequency distribution (y-axis) of the ~~monthly~~ annual KGE (x-axis). KGE values lower than -1 are highlighted in ~~yellow~~orange. The ~~yellow~~ red vertical line indicates the median value.

325

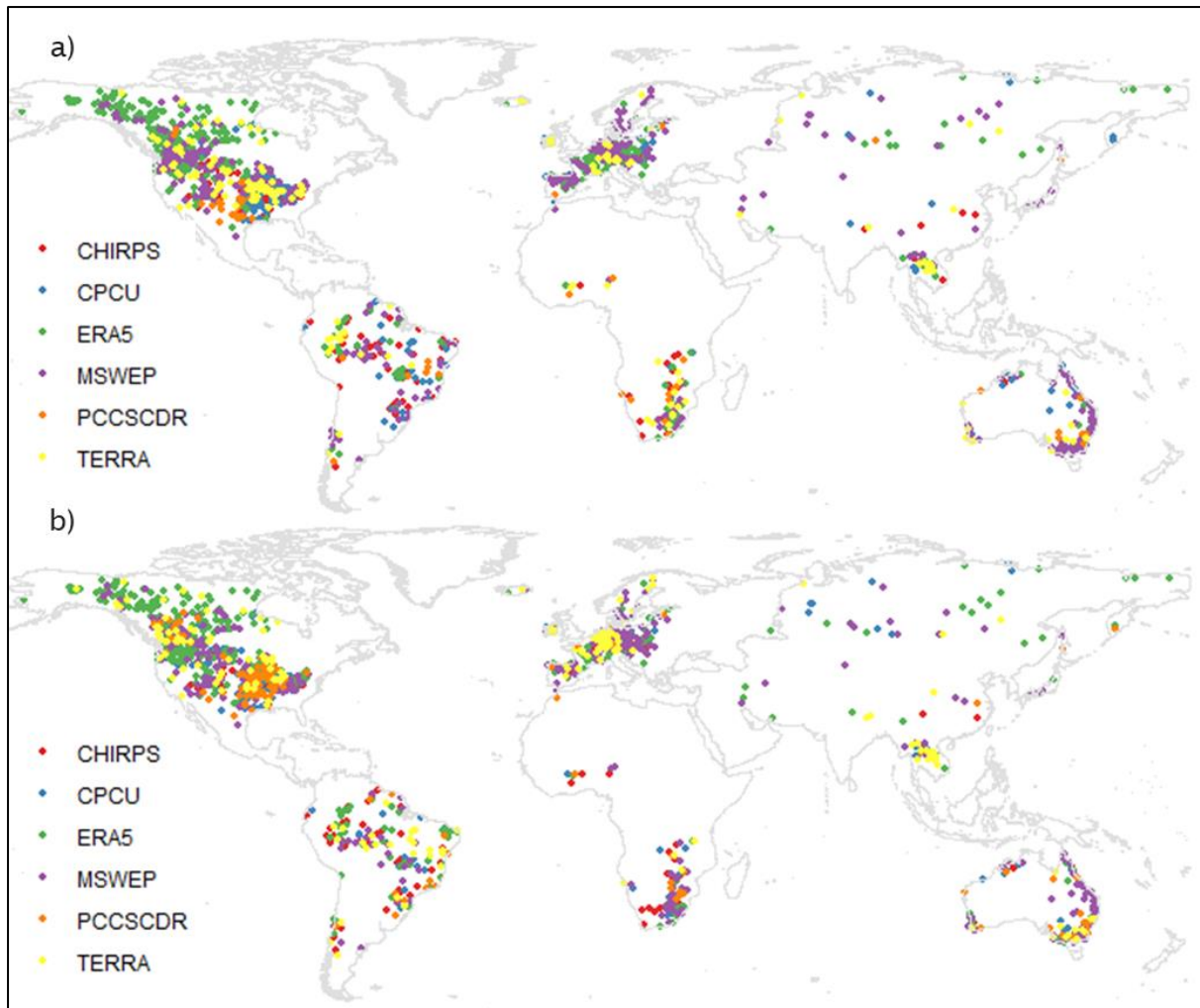



Figure 3: The best performing precipitation dataset (ERA5, CHIRPS, MSWEP, TERRA, CPCU, and PERCCDR) at each of the observed discharge stations based on annual CC (a) and KGE (b).

3.2. Performance of the six precipitation datasets for monthly discharge predictions

The six precipitation datasets consistently demonstrate high CC at a monthly scale in large parts of the world, except in some rivers of Canada and Australia (Figure 4). The monthly CC, similar to the annual CC, shows a relatively better performance of MSWEP with a median CC of 0.76. TERRA is the second-best with a median CC of 0.69. MSWEP and TERRA show a higher CC than other datasets in 35% and 28% of the stations, respectively. ERA5 and CHIRPS are ranked as the third and fourth datasets with a median CC of 0.71 and 0.75, respectively. CPCU and PERCCDR are the least well-performing datasets, which only show the highest CC in less than 6% of the stations with a median CC of 0.67 and 0.56, respectively.

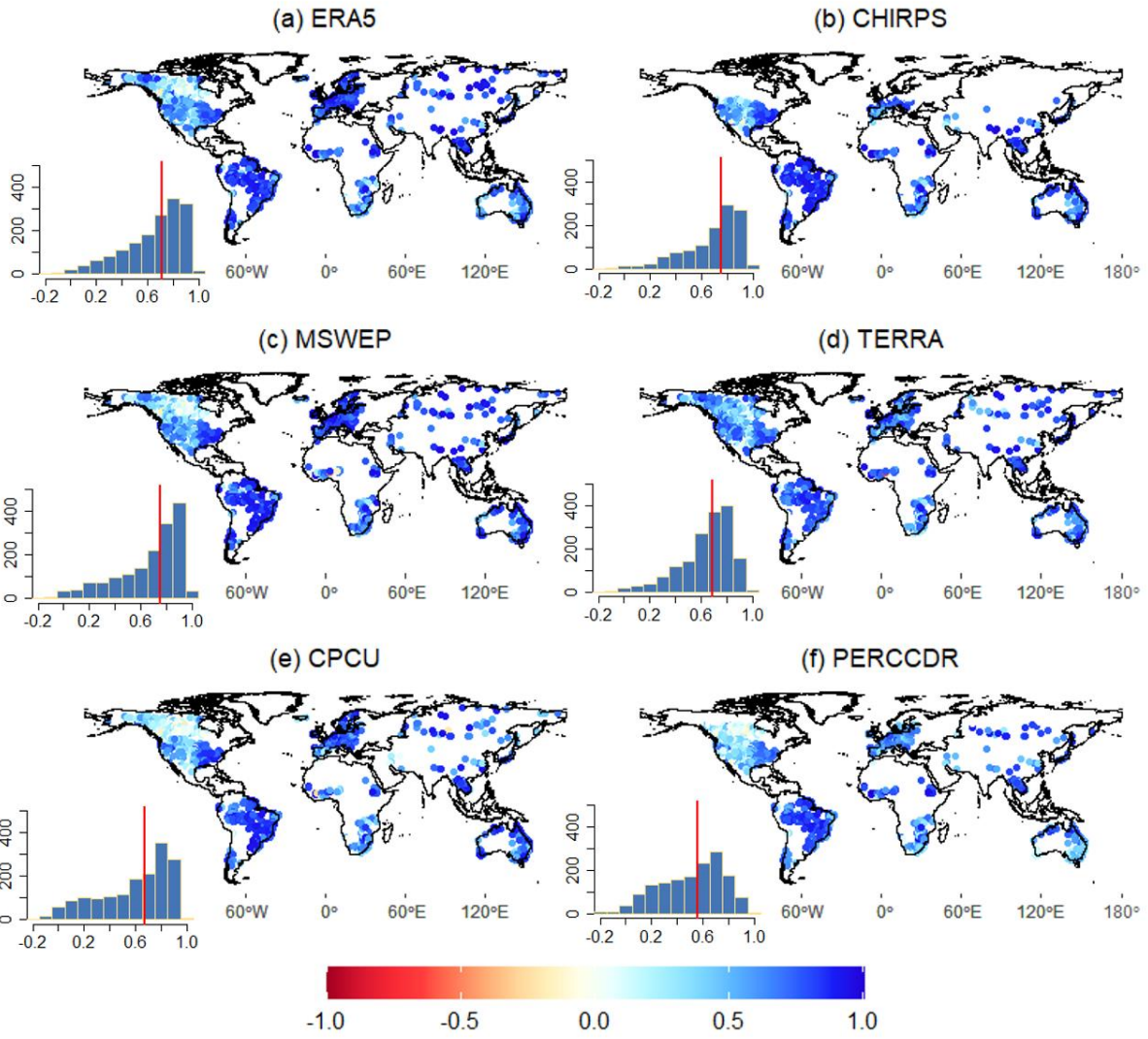
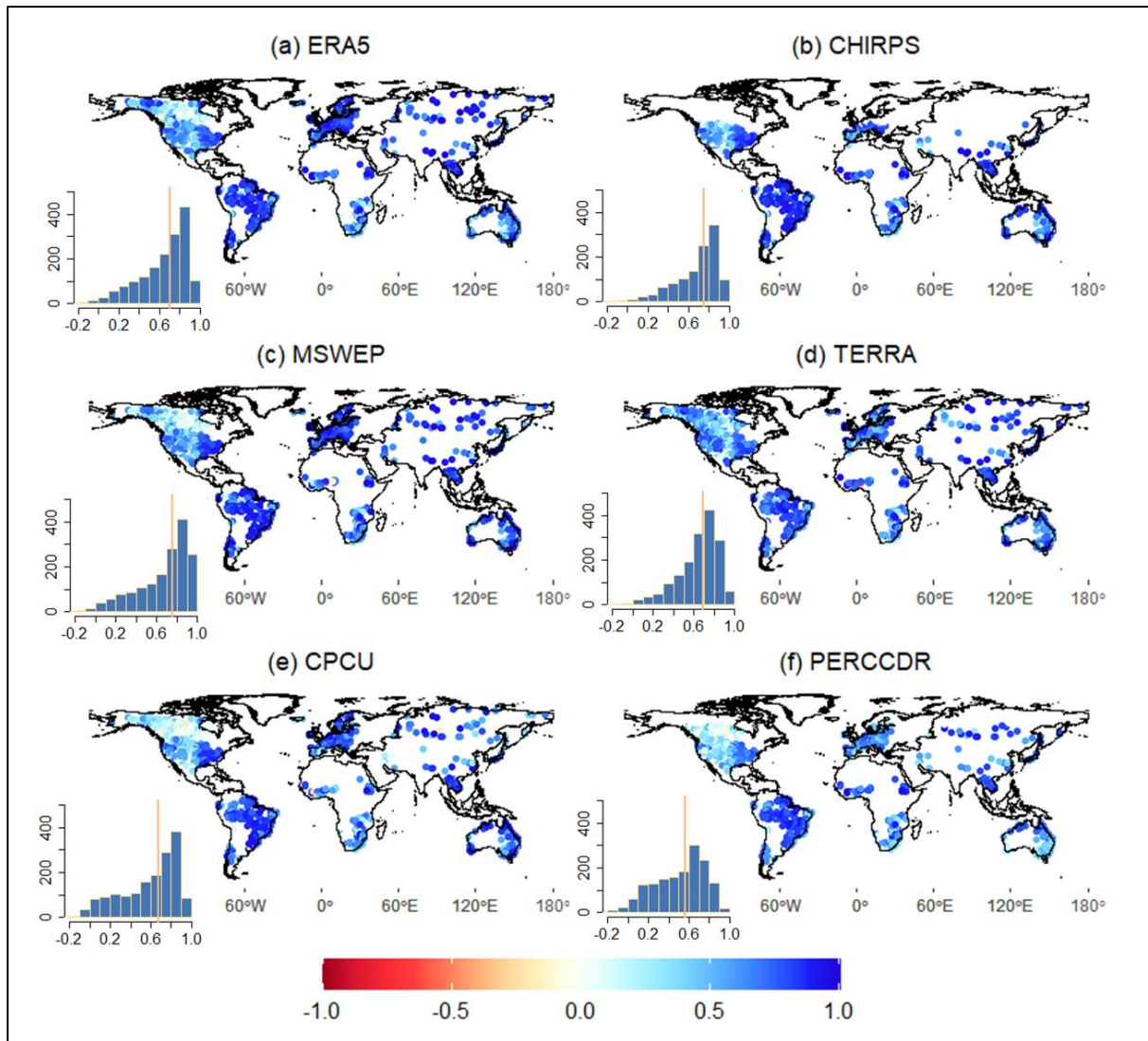



Figure 4: Correlation (CC) between monthly observed and modelled streamflow data based on a) ERA5, b) CHIRPS, c) MSWEP, d) TERRA, e) CPCU and f) PERCCDR precipitation datasets. The inset histograms show the frequency distribution (y-axis) of the monthly CC (x-axis), with the ~~yellow~~ red vertical line indicating the median value.

The monthly KGE also indicates the better performance of ERA5 and MSWEP for 26% and 24% of stations, respectively (Figure 5). MSWEP showed a lower Pbias and NRMSE than all datasets, except in 5% of the stations (Figures S3 and S4). Compared to MSWEP, ERA5 showed a larger Pbias and NRMSE in 15% and 19% of the stations. TERRA, a third-best performing dataset based on KGE (18% of stations), shows a lower monthly Pbias and RMSE in 85% of the stations compared to CHIRPS, ERA5, and PERCCDR. Compared to all datasets, the PERCCDR showed a higher NRMSE and Pbias in 55% and 28% of the stations, respectively. ~~The spatial representation of precipitation dataset performance is summarized in Figure S5, highlighting the regions where individual datasets demonstrate higher monthly CC and KGE values.~~

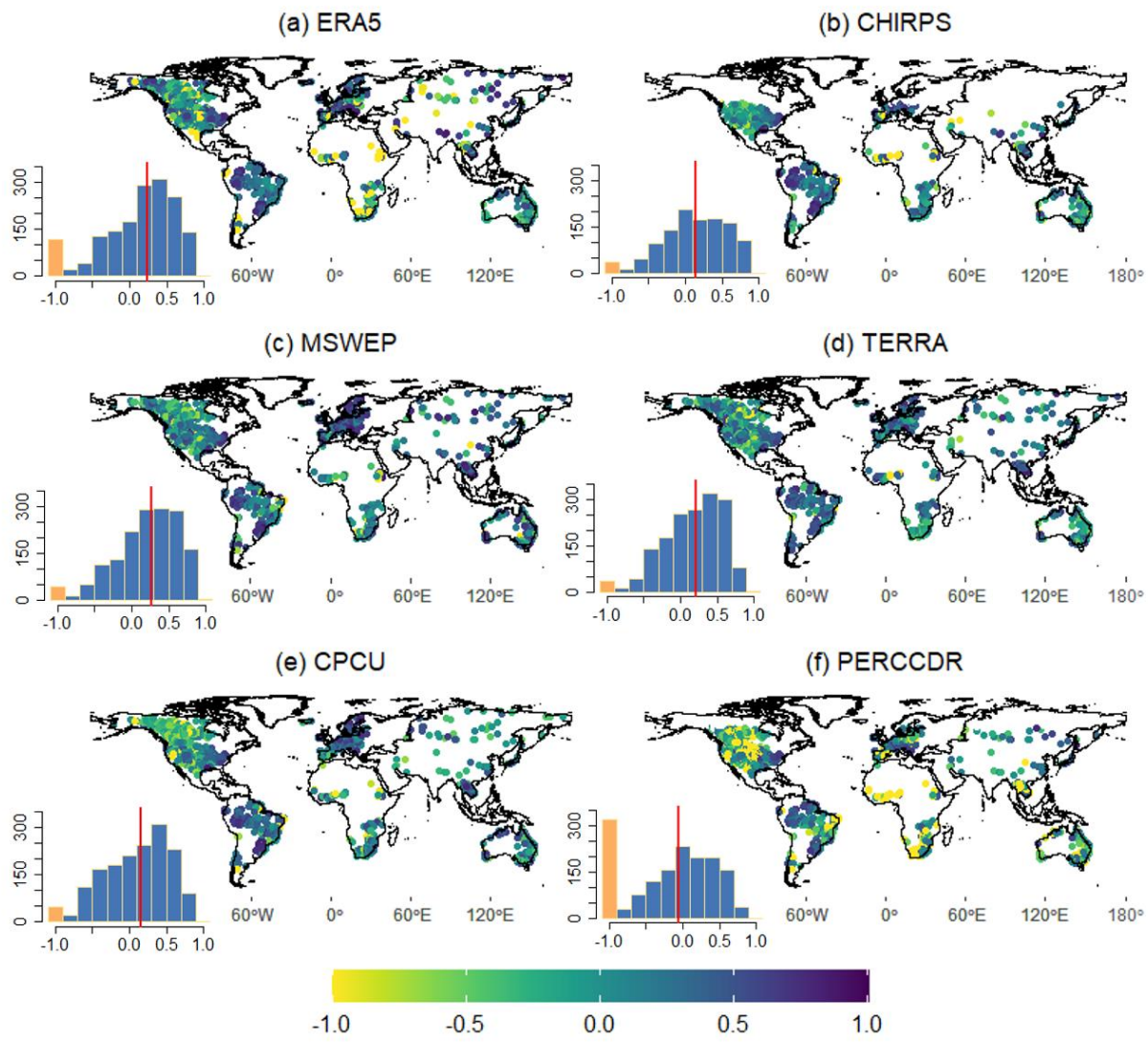
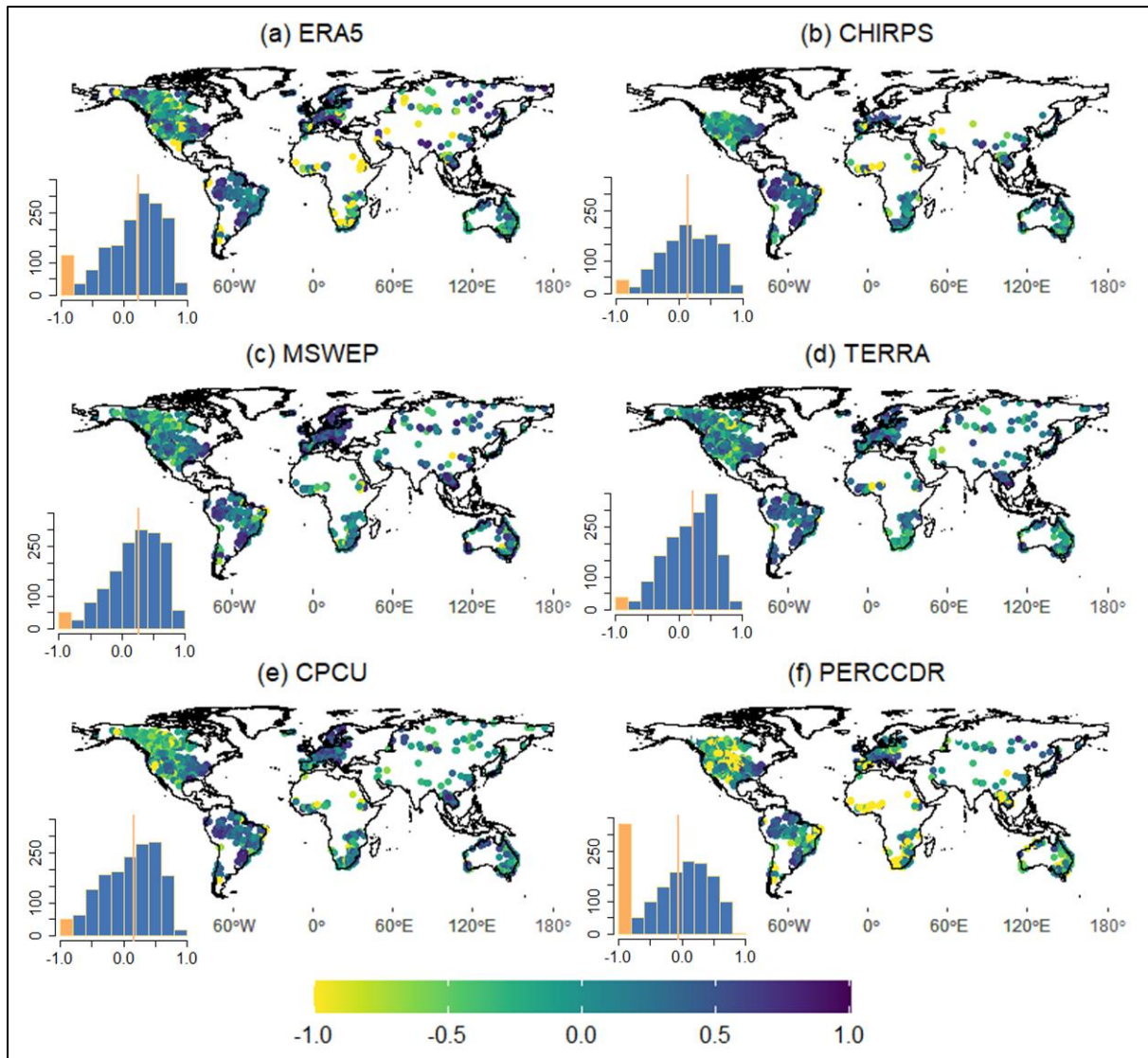
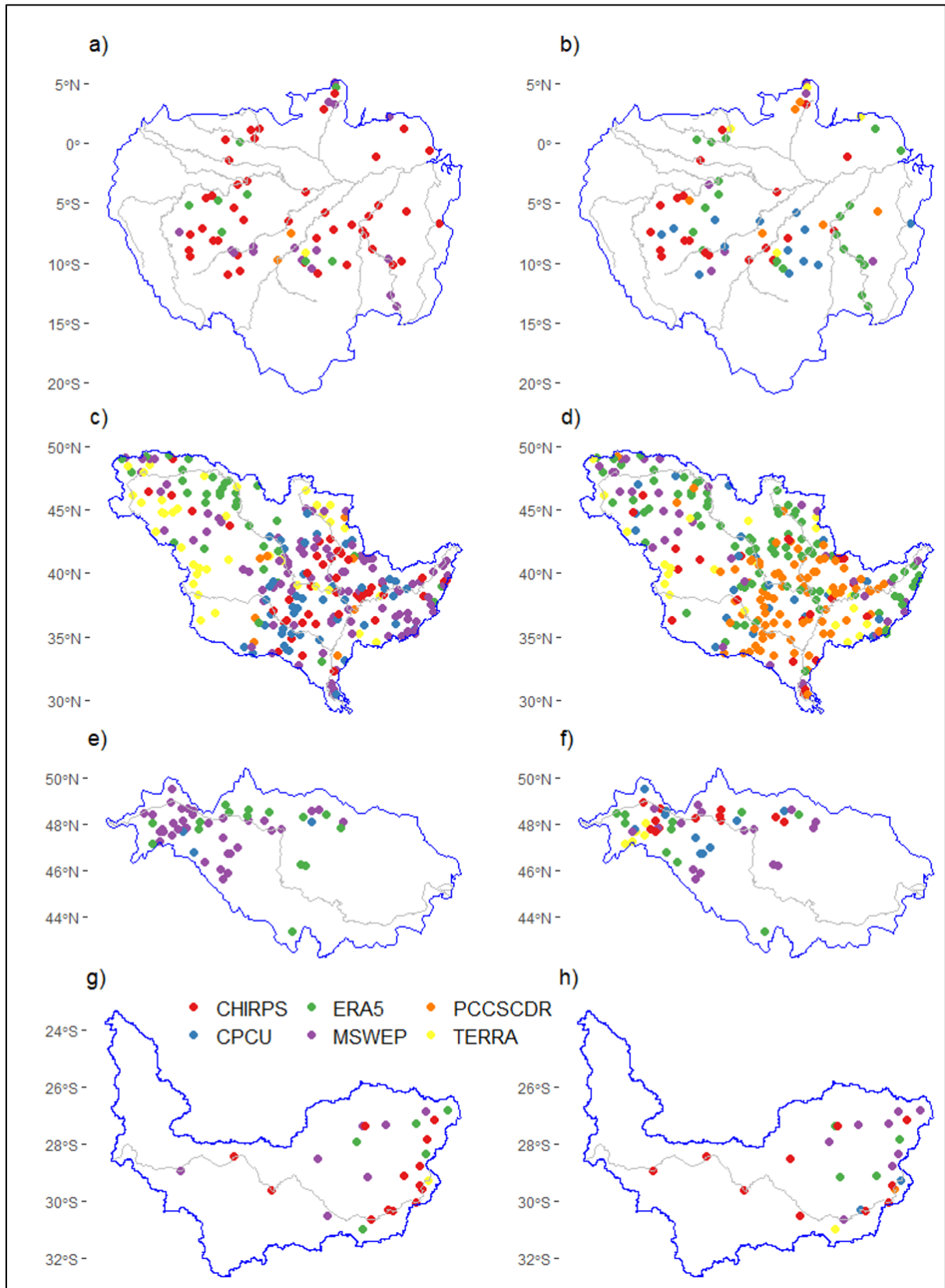



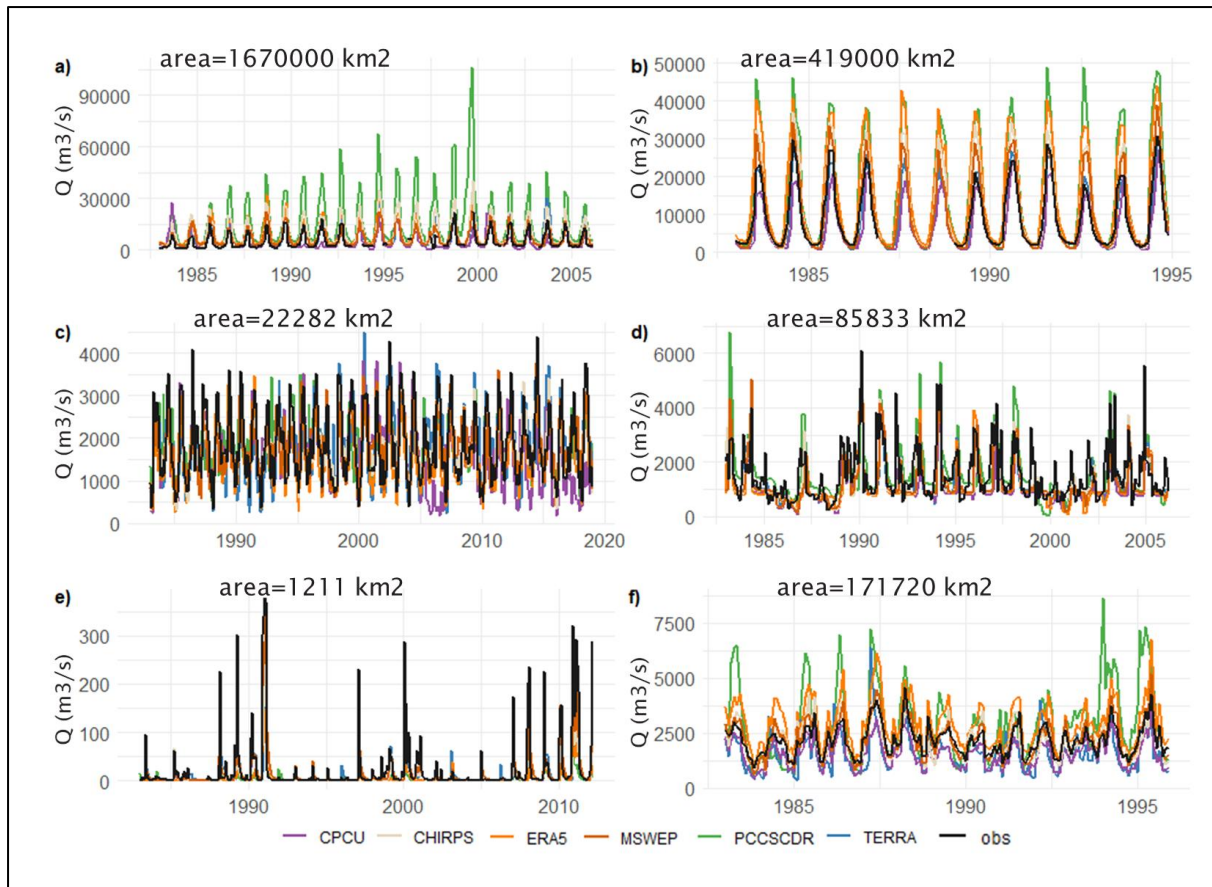
Figure 5: Monthly KGE values between observed and modelled streamflow based on a) ERA5, b) CHIRPS, c) MSWEP, d) TERRA, e) CPCU and f) PERCCDR precipitation datasets. KGE values below -0.41 indicate model performance that is worse than using the observed discharge mean as a predictor. The inset histograms show the frequency distribution (y-axis) of the monthly KGE (x-axis). KGE values lower than -1 are highlighted in ~~yellow~~orange, with the red vertical line indicating the median value.

The spatial representation of the six precipitation datasets in the Amazon, Mississippi, Danube, and Orange River basins is summarised in Figure 6, highlighting the individual datasets exhibiting the highest CC and KGE values at each hydrological station. In the Amazon basin, ERA5 (31%) and CHIRPS (29%) emerge as the top performers, while PERCCDR (8%) and TERRA (5%) rank lower among the precipitation datasets. In the Mississippi basin, MSWEP leads with higher CC in 37% of stations, and ERA5 ~~holds the top products with~~exhibits higher KGE in 31% of the stations. Notably, PCCSCDR displays higher KGE values than MSWEP, TERRA, CHIRPS, and CPCU in 30% of Mississippi stations. Across the Danube basin, MSWEP outperform the other products with a higher CC in 66% of stations and KGE in 30% of the stations, while TERRA and CPCU are the least performing products. Furthermore, CHIRPS, in 52% of stations based on CC and 37% based on KGE, outperforming other

datasets in the Orange River basin. In Orange, MSWEP ranks second with higher KGE and CC in about 27% of stations, while TERRA and PCCSCDR are the least performing datasets.

372 ~~Figure 6 shows~~Table 2 summarises the ~~the time series of~~ monthly KGE between observed and modelled
373 streamflow, based on the six precipitation datasets, for selected locations in basins of Africa (Niger, Lokoja), Asia
374 (Mekong, Khong-Chiam), South America (Amazon, Missao-Icana), North America (Mississippi, Savannah),
375 Australia (North East Coast, Mirani-Weir), and Europe (Danube, Dunaalmas). The basins were chosen to
376 represent a ~~good~~ diverse range of climatic regions and drainage areas where there was availability of a long time
377 series of observed data (Figure S5). In Niger, the observed monthly flow and variability at Lokoja station are very
378 well reproduced by CHIRPS and TERRA with a CC of 0.88 and 0.85, respectively (Figure S5a). Even though
379 CPCU showed a lower CC (0.64) at Lokoja, it showed a higher KGE (0.62) and lower Pbias (0.4%) compared to
380 the other products. At Lokoja, PERCCDR is the least well-performing dataset with the highest RMSE and Pbias
381 and lowest KGE. The monthly variability at the Khong-Chiam station is reproduced by all the precipitation
382 products with a CC of greater than 0.91, with MSWEP and TERRA showing the lowest bias and RMSE. ERA5
383 and CHIRPS performed well at station Missao-Icana in the Amazon with a CC of 0.9 and RMSE of about 610
384 m3/s. For stations Savannah, Mirani-Weir, and Dunaalmas, MSWEP is the best product with higher CC (> 0.72)
385 and KGE (> 0.62) and lower Pbias and RMSE (Figures S5d - S5f).

386 Table 2. KGE of monthly predictions for selected stations in basins of Africa (Niger), Asia (Mekong), South
387 America (Amazon), North America (Mississippi), Australia (North East Coast), and Europe (Danube).

Basin	Stations name	Longitude	Latitude	Catchment area (km^2)	ERA5	CHIRPS	MSWEP	TERRA	CPCU	PCCSCDR
Niger	Lokoja	6.8	7.8	1670000	0.21	-0.1	0.60	0.34	0.62	-0.99
Mekong	Khong Chiam	105.5	15.3	419000	0.13	0.56	0.70	0.91	0.70	-0.04
Amazon	Missao Icana	-67.6	1.1	22282	0.71	0.78	0.73	0.72	0.61	0.65
Mississippi	Savannah	-88.3	35.2	85833	0.59	0.65	0.67	0.66	0.53	0.66
North East Coast	Mirani-Weir	148.8	-21.2	1211	-0.1	0.38	0.62	0.44	0.46	-0.05
Danube	Dunaalmas	18.3	47.7	171720	0.34	0.73	0.78	0.52	0.71	-0.49

388

3.3. Performance of the precipitation datasets for daily and daily extreme discharge predictions

Based on the daily evaluation, MSWEP followed by ERA5 show a higher CC in more than 50% of the stations with median values of 0.41 and 0.39, respectively (Figure 7). ERA5 and MSWEP performed well in 31% and 31% of the stations with high KGE values (Figure 8). Similar to the monthly evaluation, PERCCDR shows poorer performance (lower CC and KGE, higher biases and errors) in almost 95% of the stations. Even though ERA5 showed a higher CC and KGE in 30% of the stations it shows a higher NRMSE (up to 250%) and Pbias (up to 100%) in 20% and 30% of the stations (Figures S6 and S7). Overall, MSWEP and CHIRPS showed lower NRMSE and Pbias compared to the other products. The CC and KGE of all the products (except CHIRPS) are lower in North America compared to stations in South America, Europe, and Australia. The spatial representation of precipitation dataset performance, highlighting the individual datasets exhibiting the highest daily CC and KGE values at each observation point, is provided in Figure S9. Additionally, Figure S10 depicts the spatial representation of each precipitation dataset for the Amazon, Mississippi, Danube, and Orange River Basins. In the Mississippi basin, ERA5 exhibited the highest KGE and CC values, followed by MSWEP and CPCU (Figure S10). In the Amazon basin, ERA5 and CHIRPS displayed the highest KGE and CC values compared to the other datasets. For the Danube basin, CPCU followed by MSWEP emerged as the best precipitation product relative to

409   ERA5, PCCSCDR, and CHIRPS. In the Orange River Basin, MSWEP based on CC and CHIRPS based on KGE

410   were the top-performing products, while PCCSCDR performed the least.

411

412

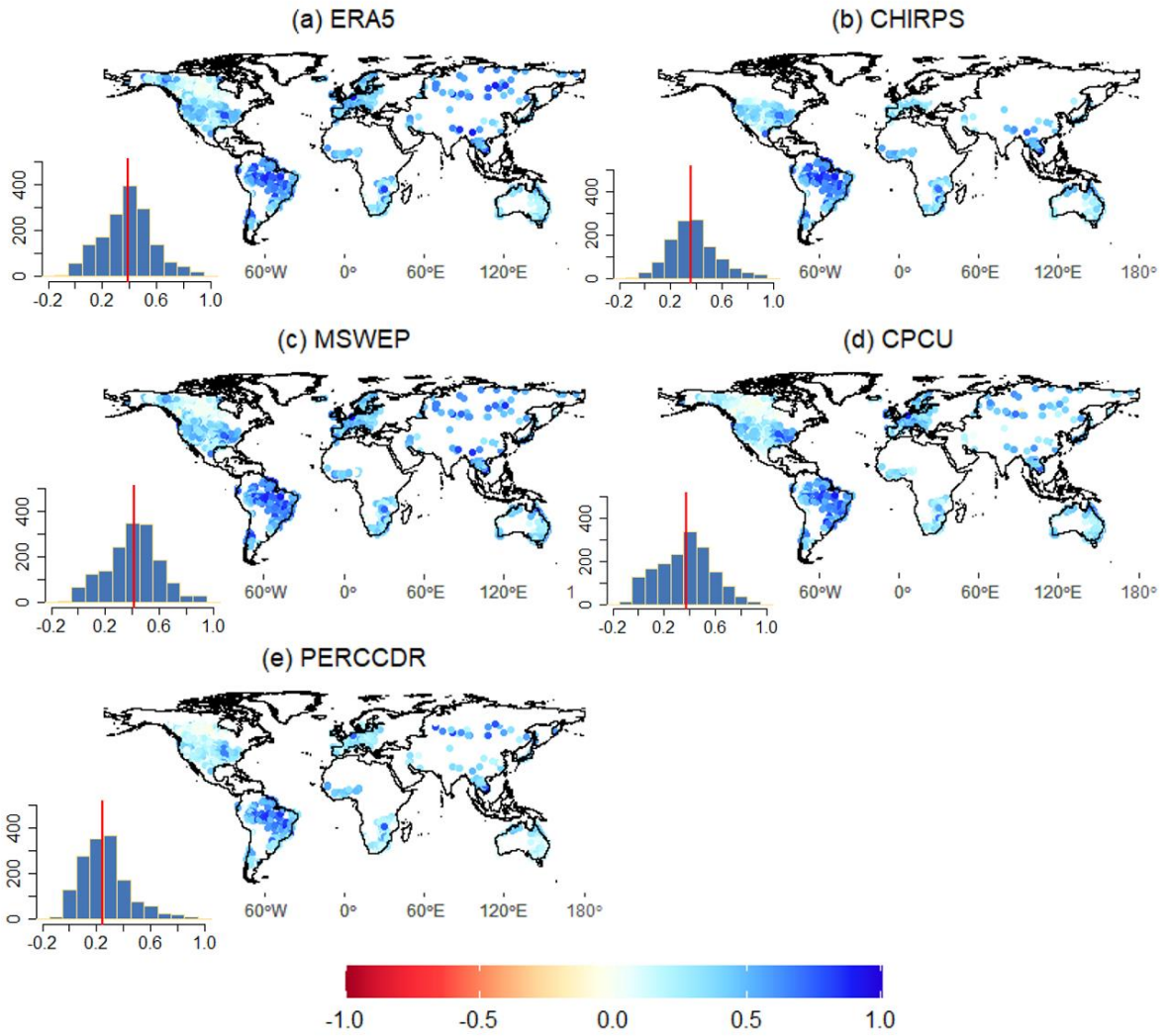
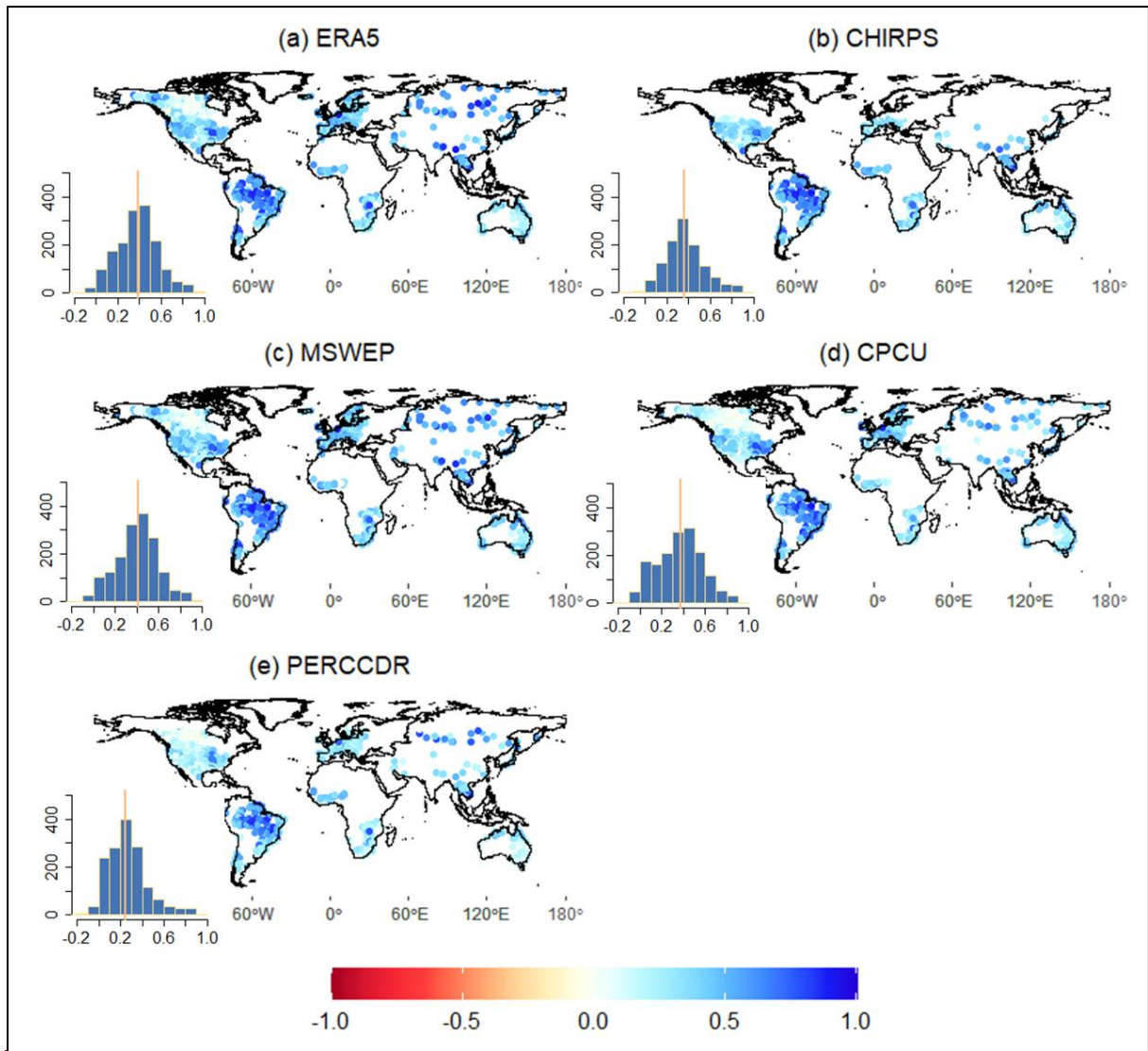



**Figure 7: Correlation (CC) between daily observed and modelled streamflow data using a) ERA5, b) CHIRPS, c)
MSWEP, d) CPCU and e) PERCCDR precipitation datasets. The inset histograms show the frequency distribution (y-
axis) of the daily CC (x-axis), with the ~~yellow~~ red vertical line indicating the median value.**

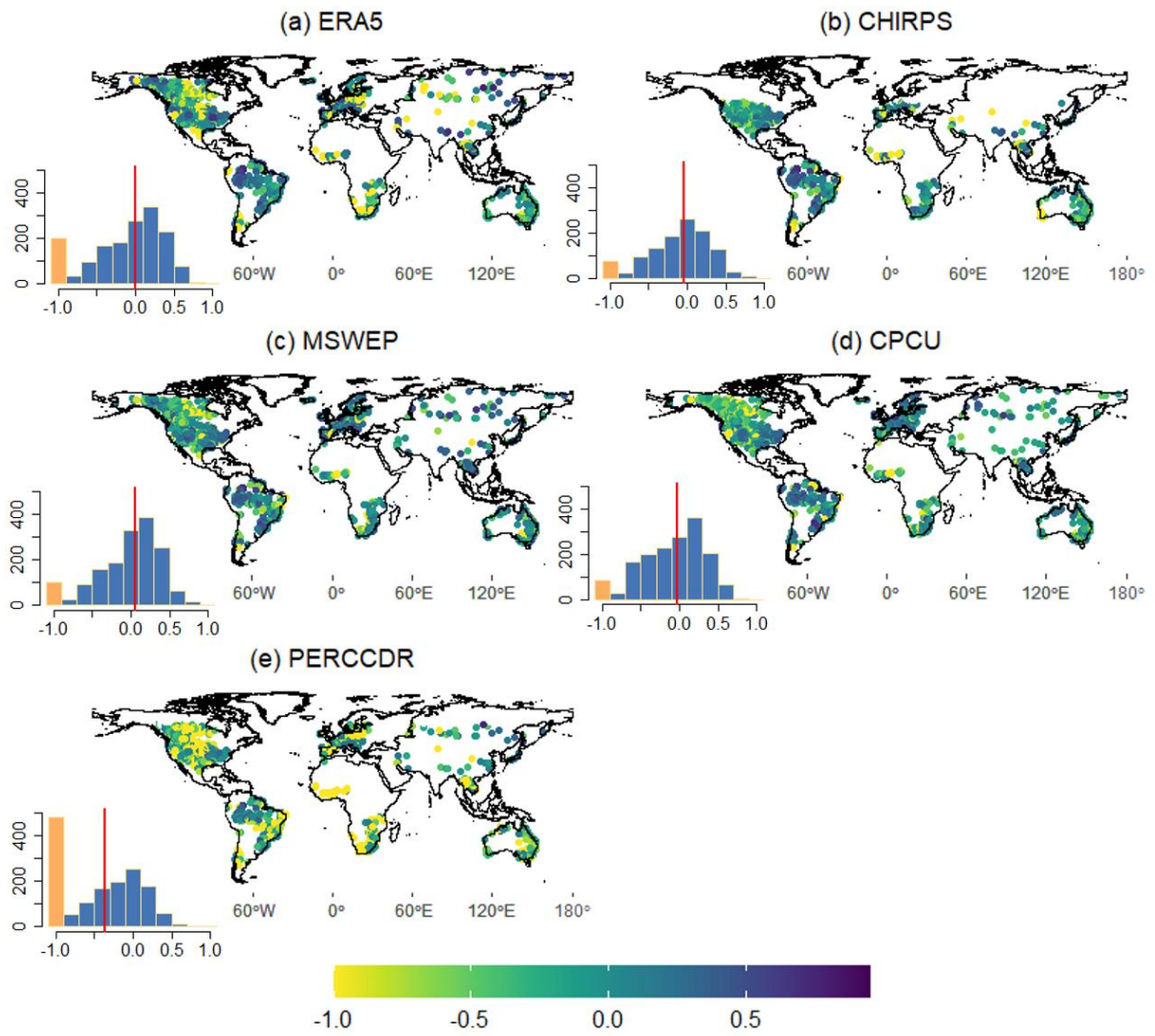
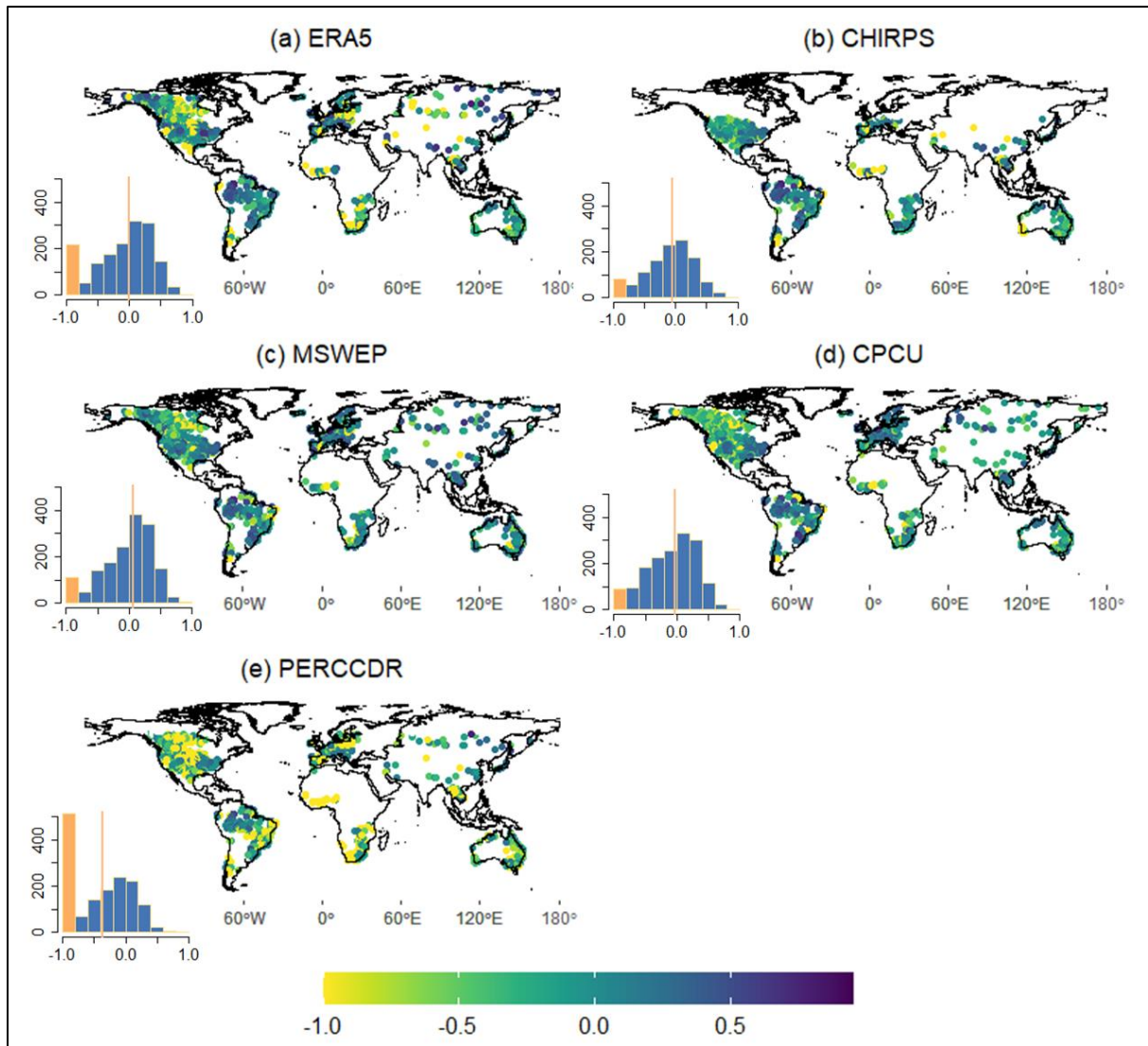



Figure 8: Daily KGE values between observed and modelled streamflow based on a) ERA5, b) CHIRPS, c) MSWEP, d) CPCU, and e) PERCCDR precipitation datasets. KGE values below -0.41 indicate bad model performance than using observed discharge mean as a predictor. The inset histograms show the frequency distribution (y-axis) of the daily KGE (y-axis). KGE values lower than -1 are highlighted in ~~yellow~~orange, with the~~. The yellow~~ ~~red~~ vertical line ~~indicates~~ indicating the median value.

The performance of the daily precipitation products is also assessed for daily extremes in terms of the Q10 and Q90 values. Based on the CC, MSWEP is the best-performing dataset for Q10 (Figure 9) and Q90 (Figure S8). ~~and Q10 (Figure S9)~~. For Q10, MSWEP and CPCU exhibited a higher CC than other datasets at 38% and 32% of the stations, respectively. Similarly, for Q90, MSWEP and ERA demonstrated a higher CC compared to other datasets at 35% and 30% of the stations. The median CC for Q10 (Q90) is 0.32 (0.41), 0.28 (0.36), 0.27 (0.35), 0.26 (0.38), and 0.16 (0.23) for MSWEP, CPCU, CHIRPS, ERA5, CHIRPS, and PERCCDR, respectively. Similar to the annual, monthly and daily evaluations, PERCCDR showed poor performance for the two extremes (Q90 and Q10). Overall, the performance of the datasets is lower for extremes compared to the annual, monthly and daily scales.

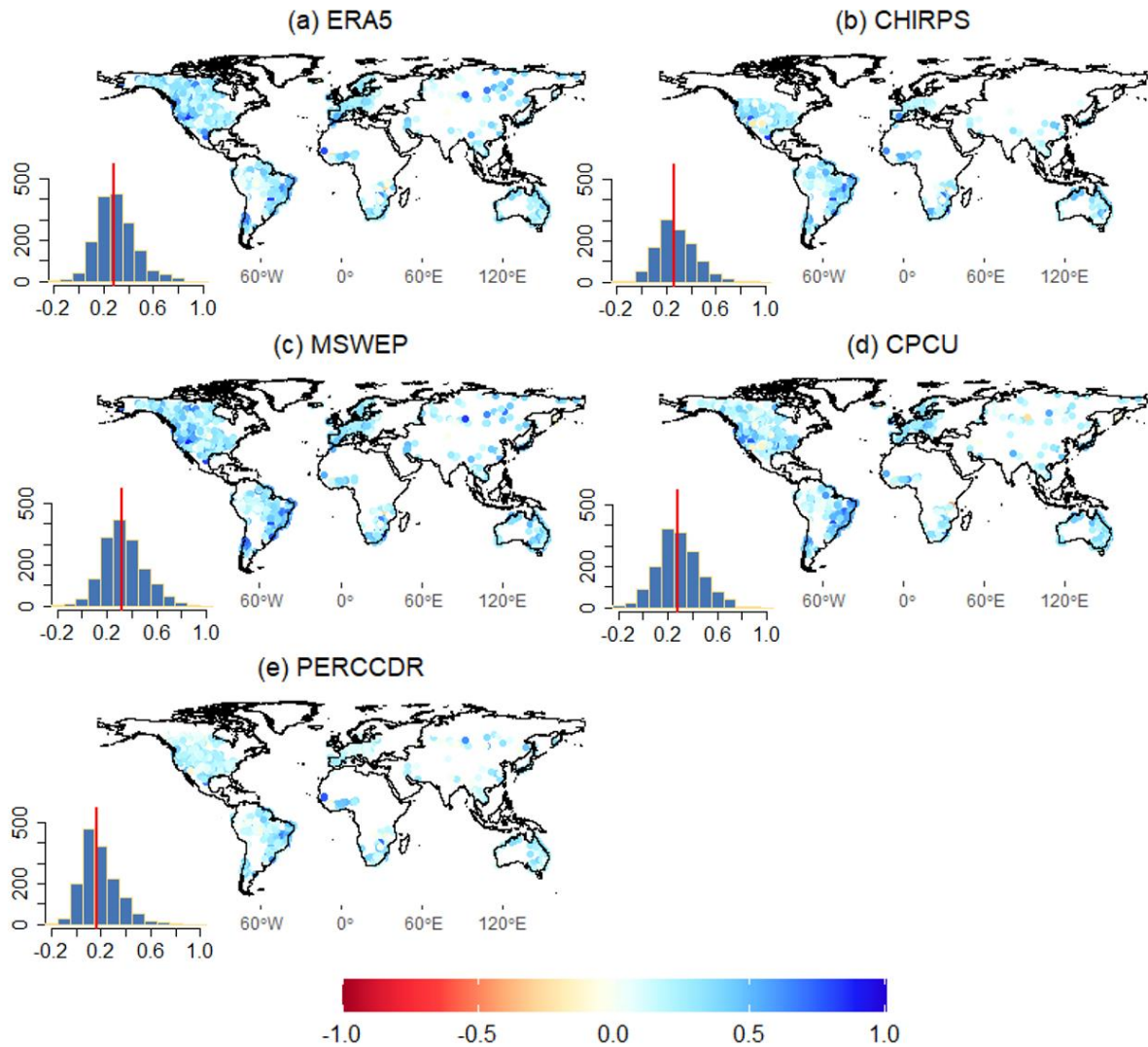
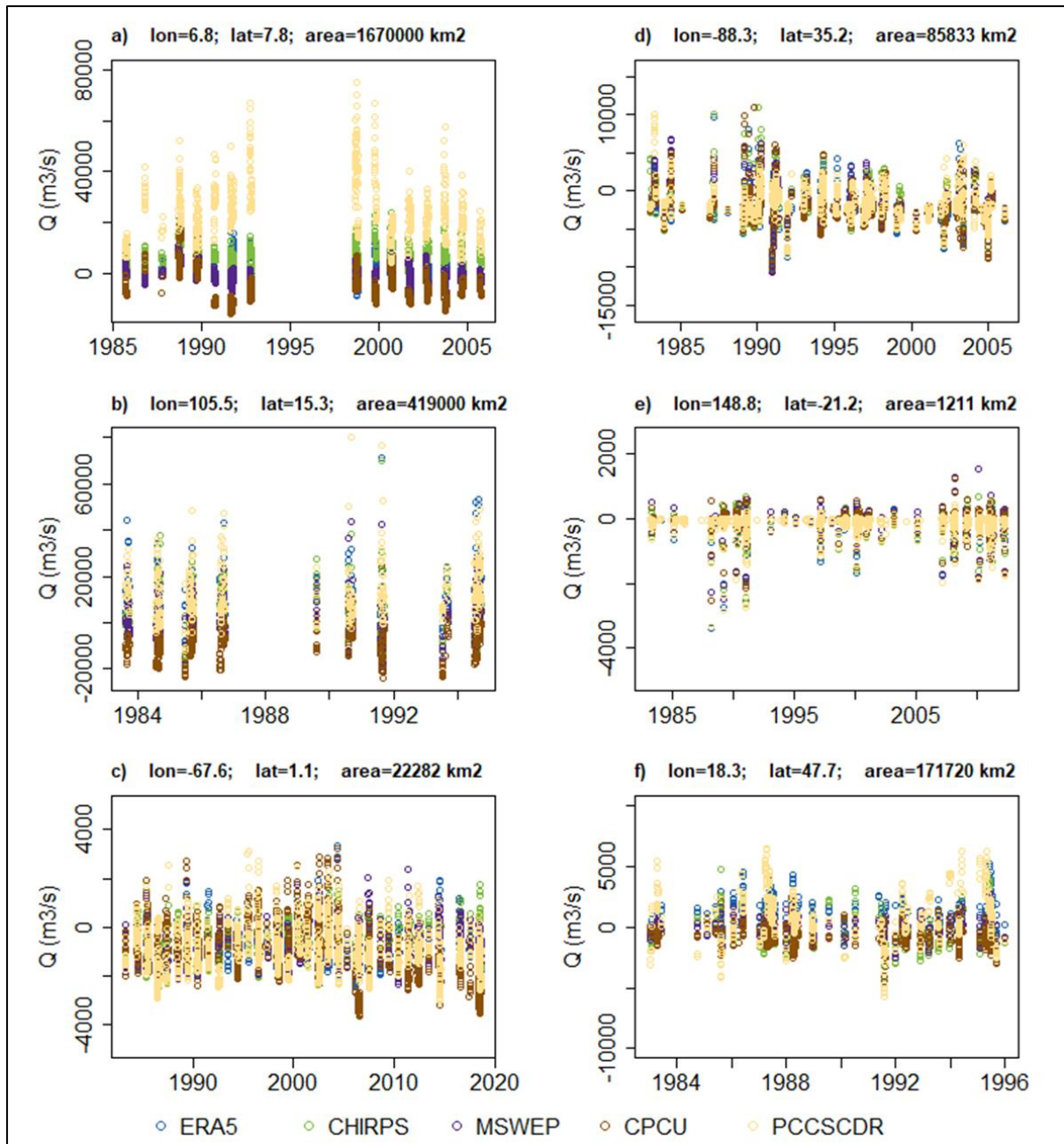



Figure 9: Correlation (CC) between observed and modelled daily extremes (Q10, high flow) streamflow data a) ERA5, b) CHIRPS, c) MSWEP, d) CPCU and e) PERCCDR precipitation datasets. The inset histograms show the frequency distribution (y-axis) of the daily Q10 CC (x-axis), with the red vertical line indicating the median value.

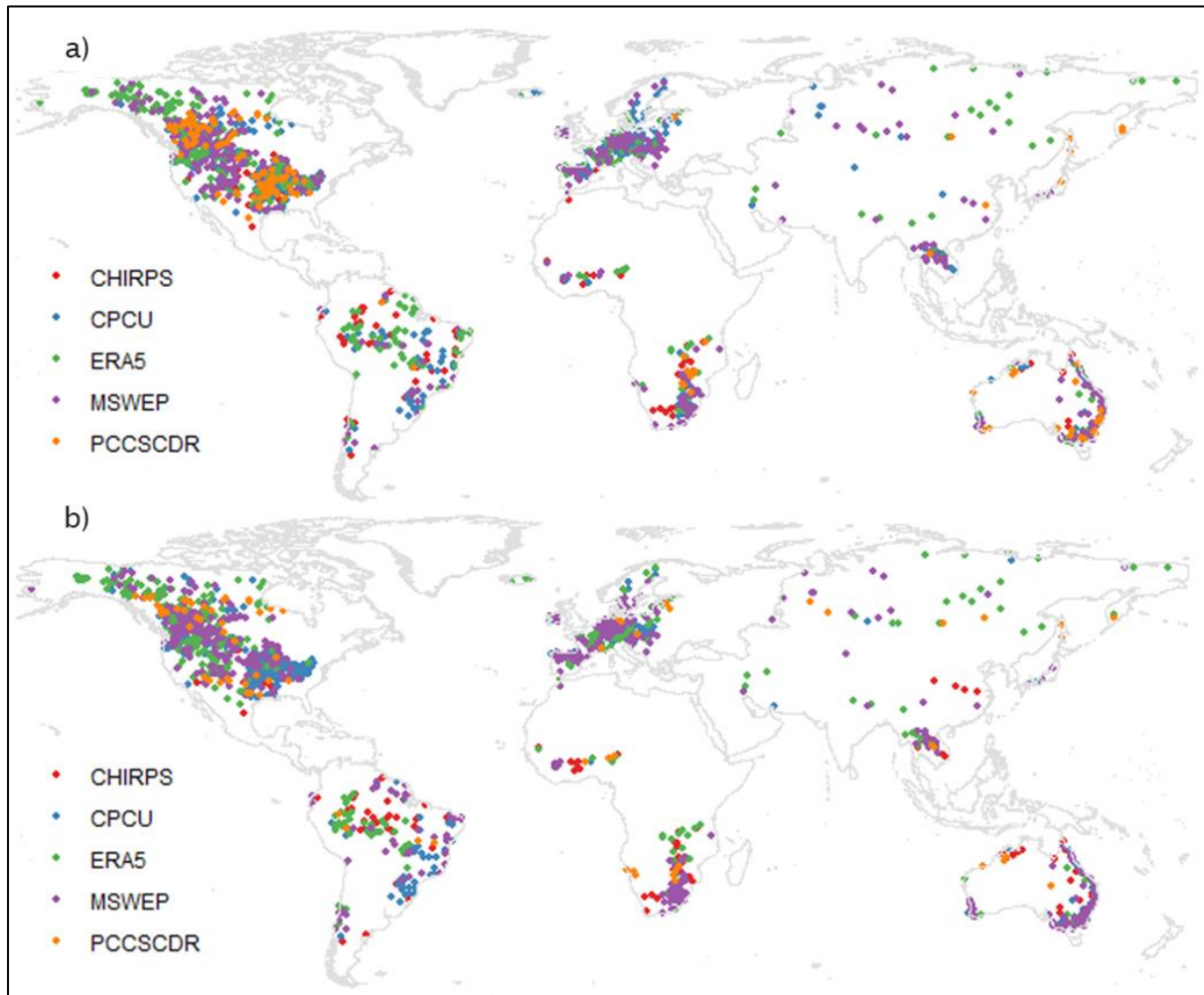
Figure 9 displays differences between the observed and modelled Q10 for selected stations of Lokoja, Khong-Chiam, Missao Icana, Savannah, Mirani Weir, and Dunaalmas (Table 2). Compared to ERA5, CPCU, and PERCCDR, MSWEP followed by CHIRPS showed a higher CC (0.21-0.62) and lower Pbias and RMSE in all stations. At stations Missao Icana, Savannah, and Mirani Weir, the observed Q10 is underestimated with a Pbias of between -20% to -80%. In all the stations, the positive and negative bias is large in PERCCDR and CPCU datasets.

4. Discussion and Conclusion

Given the challenges in representing precipitation at global scales, satellite, climate model, and reanalysis based precipitation datasets can form the basis for monitoring and prediction of water resources and hydrological extremes, particularly in data scarce regions of the world (Sheffield et al., 2018; Dembélé et al., 2020). Nevertheless, uncertainties and errors in these datasets require careful analysis to assess their suitability for a

452 ~~specific use. Error in satellite-based precipitation estimates can be due to errors in the sensor measurements, the~~
453 ~~frequency of sampling, and the retrieval algorithms, including the representation of cloud physics (Dembélé et al.,~~
454 ~~2020; Laiti et al., 2018; Alazzy et al., 2017). Climate model-based datasets, including reanalyses, have large~~
455 ~~uncertainty due to their coarse spatial resolution and ambiguity associated with model parameters (Gebrechorkos~~
456 ~~et al., 2018; AL-Falahi et al., 2020; Dembélé et al., 2020; Her et al., 2019). Reanalysis datasets may correct for~~
457 ~~some of these errors via the assimilation of observational data, but this comes with its own uncertainties due to~~
458 ~~the error characteristics of the assimilated observations and the assimilation scheme (Sheffield et al., 2006; Parker,~~
459 ~~2016). In hydrological modelling, errors and biases in precipitation data result in poor representation of the~~
460 ~~hydrological responses and affect applications (Maggioni and Massari, 2018; Zambrano-Bigiarini et al., 2016).~~
461 ~~For example, according to Bárdossy et al. (2022), uncertainty in precipitation can lead to hydrological model~~
462 ~~errors of up to 50%. Hence, it is important to assess the quality and accuracy of the precipitation products before~~
463 ~~using them in global or basin-scale hydrological models. In data-limited regions, hydrological models driven by~~
464 ~~precipitation datasets developed from satellite sources, reanalysis or climate models are the only plausible way to~~
465 ~~represent the terrestrial water cycle (van Huijgevoort et al., 2013).~~

466 ~~In light of the above, this study assesses the performance of selected global and quasi-global precipitation datasets~~
467 ~~for global hydrological modelling. It is important to note that this study assesses the precipitation datasets without~~
468 ~~calibration of the WBMsed model for each dataset, which could theoretically improve their performance in~~
469 ~~replicating observed river discharge. Within this context, our objective is not to evaluate the absolute performance~~
470 ~~of the hydrological model, which can be influenced by local factors, rather our focus is on comparing the relative~~
471 ~~performance of these datasets at individual locations across various precipitation datasets.~~ Based on the evaluation
472 at annual, monthly and daily time scales and analysis of daily extremes, no single precipitation dataset consistently
473 exhibits high accuracy across all geographical regions, nor does~~is~~ one consistently ~~better~~ outperform ~~than~~ the other
474 datasets. This finding is in ~~line~~ agreement with previous studies (Beck et al., 2017a; Dembélé et al., 2020). A
475 similar pattern of varied performance (e.g., lower in Africa and the central United States and better in Europe) by
476 different global hydrological models and precipitation datasets has been presented (Beck et al., 2017a; Lin et al.,
477 2019; Harrigan et al., 2020). In addition to the uncertainty in the precipitation datasets, the poorer performance in
478 some regions presented in this and previous studies (Beck et al., 2017a; Lin et al., 2019; Harrigan et al., 2020) can
479 be ~~due~~ exacerbated by~~to~~ the lack of representation in the hydrological models of anthropogenic influences, such
480 as for agriculture, irrigation, water supply, and energy production.

481 Comparably, MSWEP and ERA5 consistently exhibited higher CC and KGE values at over 50% of the stations
482 across annual, monthly, and daily time scales. According to Gu et al. (2023), satellite- and reanalysis-based
483 precipitation datasets, such as MSWEP and ERA5, can provide satisfactory performance for simulating discharge
484 globally. The higher performance of MSWEP indicates the advantage of incorporating a large number of daily
485 observations from field-based meteorological stations, in addition to a large set of satellite and reanalysis datasets
486 (Beck et al., 2017a, 2019a). Other studies have also shown the good performance of MSWEP for hydrological
487 modelling in different parts of the world (Beck et al., 2017a; Lakew, 2020; Li et al., 2022a; Reis et al., 2022; Gu
488 et al., 2023; López López et al., 2017; Satgé et al., 2019; Ibrahim et al., 2022). For example, Satgé et al. (2019)
489 evaluated 12 satellite-based precipitation estimates such as MSWEP, CHIRPS and PERSIANN-CDR in South
490 America (Lake Titicaca region) and found MSWEP was the best precipitation dataset for realistic simulation of

491 river discharge. MSWEP was also found to be the most reliable precipitation dataset compared to multiple datasets
492 such as CHIRPS and CMORPH for hydrological and climate studies in basins of Eastern China (Shaowei et al.,
493 2022; Wu et al., 2018). ~~Figure 10 displays the datasets with higher CC and KGE values for modelling daily~~
494 ~~discharge, offering a clear depiction of the spatial variability in precipitation dataset performance.~~

495

498 Even though ERA5 showed a higher KGE and CC than MSWEP, CHIRPS and TERRA in about 32% of the
499 stations it showed a higher error and biases. Previous studies have revealed bias and errors in ERA5 precipitation
500 (Lavers et al., 2021; Bechtold et al., 2020; AL-Falahi et al., 2020; Jiang et al., 2023; Lavers et al., 2022), which
501 leads to propagated errors and bias in hydrological modelling outputs. Harrigan et al. (2020) also reported large
502 biases in ERA5-driven hydrological simulations in the Central United States, South America (e.g., Brazil), and
503 Africa. According to Lavers et al. (2022), ERA5 precipitation is more reliable in extratropical areas compared to
504 tropical areas. Despite CPCU being a gauge-based precipitation dataset, it did not show as good performance as
505 MSWEP and ERA5 on annual, monthly, and daily timescales. In addition to the lower KGE and CC, CPCU
506 showed higher bias and error, particularly on annual and monthly time scales. The bias and errors in CPCU can
507 be due to the coarse resolution (0.5°) and the limited number of stations used to develop the datasets, particularly

in Africa and South America. According to Beck et al. (2017a), CPCU can be used in large river basins with dense meteorological stations but can be disadvantageous in Africa and South America. This highlights the need to expand and maintain the meteorological stations in these regions, but also the need to draw from satellite and model data sources. The PERSIANN-CDR is the least-performing product with lower KGE and higher errors and biases, which has been highlighted elsewhere in terms of its inability to represent precipitation extremes (Miao et al., 2015; Solakian et al., 2020).

The precipitation datasets show limited skill overall in reproducing daily extremes (high and low flows), relative to the annual and monthly time scales. MSWEP and CPCU have shown a high CC in about 38% of the stations. This is consistent with the findings of Tang et al., (2019) for the Mekong River Basin. CHIRPS and PERSIANN-CDR are the least skilful in capturing extremes with a very low CC and large positive and negative biases (Araujo Palharini et al., 2021). For instance, numerous precipitation products have been observed to both underestimate and overestimate low and high precipitation values in Brazil (Palharini et al., 2020), consequently resulting in corresponding underestimations and overestimations of low and high streamflows. In general, several studies have concluded that precipitation datasets exhibit a substantial disparity in daily extreme precipitation events (e.g., Araujo Palharini et al., 2021; Jiang et al., 2019; Huang et al., 2022), which can be attributed to factors such as inaccuracies in satellite sensors, retrieval algorithms, temporal sampling, and satellite-observation merging and bias correction procedures used, particularly in gauge-limited regions (Miao et al., 2015; El Kenawy et al., 2015; Shen et al., 2010; Jiang et al., 2019). In addition to the uncertainty of the precipitation datasets, the limited availability of hydrological observations limits the ability to assess these datasets globally, especially for extreme flood and drought events (Brunner et al., 2021).

~~While our study evaluates six global precipitation datasets for hydrological modelling using~~Whilst the WBMsed model has been shown to perform well in predicting mean flows when compared to historical observations, for example Cohen et al. (2022) ~~, which show~~ report an R^2 of 0.99 in 30-year average prediction against USGS gauge data and global river datasets ~~(Cohen et al., 2022)~~, it is important to acknowledge the uncertainties and limitations in the ~~both the precipitation data and~~ model parameterisation and initial conditions, including the role of anthropogenic impacts, which are greatly simplified in WBMsed. ~~Many of the s.~~ In addition, ~~u~~Uncertainties in the precipitation input data, such as those derived from satellite-based precipitation datasets, including retrieval errors, can propagate through the hydrological model, potentially affecting the accuracy of simulated discharge. ~~Additionally, g~~Globally calibrated model parameters may introduce further uncertainty, particularly in regions with limited observational data coverage. Due to the limited availability of observed discharge in Africa and Asia, evaluation predominantly focuses on North and South America and Europe. Hence, further evaluation in Africa and Asia could be essential to enhance the robustness of global hydrological models.

Overall, the evaluation presented in this paper underlines the importance of selecting high-quality precipitation datasets to drive hydrological models. Since no single precipitation dataset was found to be adequately accurate everywhere, across different temporal scales, this study can help identify the best precipitation products for any basin or region under consideration. Based on our results, MSWEP is the best overall choice but there are regions where ERA5, CHIRPS and CPCU were better overall ~~(e.g., see Figure 10)~~. All the precipitation datasets, particularly ERA5 and PERCCDR, require bias correction before being used to drive hydrological models in

546 regions like North America, Asia, Africa, and Australia. For data-scarce regions such as Africa and Asia, it is

547 difficult to recommend a precipitation dataset due to the limited number of hydrological stations used for

548 validation in this study. Finally, improving the precipitation datasets by adding more ground observations, for

549 example, and by better representing anthropogenic drivers in hydrological models has the potential of

550 considerably improving global and regional hydrological predictions.

551 **Data availability**

552 The selected precipitation datasets used in this study are openly accessible to the public. ERA5 is freely available

553 from the Copernicus Climate Data Store (CDS; https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-

554 era5-land?tab=overview). CHIRPS can be obtained from the Climate Hazards Group (CHG;

555 https://www.chc.ucsb.edu/data/chirps/). Access to the MSWEP precipitation dataset is provided through the

556 GloH2O website (https://www.gloh2o.org/mswep/). TERRA is accessible from the Climatology Lab website

557 (https://www.climatologylab.org/). CPCU is publicly available through the NOAA Physical Sciences Laboratory

558 (PSL; https://downloads.psl.noaa.gov/Datasets/cpc_global_precip/), and PERCCDR can be freely accessed

559 through the Center for Hydrometeorology and Remote Sensing (CHRS; https://chrsdata.eng.uci.edu/).

560 **Author contribution**

561 SG, JL, and SJD conceived the study, incorporating input from all co-authors. SG led the global hydrological

562 modelling, while JL, SJD, and LS assisted with data management and computational resources. SG was

563 responsible for evaluating various precipitation datasets for hydrological modelling and drafted the initial

564 manuscript. SC provided the hydrological model and input parameters. MW, GB, RB, PD, HG, EV, YL, RH, LH,

565 SM, and JN executed extensive data quality control and identified stations for evaluation. PA, HC, AN, AT, and

566 JS provided code, methods, and guidance. DP, SJD, and SED supervised the research and secured funding. All

567 authors contributed to investigating research findings and played integral roles in manuscript writing and editing.

568 **Competing interests**

569 We declare that Louise Slater is a topical editor of Hydrology and Earth System Sciences (HESS).

570 **Acknowledgements**

571 This work is part of the Evolution of Global Flood Hazard and Risk (EVOFLOOD) project [NE/S015817/1]

572 supported by the Natural Environment Research Council (NERC), the UK Foreign, Commonwealth and

573 Development Office (FCDO) for the benefit of developing countries (Programme Code 201880) and the UK's

574 Natural Environment Research Council (NERC; NE/S017380/1).

575

576

577

578

579   Reference

580   Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.: TerraClimate, a high-resolution
581   global dataset of monthly climate and climatic water balance from 1958–2015, Sci Data, 5, 170191,
582   https://doi.org/10.1038/sdata.2017.191, 2018.

583   Acharya, S. C., Nathan, R., Wang, Q. J., Su, C.-H., and Eizenberg, N.: An evaluation of daily
584   precipitation from a regional atmospheric reanalysis over Australia, Hydrology and Earth System
585   Sciences, 23, 3387–3403, https://doi.org/10.5194/hess-23-3387-2019, 2019.

586   Ahmed, K., Shahid, S., Wang, X., Nawaz, N., and Khan, N.: Evaluation of Gridded Precipitation
587   Datasets over Arid Regions of Pakistan, Water, 11, 210, https://doi.org/10.3390/w11020210, 2019.

588   Alazzy, A. A., Lü, H., Chen, R., Ali, A. B., Zhu, Y., and Su, J.: Evaluation of Satellite Precipitation
589   Products and Their Potential Influence on Hydrological Modeling over the Ganzi River Basin of the
590   Tibetan Plateau, Advances in Meteorology, 2017, e3695285, https://doi.org/10.1155/2017/3695285,
591   2017.

592   AL-Falahi, A. H., Saddique, N., Spank, U., Gebrechorkos, S. H., and Bernhofer, C.: Evaluation the
593   Performance of Several Gridded Precipitation Products over the Highland Region of Yemen for
594   Water Resources Management, Remote Sensing, 12, 2984, https://doi.org/10.3390/rs12182984,
595   2020.

596   Araujo Palharini, R. S., Vila, D. A., Rodrigues, D. T., Palharini, R. C., Mattos, E. V., and Pedra, G. U.:
597   Assessment of extreme rainfall estimates from satellite-based: Regional analysis, Remote Sensing
598   Applications: Society and Environment, 23, 100603, https://doi.org/10.1016/j.rsase.2021.100603,
599   2021.

600   Bárdossy, A., Kilsby, C., Birkinshaw, S., Wang, N., and Anwar, F.: Is Precipitation Responsible for the
601   Most Hydrological Model Uncertainty?, Frontiers in Water, 4, 2022.

602   Bechtold, P., R Forbes, I Sandu, S Lang, and M Ahlgrimm: A major moist physics upgrade for the IFS, ,
603   24–32, 2020.

604   Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., Dijk, A. I. J. M. van, Weedon, G. P., Brocca, L.,
605   Pappenberger, F., Huffman, G. J., and Wood, E. F.: Global-scale evaluation of 22 precipitation
606   datasets using gauge observations and hydrological modeling, Hydrology and Earth System Sciences,
607   21, 6201–6217, https://doi.org/10.5194/hess-21-6201-2017, 2017a.

608   Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo,
609   A.: MSWEP: 3-hourly 0.25°global gridded precipitation (1979-2015) by merging gauge, satellite, and
610   reanalysis data, Hydrology and Earth System Sciences, 21, 589–615, https://doi.org/10.5194/hess-
611   21-589-2017, 2017b.

612   Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., Huffman, G. J.,
613   Adler, R. F., and Wood, E. F.: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar
614   data for the CONUS, Hydrology and Earth System Sciences, 23, 207–224,
615   https://doi.org/10.5194/hess-23-207-2019, 2019a.

616   Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Dijk, A. I. J. M. van, McVicar, T. R., and
617   Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative

618   Assessment, Bulletin of the American Meteorological Society, 100, 473–500,
619   https://doi.org/10.1175/BAMS-D-17-0138.1, 2019b.

620   Brunner, M. I., Slater, L., Tallaksen, L. M., and Clark, M.: Challenges in modeling and predicting floods
621   and droughts: A review, WIREs Water, 8, e1520, https://doi.org/10.1002/wat2.1520, 2021.

622   Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Wayne Higgins, R., and Janowiak, J. E.:
623   Assessing objective techniques for gauge-based analyses of global daily precipitation, Journal of
624   Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009132, 2008.

625   Chen, Y., Hu, D., Duan, X., Zhang, Y., Liu, M., and Shasha, W.: Rainfall-runoff simulation and flood
626   dynamic monitoring based on CHIRPS and MODIS-ET, International Journal of Remote Sensing, 41,
627   4206–4225, https://doi.org/10.1080/01431161.2020.1714779, 2020.

628   Cohen, S., Kettner, A. J., Syvitski, J. P. M., and Fekete, B. M.: WBMsed, a distributed global-scale
629   riverine sediment flux model: Model description and validation, Computers & Geosciences, 53, 80–
630   93, https://doi.org/10.1016/j.cageo.2011.08.011, 2013.

631   Cohen, S., Kettner, A. J., and Syvitski, J. P. M.: Global suspended sediment and water discharge
632   dynamics between 1960 and 2010: Continental trends and intra-basin sensitivity, Global and
633   Planetary Change, 115, 44–58, https://doi.org/10.1016/j.gloplacha.2014.01.011, 2014.

634   Cohen, S., Syvitski, J., Ashley, T., Lammers, R., Fekete, B., and Li, H.-Y.: Spatial Trends and Drivers of
635   Bedload and Suspended Sediment Fluxes in Global Rivers, Water Resources Research, 58,
636   e2021WR031583, https://doi.org/10.1029/2021WR031583, 2022.

637   Day, C. A. and Howarth, D. A.: Modeling Climate Change Impacts on the Water Balance of a Medium-
638   Scale Mixed-Forest Watershed, SE USA, Southeastern Geographer, 59, 110–129, 2019.

639   Dembélé, M., Schaefli, B., van de Giesen, N., and Mariéthoz, G.: Suitability of 17 gridded rainfall and
640   temperature datasets for large-scale hydrological modelling in West Africa, Hydrology and Earth
641   System Sciences, 24, 5379–5406, https://doi.org/10.5194/hess-24-5379-2020, 2020.

642   Dunn, F. E., Darby, S. E., Nicholls, R. J., Cohen, S., Zarfl, C., and Fekete, B. M.: Projections of declining
643   fluvial sediment delivery to major deltas worldwide in response to climate change and
644   anthropogenic stress, Environ. Res. Lett., 14, 084034, https://doi.org/10.1088/1748-9326/ab304e,
645   2019.

646   Eini, M. R., Rahmati, A., and Piniewski, M.: Hydrological application and accuracy evaluation of
647   PERSIANN satellite-based precipitation estimates over a humid continental climate catchment,
648   Journal of Hydrology: Regional Studies, 41, 101109, https://doi.org/10.1016/j.ejrh.2022.101109,
649   2022.

650   El Kenawy, A. M., Lopez-Moreno, J. I., McCabe, M. F., and Vicente-Serrano, S. M.: Evaluation of the
651   TMPA-3B42 precipitation product using a high-density rain gauge network over complex terrain in
652   northeastern Iberia, Global and Planetary Change, 133, 188–200,
653   https://doi.org/10.1016/j.gloplacha.2015.08.013, 2015.

654   Fallah, A., Rakhshandehroo, G. R., Berg, P., O, S., and Orth, R.: Evaluation of precipitation datasets
655   against local observations in southwestern Iran, International Journal of Climatology, 40, 4102–4116,
656   https://doi.org/10.1002/joc.6445, 2020.

657 Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J.,
658 Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations—a
659 new environmental record for monitoring extremes, Scientific Data, 2, 150066,
660 https://doi.org/10.1038/sdata.2015.66, 2015.

661 Gebrechorkos, S. H., Hülsmann, S., and Bernhofer, C.: Evaluation of multiple climate data sources for
662 managing environmental resources in East Africa, Hydrology and Earth System Sciences, 22, 4547–
663 4564, https://doi.org/10.5194/hess-22-4547-2018, 2018.

664 Gebrechorkos, S. H., Bernhofer, C., and Hülsmann, S.: Impacts of projected change in climate on
665 water balance in basins of East Africa, Science of The Total Environment,
666 https://doi.org/10.1016/j.scitotenv.2019.05.053, 2019.

667 Gebrechorkos, S. H., Bernhofer, C., and Hülsmann, S.: Climate change impact assessment on the
668 hydrology of a large river basin in Ethiopia using a local-scale climate modelling approach, Science of
669 The Total Environment, 742, 140504, https://doi.org/10.1016/j.scitotenv.2020.140504, 2020.

670 Gebrechorkos, S. H., Leyland, J., Darby, S., and Parsons, D.: High-resolution daily global climate
671 dataset of BCCAQ statistically downscaled CMIP6 models for the EVOFLOOD project. NERC EDS
672 Centre for Environmental Data Analysis,
673 https://doi.org/10.5285/C107618F1DB34801BB88A1E927B82317, 2022a.

674 Gebrechorkos, S. H., Pan, M., Beck, H. E., and Sheffield, J.: Performance of State-of-the-Art C3S
675 European Seasonal Climate Forecast Models for Mean and Extreme Precipitation Over Africa, Water
676 Resources Research, 58, e2021WR031480, https://doi.org/10.1029/2021WR031480, 2022b.

677 Gebrechorkos, S. H., Pan, M., Lin, P., Anghileri, D., Forsythe, N., Pritchard, D. M. W., Fowler, H. J.,
678 Obuobie, E., Darko, D., and Sheffield, J.: Variability and changes in hydrological drought in the Volta
679 Basin, West Africa, Journal of Hydrology: Regional Studies, 42, 101143,
680 https://doi.org/10.1016/j.ejrh.2022.101143, 2022c.

681 Gebrechorkos, S. H., Peng, J., Dyer, E., Miralles, D. G., Vicente-Serrano, S. M., Funk, C., Beck, H. E.,
682 Asfaw, D. T., Singer, M. B., and Dadson, S. J.: Global High-Resolution Drought Indices for
683 1981–2022, Earth System Science Data Discussions, 1–28, https://doi.org/10.5194/essd-
684 2023-276, 2023.

685 Geleta, C. D. and Deressa, T. A.: Evaluation of Climate Hazards Group InfraRed Precipitation Station
686 (CHIRPS) satellite-based rainfall estimates over Finchaa and Neshe Watersheds, Ethiopia,
687 Engineering Reports, 3, e12338, https://doi.org/10.1002/eng2.12338, 2021.

688 GRDC: The Global Runoff Data Centre, 56068 Koblenz, Germany, 2023.

689 Grogan, D. S., Zuidema, S., Prusevich, A., Wollheim, W. M., Glidden, S., and Lammers, R. B.: Water
690 balance model (WBM) v.1.0.0: a scalable gridded global hydrologic model with water-tracking
691 functionality, Geoscientific Model Development, 15, 7287–7323, https://doi.org/10.5194/gmd-15-
692 7287-2022, 2022.

693 Gu, L., Yin, J., Wang, S., Chen, J., Qin, H., Yan, X., He, S., and Zhao, T.: How well do the multi-satellite
694 and atmospheric reanalysis products perform in hydrological modelling, Journal of Hydrology, 617,
695 128920, https://doi.org/10.1016/j.jhydrol.2022.128920, 2023.

696   Guo, B., Zhang, J., Xu, T., Croke, B., Jakeman, A., Song, Y., Yang, Q., Lei, X., and Liao, W.: Applicability
697   Assessment and Uncertainty Analysis of Multi-Precipitation Datasets for the Simulation of Hydrologic
698   Models, Water, 10, 1611, https://doi.org/10.3390/w10111611, 2018.

699   Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error
700   and NSE performance criteria: Implications for improving hydrological modelling, Journal of
701   Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

702   Hafizi, H. and Sorman, A. A.: Assessment of 13 Gridded Precipitation Datasets for Hydrological
703   Modeling in a Mountainous Basin, Atmosphere, 13, 143, https://doi.org/10.3390/atmos13010143,
704   2022.

705   Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H.,
706   and Pappenberger, F.: GloFAS-ERA5 operational global river discharge reanalysis 1979–present,
707   Earth System Science Data, 12, 2043–2060, https://doi.org/10.5194/essd-12-2043-2020, 2020.

708   He, Q., Shen, Z., Wan, M., and Li, L.: Precipitable Water Vapor Converted from GNSS-ZTD and ERA5
709   Datasets for the Monitoring of Tropical Cyclones, IEEE Access, 8, 87275–87290,
710   https://doi.org/10.1109/ACCESS.2020.2991094, 2020.

711   Her, Y., Yoo, S.-H., Cho, J., Hwang, S., Jeong, J., and Seong, C.: Uncertainty in hydrological analysis of
712   climate change: multi-parameter vs. multi-GCM ensemble predictions, Sci Rep, 9, 4974,
713   https://doi.org/10.1038/s41598-019-41334-7, 2019.

714   Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey,
715   C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P.,
716   Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R.,
717   Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E.,
718   Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I.,
719   Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the
720   Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

721   Hong, Y., Xuan Do, H., Kessler, J., Fry, L., Read, L., Rafieei Nasab, A., Gronewold, A. D., Mason, L., and
722   Anderson, E. J.: Evaluation of gridded precipitation datasets over international basins and large
723   lakes, Journal of Hydrology, 607, 127507, https://doi.org/10.1016/j.jhydrol.2022.127507, 2022.

724   Hou, D., Charles, M., Luo, Y., Toth, Z., Zhu, Y., Krzysztofowicz, R., Lin, Y., Xie, P., Seo, D.-J., Pena, M.,
725   and Cui, B.: Climatology-Calibrated Precipitation Analysis at Fine Scales: Statistical Adjustment of
726   Stage IV toward CPC Gauge-Based Analysis, Journal of Hydrometeorology, 15, 2542–2557,
727   https://doi.org/10.1175/JHM-D-11-0140.1, 2014.

728   Huang, Z., Zhang, Y., Xu, J., Fang, X., and Ma, Z.: Can satellite precipitation estimates capture the
729   magnitude of extreme rainfall Events?, Remote Sensing Letters, 13, 1048–1057,
730   https://doi.org/10.1080/2150704X.2022.2123258, 2022.

731   van Huijgevoort, M. H. J., Hazenberg, P., van Lanen, H. a. J., Teuling, A. J., Clark, D. B., Folwell, S.,
732   Gosling, S. N., Hanasaki, N., Heinke, J., Koirala, S., Stacke, T., Voss, F., Sheffield, J., and Uijlenhoet, R.:
733   Global Multimodel Analysis of Drought in Runoff for the Second Half of the Twentieth Century, J.
734   Hydrometeor., 14, 1535–1552, https://doi.org/10.1175/JHM-D-12-0186.1, 2013.

735   Ibrahim, A. H., Molla, D. D., and Lohani, T. K.: Performance evaluation of satellite-based rainfall
736   estimates for hydrological modeling over Bilate river basin, Ethiopia, World Journal of Engineering,
737   ahead-of-print, https://doi.org/10.1108/WJE-03-2022-0106, 2022.

738 Jiang, Q., Li, W., Wen, J., Fan, Z., Chen, Y., Scaioni, M., and Wang, J.: Evaluation of satellite-based
739 products for extreme rainfall estimations in the eastern coastal areas of China, Journal of Integrative
740 Environmental Sciences, 16, 191–207, https://doi.org/10.1080/1943815X.2019.1707233, 2019.

741 Jiang, S., Wei, L., Ren, L., Zhang, L., Wang, M., and Cui, H.: Evaluation of IMERG, TMPA, ERA5, and
742 CPC precipitation products over mainland China: Spatiotemporal patterns and extremes, Water
743 Science and Engineering, 16, 45–56, https://doi.org/10.1016/j.wse.2022.05.001, 2023.

744 Jiao, D., Xu, N., Yang, F., and Xu, K.: Evaluation of spatial-temporal variation performance of ERA5
745 precipitation data in China, Sci Rep, 11, 17956, https://doi.org/10.1038/s41598-021-97432-y, 2021.

746 Kidd, C. and Levizzani, V.: Status of satellite precipitation retrievals, Hydrology and Earth System
747 Sciences, 15, 1109–1116, https://doi.org/10.5194/hess-15-1109-2011, 2011.

748 Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D.
749 B.: So, How Much of the Earth's Surface Is Covered by Rain Gauges?, Bulletin of the American
750 Meteorological Society, 98, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1, 2017.

751 Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not?
752 Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrology and Earth System Sciences,
753 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

754 Laiti, L., Mallucci, S., Piccolroaz, S., Bellin, A., Zardi, D., Fiori, A., Nikulin, G., and Majone, B.: Testing
755 the Hydrological Coherence of High-Resolution Gridded Precipitation and Temperature Data Sets,
756 Water Resources Research, 54, 1999–2016, https://doi.org/10.1002/2017WR021633, 2018.

757 Lakew, H. B.: Investigating the effectiveness of bias correction and merging MSWEP with gauged
758 rainfall for the hydrological simulation of the upper Blue Nile basin, Journal of Hydrology: Regional
759 Studies, 32, 100741, https://doi.org/10.1016/j.ejrh.2020.100741, 2020.

760 Lavers, D. A., Harrigan, S., and Prudhomme, C.: Precipitation Biases in the ECMWF Integrated
761 Forecasting System, Journal of Hydrometeorology, 22, 1187–1198, https://doi.org/10.1175/JHM-D-
762 20-0308.1, 2021.

763 Lavers, D. A., Simmons, A., Vamborg, F., and Rodwell, M. J.: An evaluation of ERA5 precipitation for
764 climate monitoring, Quarterly Journal of the Royal Meteorological Society, 148, 3152–3165,
765 https://doi.org/10.1002/qj.4351, 2022.

766 Lehner, B., Verdin, K. L., and Jarvis, A.: New global hydrography derived from spaceborne elevation
767 data, Eos, Transactions, American Geophysical Union, 89, 2, https://doi.org/10.1029/2008EO100001,
768 2008.

769 Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M.,
770 Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wisser, D.: High-
771 resolution mapping of the world's reservoirs and dams for sustainable river-flow management,
772 Frontiers in Ecology and the Environment, 9, 494–502, https://doi.org/10.1890/100125, 2011.

773 Li, L., Wang, Y., Wang, L., Hu, Q., Zhu, Z., Li, L., and Li, C.: Spatio-temporal accuracy evaluation of
774 MSWEP daily precipitation over the Huaihe River Basin, China: A comparison study with
775 representative satellite- and reanalysis-based products, J. Geogr. Sci., 32, 2271–2290,
776 https://doi.org/10.1007/s11442-022-2047-9, 2022a.

777   Li, M., Lv, X., Zhu, L., Uchenna Ochege, F., and Guo, H.: Evaluation and Application of MSWEP in
778   Drought Monitoring in Central Asia, Atmosphere, 13, 1053,
779   https://doi.org/10.3390/atmos13071053, 2022b.

780   Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T.
781   M., Allen, G. H., Gleason, C. J., and Wood, E. F.: Global Reconstruction of Naturalized River Flows at
782   2.94 Million Reaches, Water Resources Research, 55, 6499–6516,
783   https://doi.org/10.1029/2019WR025287, 2019.

784   López López, P., Sutanudjaja, E. H., Schellekens, J., Sterk, G., and Bierkens, M. F. P.: Calibration of a
785   large-scale hydrological model using satellite-based soil moisture and evapotranspiration products,
786   Hydrology and Earth System Sciences, 21, 3125–3144, https://doi.org/10.5194/hess-21-3125-2017,
787   2017.

788   Luo, X., Wu, W., He, D., Li, Y., and Ji, X.: Hydrological Simulation Using TRMM and CHIRPS
789   Precipitation Estimates in the Lower Lancang-Mekong River Basin, Chin. Geogr. Sci., 29, 13–25,
790   https://doi.org/10.1007/s11769-019-1014-6, 2019.

791   Maggioni, V. and Massari, C.: On the performance of satellite precipitation products in riverine flood
792   modeling: A review, Journal of Hydrology, 558, 214–224,
793   https://doi.org/10.1016/j.jhydrol.2018.01.039, 2018.

794   Mazzoleni, M., Brandimarte, L., and Amaranto, A.: Evaluating precipitation datasets for large-scale
795   distributed hydrological modelling, Journal of Hydrology, 578, 124076,
796   https://doi.org/10.1016/j.jhydrol.2019.124076, 2019.

797   Mehran, A. and AghaKouchak, A.: Capabilities of satellite precipitation datasets to estimate heavy
798   precipitation rates at different temporal accumulations, Hydrological Processes, 28, 2262–2270,
799   https://doi.org/10.1002/hyp.9779, 2014.

800   Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An Overview of the Global
801   Historical Climatology Network-Daily Database, Journal of Atmospheric and Oceanic Technology, 29,
802   897–910, https://doi.org/10.1175/JTECH-D-11-00103.1, 2012.

803   Mianabadi, A., Salari, K., and Pourmohamad, Y.: Drought monitoring using the long-term CHIRPS
804   precipitation over Southeastern Iran, Appl Water Sci, 12, 183, https://doi.org/10.1007/s13201-022-
805   01705-4, 2022.

806   Miao, C., Ashouri, H., Hsu, K.-L., Sorooshian, S., and Duan, Q.: Evaluation of the PERSIANN-CDR Daily
807   Rainfall Estimates in Capturing the Behavior of Extreme Precipitation Events over China, Journal of
808   Hydrometeorology, 16, 1387–1396, https://doi.org/10.1175/JHM-D-14-0174.1, 2015.

809   Miao, Q., Pan, B., Wang, H., Hsu, K., and Sorooshian, S.: Improving Monsoon Precipitation Prediction
810   Using Combined Convolutional and Long Short Term Memory Neural Network, Water, 11, 977,
811   https://doi.org/10.3390/w11050977, 2019.

812   Michaelides, S., Levizzani, V., Anagnostou, E., Bauer, P., Kasparis, T., and Lane, J. E.: Precipitation:
813   Measurement, remote sensing, climatology and modeling, Atmospheric Research, 94, 512–533,
814   https://doi.org/10.1016/j.atmosres.2009.08.017, 2009.

815   Moazami, S., Golian, S., Kavianpour, M. R., and Hong, Y.: Comparison of PERSIANN and V7 TRMM
816   Multi-satellite Precipitation Analysis (TMPA) products with rain gauge data over Iran, International
817   Journal of Remote Sensing, 34, 8156–8171, https://doi.org/10.1080/01431161.2013.833360, 2013.

818 Moragoda, N. and Cohen, S.: Climate-induced trends in global riverine water discharge and
819 suspended sediment dynamics in the 21st century, Global and Planetary Change, 191, 103199,
820 https://doi.org/10.1016/j.gloplacha.2020.103199, 2020.

821 Nguyen, P., Thorstensen, A., Sorooshian, S., Hsu, K., Aghakouchak, A., Ashouri, H., Tran, H., and
822 Braithwaite, D.: Global Precipitation Trends across Spatial Scales Using Satellite Observations,
823 Bulletin of the American Meteorological Society, 99, 689–697, https://doi.org/10.1175/BAMS-D-17-
824 0065.1, 2018.

825 Opere, A. O., Waswa, R., and Mutua, F. M.: Assessing the Impacts of Climate Change on Surface
826 Water Resources Using WEAP Model in Narok County, Kenya, Frontiers in Water, 3, 2022.

827 Palharini, R. S. A., Vila, D. A., Rodrigues, D. T., Quispe, D. P., Palharini, R. C., de Siqueira, R. A., and de
828 Sousa Afonso, J. M.: Assessment of the Extreme Precipitation by Satellite Estimates over South
829 America, Remote Sensing, 12, 2085, https://doi.org/10.3390/rs12132085, 2020.

830 Parker, W. S.: Reanalyses and Observations: What's the Difference?, Bulletin of the American
831 Meteorological Society, 97, 1565–1572, https://doi.org/10.1175/BAMS-D-14-00226.1, 2016.

832 Peng, J., Dadson, S., Hirpa, F., Dyer, E., Lees, T., Miralles, D. G., Vicente-Serrano, S. M., and Funk, C.:
833 A pan-African high-resolution drought index dataset, Earth System Science Data, 12, 753–769,
834 https://doi.org/10.5194/essd-12-753-2020, 2020.

835 Raimonet, M., Oudin, L., Thieu, V., Silvestre, M., Vautard, R., Rabouille, C., and Moigne, P. L.:
836 Evaluation of Gridded Meteorological Datasets for Hydrological Modeling, Journal of
837 Hydrometeorology, 18, 3027–3041, https://doi.org/10.1175/JHM-D-17-0018.1, 2017.

838 Reichle, R. H., Koster, R. D., Lannoy, G. J. M. D., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Touré,
839 A.: Assessment and Enhancement of MERRA Land Surface Hydrology Estimates, Journal of Climate,
840 24, 6322–6338, https://doi.org/10.1175/JCLI-D-10-05033.1, 2011.

841 Reis, A. A. dos, Weerts, A., Ramos, M.-H., Wetterhall, F., and Fernandes, W. dos S.: Hydrological data
842 and modeling to combine and validate precipitation datasets relevant to hydrological applications,
843 Journal of Hydrology: Regional Studies, 44, 101200, https://doi.org/10.1016/j.ejrh.2022.101200,
844 2022.

845 Sadeghi, M., Nguyen, P., Naeini, M. R., Hsu, K., Braithwaite, D., and Sorooshian, S.: PERSIANN-CCS-
846 CDR, a 3-hourly 0.04° global precipitation climate data record for heavy precipitation studies, Sci
847 Data, 8, 157, https://doi.org/10.1038/s41597-021-00940-9, 2021.

848 Salehi, H., Sadeghi, M., Golian, S., Nguyen, P., Murphy, C., and Sorooshian, S.: The Application of
849 PERSIANN Family Datasets for Hydrological Modeling, Remote Sensing, 14, 3675,
850 https://doi.org/10.3390/rs14153675, 2022.

851 Satgé, F., Ruelland, D., Bonnet, M.-P., Molina, J., and Pillco, R.: Consistency of satellite-based
852 precipitation products in space and over time compared with gauge observations and snow-
853 hydrological modelling in the Lake Titicaca region, Hydrology and Earth System Sciences, 23, 595–
854 619, https://doi.org/10.5194/hess-23-595-2019, 2019.

855 Seyyedi, H., Anagnostou, E. N., Beighley, E., and McCollum, J.: Hydrologic evaluation of satellite and
856 reanalysis precipitation datasets over a mid-latitude basin, Atmospheric Research, 164–165, 37–48,
857 https://doi.org/10.1016/j.atmosres.2015.03.019, 2015.

Shaowei, N., Jie, W., Juliang, J., Xiaoyan, X., Yuliang, Z., Fan, S., and Linlin, Z.: Comprehensive evaluation of satellite-derived precipitation products considering spatial distribution difference of daily precipitation over eastern China, Journal of Hydrology: Regional Studies, 44, 101242, https://doi.org/10.1016/j.ejrh.2022.101242, 2022.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling, J. Climate, 19, 3088–3111, https://doi.org/10.1175/JCLI3790.1, 2006.

Sheffield, J., Wood, E. F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., and Verbist, K.: Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions, Water Resources Research, 54, 9724–9758, https://doi.org/10.1029/2017WR022437, 2018.

Shen, Y., Xiong, A., Wang, Y., and Xie, P.: Performance of high-resolution satellite precipitation products over China, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2009JD012097, 2010.

Solakian, J., Maggioni, V., and Godrej, A. N.: On the Performance of Satellite-Based Precipitation Products in Simulating Streamflow and Water Quality During Hydrometeorological Extremes, Frontiers in Environmental Science, 8, 2020.

Sun, G., Wei, Y., Wang, G., Shi, R., Chen, H., and Mo, C.: Downscaling Correction and Hydrological Applicability of the Three Latest High-Resolution Satellite Precipitation Products (GPM, GSMAP, and MSWEP) in the Pingtang Catchment, China, Advances in Meteorology, 2022, e6507109, https://doi.org/10.1155/2022/6507109, 2022.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.-L.: A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, Reviews of Geophysics, 56, 79–107, https://doi.org/10.1002/2017RG000574, 2018.

Tang, X., Zhang, J., Gao, C., Ruben, G. B., and Wang, G.: Assessing the Uncertainties of Four Precipitation Products for Swat Modeling in Mekong River Basin, Remote Sensing, 11, 304, https://doi.org/10.3390/rs11030304, 2019.

Ursulak, J. and Coulibaly, P.: Integration of hydrological models with entropy and multi-objective optimization based methods for designing specific needs streamflow monitoring networks, Journal of Hydrology, 593, 125876, https://doi.org/10.1016/j.jhydrol.2020.125876, 2021.

Voisin, N., Wood, A. W., and Lettenmaier, D. P.: Evaluation of Precipitation Products for Global Hydrological Prediction, Journal of Hydrometeorology, 9, 388–407, https://doi.org/10.1175/2007JHM938.1, 2008.

Wang, M., Rezaie-Balf, M., Naganna, S. R., and Yaseen, Z. M.: Sourcing CHIRPS precipitation data for streamflow forecasting using intrinsic time-scale decomposition based machine learning models, Hydrological Sciences Journal, 66, 1437–1456, https://doi.org/10.1080/02626667.2021.1928138, 2021.

Wang, N., Liu, W., Sun, F., Yao, Z., Wang, H., and Liu, W.: Evaluating satellite-based and reanalysis precipitation datasets with gauge-observed data and hydrological modeling in the Xihe River Basin, China, Atmospheric Research, 234, 104746, https://doi.org/10.1016/j.atmosres.2019.104746, 2020.

898   Wati, T., Hadi, T. W., Sopaheluwakan, A., and Hutasoit, L. M.: Statistics of the Performance of
899   Gridded Precipitation Datasets in Indonesia, Advances in Meteorology, 2022, e7995761,
900   https://doi.org/10.1155/2022/7995761, 2022.

901   Wisser, D., Fekete, B. M., Vörösmarty, C. J., and Schumann, A. H.: Reconstructing 20th century global
902   hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H), Hydrology and
903   Earth System Sciences, 14, 1–24, https://doi.org/10.5194/hess-14-1-2010, 2010.

904   Wollheim, W. M., Vörösmarty, C. J., Bouwman, A. F., Green, P., Harrison, J., Linder, E., Peterson, B. J.,
905   Seitzinger, S. P., and Syvitski, J. P. M.: Global N removal by freshwater aquatic systems using a
906   spatially distributed, within-basin approach, Global Biogeochemical Cycles, 22,
907   https://doi.org/10.1029/2007GB002963, 2008.

908   Wu, Z., Xu, Z., Wang, F., He, H., Zhou, J., Wu, X., and Liu, Z.: Hydrologic Evaluation of Multi-Source
909   Satellite Precipitation Products for the Upper Huaihe River Basin, China, Remote Sensing, 10, 840,
910   https://doi.org/10.3390/rs10060840, 2018.

911   Xiang, Y., Chen, J., Li, L., Peng, T., and Yin, Z.: Evaluation of Eight Global Precipitation Datasets in
912   Hydrological Modeling, Remote Sensing, 13, 2831, https://doi.org/10.3390/rs13142831, 2021.

913   Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., and Ribbe, L.: Temporal and spatial
914   evaluation of satellite-based rainfall estimates across the complex topographical and climatic
915   gradients of Chile, Hydrology and Earth System Sciences Discussions, 1–43,
916   https://doi.org/10.5194/hess-2016-453, 2016.

917   Zhu, D., Ilyas, A. M., Wang, G., and Zeng, B.: Long-term hydrological assessment of remote sensing
918   precipitation from multiple sources over the lower Yangtze River basin, China, Meteorological
919   Applications, 28, e1991, https://doi.org/10.1002/met.1991, 2021.

920   Zhu, H., Li, Y., Huang, Y., Li, Y., Hou, C., and Shi, X.: Evaluation and hydrological application of
921   satellite-based precipitation datasets in driving hydrological models over the Huifa river basin in
922   Northeast China, Atmospheric Research, 207, 28–41,
923   https://doi.org/10.1016/j.atmosres.2018.02.022, 2018.

924