Received: 13 June 2019 | Revised: 3 December 2019 | Accepted: 4 December 2019 | Published on: 5 January 2020

DOI: 10.1002/joc.6445

International Journal of Climatology RMetS

RESEARCH ARTICLE

Evaluation of precipitation datasets against local observations in southwestern Iran

Ali Fallah[1,2,3] | **Gholam Reza Rakhshandehroo**[1] | **Peter Berg**[2] | **Sungmin O**[3] | **Rene Orth**[3]

[1]Department of Civil and Environmental Engineering, Shiraz University, Shiraz, Iran

[2]Hydrology Research Unit, SMHI, Norrköping, Sweden

[3]Max Planck Institute for Biogeochemistry, Jena, Germany

Correspondence
Ali Fallah, Department of Civil and Environmental Engineering, Faculty of Engineering, Shiraz University, Shiraz, Iran.
Email: alifallah@shirazu.ac.ir, afallah@bgc-jena.mpg.de

Funding information
German Research Foundation, Grant/Award Number: 391059971; Ministry of Science, Research and technology I.R. of Iran; Sveriges Meteorologiska och Hydrologiska Institut

Abstract

This study provides a comprehensive evaluation of a great variety of state-of-the-art precipitation datasets against gauge observations over the Karun basin in southwestern Iran. In particular, we consider (a) gauge-interpolated datasets (GPCCv8, CRU TS4.01, PREC/L, and CPC-Unified), (b) multi-source products (PERSIANN-CDR, CHIRPS2.0, MSWEP V2, HydroGFD2.0, and SM2RAIN-CCI), and (c) reanalyses (ERA-Interim, ERA5, CFSR, and JRA-55). The spatiotemporal performance of each product is evaluated against monthly precipitation observations from 155 gauges distributed across the basin during the period 2000–2015. This way, we find that overall the GPCCv8 dataset agrees best with the measurements. Most datasets show significant underestimations, which are largest for the interpolated datasets. These underestimations are usually smallest at low altitudes and increase towards more mountainous areas, although there is large spread across the products. Interestingly, no overall performance difference can be found between precipitation datasets for which gauge observations from Karun basin were used, versus products that were derived without these measurements, except in the case of GPCCv8. In general, our findings highlight remarkable differences between state-of-the-art precipitation products over regions with comparatively sparse gauge density, such as Iran. Revealing the best-performing datasets and their remaining weaknesses, we provide guidance for monitoring and modelling applications which rely on high-quality precipitation input.

KEYWORDS

evaluation, interpolated dataset, Karun basin, precipitation datasets, reanalysis dataset, satellite rainfall estimate

1 | INTRODUCTION

In many regions of the world, changes in water availability have severe impacts on society and economy. These changes will be intensified by climate change in some regions (Kirtman *et al.*, 2013). This jeopardizes water and food security, especially in developing countries with highly agricultural-oriented economies (Vaghefi *et al.*, 2019; Hameed et al., 2020). In this context, a lack of reliable precipitation information, which is key to monitor these water dynamics, has been a serious barrier for supporting decision-makers. To compensate for this

deficit, different sources of data, including gauge or satellite-based and reanalyses, have been utilized by researchers to monitor and predict extreme events across the world (Naumann *et al*., 2014; AghaKouchak *et al*., 2015; Zhan *et al*., 2016; Balsamo *et al*., 2018). However, without comprehensive and comparative validation against ground-based observations the usefulness of these data products is unclear.

In spite of this necessity for accurate ground-based precipitation dataset, reliable gauge station datasets are not widely available. Further, existing gauge observations often suffer from artificial disturbances because the stations' reliability varies with time due to, for example, instrument deterioration or relocations. As a result, there may be substantial questionable or missing data records. To address these problems, some studies have focused on reconstruction, quality control and homogeneity of the time series (e.g., González-Rouco *et al*., 2001 in south-western Europe; Beaulieu *et al*., 2008 in Québec, Canada; Vicente-Serrano *et al*., 2010 in north-eastern Spain).

As a result of sparse gauge measurements, datasets of near-global coverage have been generated with various approaches. Some make direct use of gauge measurements together with statistical techniques for interpolating the observations (e.g., Chen *et al*., 2002; Xie *et al*., 2007; Chen *et al*., 2008; Becker *et al*., 2013; Harris *et al*., 2014). Others use remote sensing from satellites with high spatiotemporal resolution and near real-time availability, making them suitable especially for data-sparse or ungauged basins. Because of the indirect nature of the precipitation estimates from satellites, the products are subject to a variety of potential errors (Brocca *et al*., 2014; Moazami *et al*., 2014; Koster *et al*., 2016; Sun *et al*., 2018; Salmani-Dehaghi and Samani, 2019). The satellite estimates are therefore often blended with gauge data (Ashouri *et al*., 2015; Funk *et al*., 2015), which also enhances their usefulness in areas with insufficient gauge coverage.

Besides these direct and indirect measurements of precipitation, there are also modelling approaches such as through reanalyses which assimilate meteorological observations from various sources, for example, ground-based stations, ships, airplanes, and satellites (Parker, 2016) with forecasts from numerical weather prediction models to infer precipitation estimates (e.g., Saha *et al*., 2010, 2014; Dee *et al*., 2011; Kobayashi *et al*., 2015; Copernicus Climate Change Service, 2017). To alleviate the major bias in reanalysis models, because of not making any direct use of gauge information, they are sometimes combined with bias adjustment techniques (Weedon *et al*., 2011; Berg *et al*., 2018) and often merged with other data products such as satellite remote sensing and interpolated gauge datasets (Beck *et al*., 2017a, 2017b). In regions with poor station coverage, satellite rainfall estimates and reanalysis products may compensate for the lack of gauge stations provided a meaningful calibration of the underlying models. This is particularly challenging in mountainous regions (Dinku *et al*., 2008; Hu *et al*., 2016; Alijanian *et al*., 2017; Beck *et al*., 2017a).

This growing number of precipitation datasets derived through various approaches implies a need for a comparative performance assessment. In this study, we compare the performance of various precipitation datasets derived from multiple sources, against gauge measurements over Karun basin, southwestern Iran. The outcome can inform dataset developers about respective strengths and weaknesses, and also provide guidance to users for their choice from the variety of state-of-the-art products. This is especially important in relatively data-sparse regions such as Iran. Although there have been some studies on the evaluation of precipitation products over Iran's climatic zones (Moazami *et al*., 2014; Katiraie-Boroujerdy *et al*., 2013, 2017, 2019; Ghajarnia *et al*., 2015; Sharifi *et al*., 2016; Khodadoust Siuki *et al*., 2017; Alijanian *et al*., 2017, 2019; Hosseini-Moghari *et al*., 2018; Dezfooli *et al*., 2018; Saeidizand *et al*., 2018), comprehensive evaluations of reanalysis, satellite-based, and interpolated precipitation data are lacking.

Sections 2 and 3 introduce the study area and the considered precipitation datasets. Section 4 illustrates the methodology, and Section 5 presents results and discussion. Finally, in Section 6 the conclusions of this study are presented.

2 | STUDY AREA

Karun basin, with an area of 65,230 km^2, is one of the largest basins in Iran (Figure 1), hosting the Karun River with an average annual discharge rate of 575 $m^3 \cdot s^{-1}$. Variation in topography is significant over the basin; surface elevation varies from zero at Persian Gulf coast to 4,400 m over the Zagros mountain chains. The basin encompasses various climate zones; this climate variability is controlled by geographical latitude, proximity to the Persian Gulf, and elevation. Average annual precipitation over the basin is about 632 mm, however, with a large spatial variability illustrated by values ranging from 153 mm in southern plain regions to >2000 mm in mountainous regions. Daily temperature varies over the basin from a minimum of −30.6°C at Koohrang station to a maximum of 52.2°C at Ahvaz station.

3 | DATA

In this section, we introduce the ground-truth reference data (Section 3.1) which we use to validate a multitude of established gridded state-of-the-art precipitation datasets (Section 3.2).

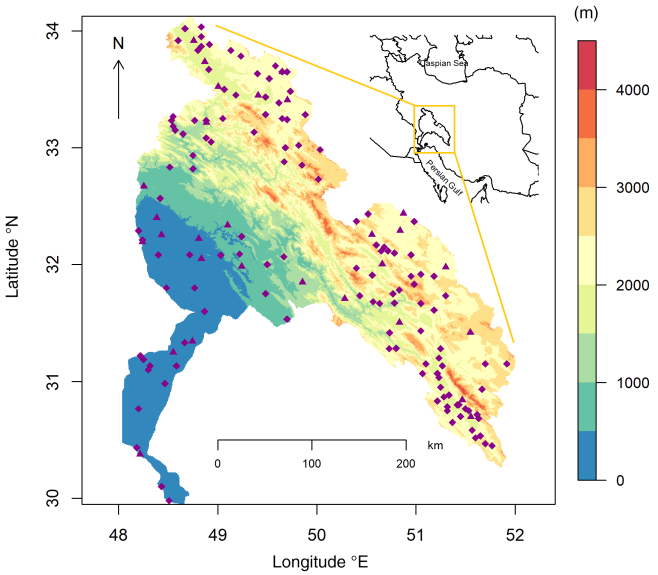

FIGURE 1 Topographical map of the Karun basin in Iran. The location of precipitation gauge stations utilized in this study is marked with purple diamonds (daily rain gauges) and triangles (synoptic stations)

3.1 | Reference data

In this study, in situ data from rain gauges operated by IRIMO (Islamic Republic of Iran Meteorological Organization) are utilized as a ground reference for evaluating the selected gridded precipitation products (Figure 1). In essence, this dataset consists of two different sources of 3-hr synoptic and daily rain gauges, represented by triangle and diamond, respectively. It should be noted that the synoptic stations are more reliable because of occurring less human error in the process of observation. As shown in Figure S1 the distribution of gauges across altitudes matches that of the grid cells which cover Karun basin. No statistical post-processing has been applied to the gauge measurements. We focus on the time period 2000–2015 in this study since, according to plotted stations' time series, the overall availability of the station measurements is highest in these years. As most gridded datasets come with a monthly resolution, we derive monthly estimates of the gauge data by accumulating daily values.

3.2 | Gridded data products for evaluation

The gridded products are grouped into three categories: entirely gauge-based datasets, merged datasets using gauge data among other sources, and reanalysis datasets not making any direct use of gauge information. A summary of all individual datasets and their respective characteristics is shown in Table 1.

3.2.1 | Interpolated gauge data

The four interpolated gauge datasets evaluated in this study are GPCCv8 (Global Precipitation Climatology Centre; Schneider *et al.*, 2014, 2018), CPC-Unified (Climate Prediction Center Unified; Xie *et al.*, 2007; Chen *et al.*, 2008), PREC/L (PRECipitation REConstruction over Land; Chen *et al.*, 2002), and CRU TS4.01 (Climatic Research Unit; Harris *et al.*, 2014). They share a spatial resolution of 0.5°, except for GPCCv8, which provides 0.25°. These products provide gridded gauge analysis products derived from quality-controlled station data at a daily (CPC-Unified) or monthly (GPCCv8, PREC/L and CRU TS4.01) temporal resolution. PREC/L is based on an advanced method of optimal interpolation (OI) and is derived from gauge observations from over 17,000 stations collected in the Global Historical Climatology Network version 2, and the Climate Anomaly Monitoring System datasets (Chen *et al.*, 2002). CPC-Unified is derived by combining all information sources available at CPC, 16,000 quality-controlled daily stations, and by taking advantage of the OI objective analysis technique (Chen *et al.*, 2008). GPCC employs an extraordinarily large number of gauges around 85,000 stations (Schneider *et al.*, 2014, 2018). Further, it provides the number of gauges used for each grid cell, and an uncertainty estimate deduced from ordinary kriging (Yamamoto, 2000). CRU TS4.01 provides a precipitation dataset and other metrological variables from 1901 to near-present, including over 4,000 individual weather station records (Harris *et al.*, 2014).

3.2.2 | Merged multi-source data

We use three high-resolution merged products, namely CHIRPS2.0 (Climate Hazards group Infrared Precipitation with Stations), PERSIANN-CDR (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record), and SM2RAIN-CCI (SM2RAIN-Climate Change Initiative, published in July 2015). CHIRPS is a quasi-global rainfall dataset, spanning across 50°S–50°N. It extends from 1981 to near-present and blends 0.05° resolution satellite imagery with in situ station data to create gridded rainfall time series. PERSIANN-CDR provides daily rainfall estimates at a spatial resolution of 0.25° over the latitude band of

TABLE 1 Summary of the precipitation datasets evaluated in this study (NRT stands for near real-time)

Group	Dataset	Temporal coverage	Spatial coverage	Spatial resolution	Data sources	Reference
Modelled	JRA-55	1959–NRT[a]	Global	~0.56°	Reanalysis	Kobayashi *et al.* (2015); Japanese 55-year Reanalysis (2013)
	CFSR	1979–NRT[b]	Global	0.5°	Reanalysis	Saha *et al.* (2010, 2012, 2014)
	ERA-Interim	1979–NRT	Global	~0.75°	Reanalysis	Dee *et al.* (2011)
	ERA5	1979–NRT[c]	Global	~0.28°	Reanalysis	Copernicus Climate Change Service (2017)
Interpolated	CRU TS4.01	1901–2016	Global	0.5°	Gauge	Harris *et al.* (2014)
	CPC-Unified	1979–NRT[a]	Global	0.5°	Gauge	Xie *et al.* (2007); Chen *et al.* (2008)
	PREC/L	1948–NRT[c]	Global	0.5°	Gauge	Chen *et al.* (2002)
	GPCCv8	1901–2016	Global	0.25°	Gauge	Becker *et al.* (2013); Schneider *et al.* (2014, 2018)
Multi-source	MSWEP V2	1979–NRT[b]	Global	0.1°	Satellite + gauge + reanalysis	Beck *et al.* (2019)
	HydroGFD2.0	1978–NRT[c]	Global	0.5°	Satellite + gauge + reanalysis	Berg *et al.* (2018)
	CHIRPS2.0	1981–NRT[a]	50°N–50°S	0.05°	Satellite + gauge+ reanalysis	Funk *et al.* (2015)
	PERSIANN-CDR	1983–NRT[c]	60°N–60°S	0.25°	Satellite + gauge[d]	Ashouri *et al.* (2015)
	SM2RAIN-CCI	1998–2015	Global	0.25°	Inversion of the satellite soil moisture + gauge[e]	Brocca *et al.* (2014); Ciabatta *et al.* (2018)

[a]Available until the present with a delay of several days.
[b]Available until the present with a delay of several hours.
[c]Available until the present with a delay of several months.
[d]Gauge information have been utilized for bias adjustment.
[e]Gauge information have been incorporated in the calibration process.

60°S–60°N from 1983 to the near-present. PERSIANN-CDR is produced from the PERSIANN algorithm using GridSat-B1 infrared satellite data, and the training of the artificial neural network is done using NCEP stage IV radar data. The biases of PERSIANN-CDR are adjusted using 2.5° monthly GPCP (Global Precipitation Climatology Project) precipitation data (Adler *et al.*, 2003; Ashouri *et al.*, 2015). Brocca *et al.* (2014) developed a global scale rainfall product, SM2RAIN, by translating soil moisture obtained from satellite soil moisture data into precipitation estimates. Recently, the SM2RAIN method has been applied to the satellite-derived ESA CCI soil moisture product with a spatial resolution of 0.25° for the period of 1998–2015 (Ciabatta *et al.*, 2018). The algorithm was calibrated against the Global Precipitation Climatology Centre Full-Data daily dataset (Ciabatta *et al.*, 2018). Note that even though these datasets rely primarily on satellite-derived information, they also indirectly employ gauge measurements for bias adjustment (PERSIANN-CDR) or algorithm calibration (SM2RAIN-CCI).

Further, we evaluate two datasets: HydroGFD2.0 (Hydrological Global Forcing Data) and MSWEP V2 (Multi-Source Weighted-Ensemble Precipitation), which are produced by merging gauge, satellite and reanalysis data. HydroGFD2.0 is methodologically based on Weedon *et al.* (2011), and is produced in near real-time (Berg *et al.*, 2018; Arheimer et al., 2019). The HydroGFD2.0 dataset covers the period 1979 to present at a daily time scale and with a spatial resolution of 0.5°. Additionally, we use the MSWEP V2 dataset from Beck *et al.*, 2019. It provides data with high spatial (0.1°) and temporal (3-hourly) resolution, and is computed by merging precipitation estimates based on gauges, satellites, and reanalyses data. In addition, they corrected frequency and systematic biases in the precipitation data (Beck *et al.*, 2019).

3.2.3 | Reanalysis data

Climate reanalyses are derived through observation-adjusted model simulations which generate spatiotemporal consistent time series of multiple climate variables. In this study, four reanalysis precipitation datasets are evaluated: CFSR (Climate Forecast System Reanalysis), JRA-55 (Japanese 55-year Reanalysis), ERA-Interim (European Centre for Medium-range Weather Forecasts Reanalysis Interim) and ERA5. They all cover the entire global domain. CFSR extends from 1979 to 2010 with a spatial resolution of 0.5° (Saha *et al*., 2010, 2012). JRA-55 covers a period of more than 55 years from 1959 with a spatial resolution of 0.5625°, and with 3-hourly time steps. It is based on four-dimensional variational data assimilation (4D-Var) with Variational Bias Correction (VarBC) for satellite radiances (Kobayashi *et al*., 2015). ERA-Interim starts in 1979 and is continuously updated, providing 3-hourly data with a spatial resolution of ~0.75°. The data assimilation system used to produce ERA-Interim is based on a 2006 release of the integrated forecasting system (IFS), including 4D-Var with a 12-hr analysis window (Dee *et al*., 2011). ERA5 is the sequel of ERA-Interim. Using the latest IFS version, it provides hourly data on many atmospheric, land-surface and sea-state parameters together with estimates of uncertainty on a 30-km grid from 1979 to near-present (Copernicus Climate Change Service, 2017).

4 | METHODOLOGY

Before comparing the various precipitation products, they are re-gridded to a common 0.5° spatial resolution, if necessary. This was done using climate data operators (Schulzweida, 2019), namely through conservative remapping which preserves the water mass (Jones, 1999). This method is widely used for remapping precipitation datasets (Chen and Knutson, 2008; Nikulin *et al*., 2012). To examine the effect of this re-gridding, we also compute our analyses with the native resolutions of the precipitation datasets.

Gauge measurements within each grid box of the respective datasets were averaged arithmetically to produce representative time series, and to form a reference for the evaluation of the considered precipitation datasets. In this context, to assess the performance, monthly time series were computed for for gauges and datasets; G and M are defined later. Based on these time series, six measures were calculated for each grid cell of the models: mean and maximum values, spatial and temporal correlation coefficient (CC), relative error (RE), and absolute error (AE):

$$\mathrm{CC}=\frac{\sum_{i=1}^{n}(\mathrm{P}_{M_i}-\bar{p_M})(P_{G_i}-\bar{p_G})}{\sqrt{\sum_{i=1}^{n}(P_{M_i}-\bar{p_M})^2}\sqrt{\sum_{i=1}^{n}(P_{G_i}-\bar{p_G})^2}}, \quad (1)$$

$$\mathrm{RE}=\frac{\sum_{i=1}^{n}(P_{M_i}-P_{G_i})}{\sum_{i=1}^{n}P_{G_i}}\times 100, \quad (2)$$

$$\mathrm{AE}=\frac{\sum_{i=1}^{n}(P_{M_i}-P_{G_i})}{n}, \quad (3)$$

where i, n, and top bar indicate time, the number of months, and an average over time, respectively. Further, P denotes monthly precipitation data from gauge measurements (G) or the considered datasets (M). After calculating these measures, the average of all grid cell values is taken to obtain a single representative value for each metric and each dataset over the basin. In addition to temporal correlation, we also infer spatial correlations of observed versus modelled grid cell averages.

We also examine the accuracy of products in different climate regimes. These are characterized by long-term average aridity and temperature, and computed from ERA-Interim data for each grid cell. Thereby, aridity is computed as the ratio of mean annual net radiation to mean annual precipitation, converted to the same units by normalization with the latent heat of vaporization (Budyko, 1974; Orth and Destouni, 2018). Figure S2 indicates the basin climate determined by aridity values in each grid cell along with the respective stations' temperature.

5 | RESULTS AND DISCUSSION

5.1 | Seasonal variability of precipitation

Figure 2 illustrates mean monthly precipitation averages across the basin. It shows that the main differences between the datasets occur during the regional wet season (November to April) while all datasets capture the summer dry season. Most datasets, except for the reanalyses and HydroGFD2.0, underestimate precipitation during the wet season. GPCCv8, reanalyses, and merged products generally demonstrate seasonal variability of precipitation better than the other interpolated gauge products. Underestimation in the datasets incorporating satellite estimates during winter might be due to a systematic bias related to snow-covered surfaces (Gebregiorgis *et al*., 2017); this is despite the fact that they are calibrated with gauge observations. Moreover, such underestimation in SM2RAIN-CCI might be further due to surface soil moisture saturation (Brocca *et al*., 2013). We repeat this analysis for low, medium, and high

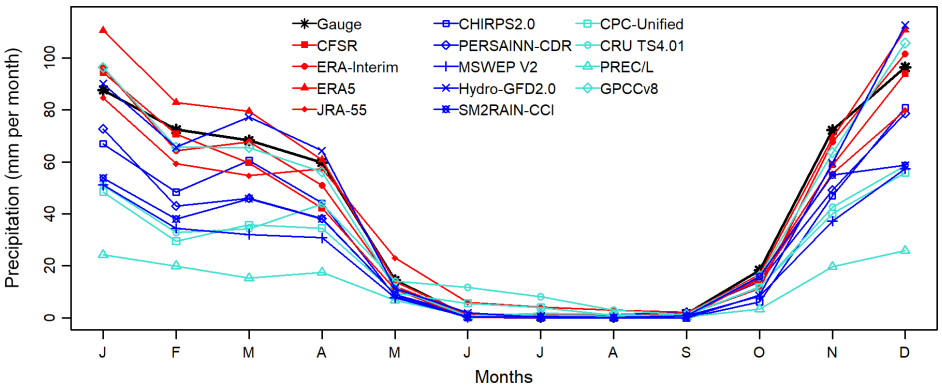

FIGURE 2 Comparison of mean monthly precipitation (mm per month) for different datasets (symbols) across all grids of Karun basin (2000–2015). Red lines denote datasets which did not employ gauge data, while blue lines indicate datasets which use some gauge records, and turquoise lines refer to datasets which extensively used gauge observations

TABLE 2 Statistical measures of the evaluated monthly datasets over Karun basin (2000–2015)

Group	Dataset	Mean (mm per month)	Maximum (mm per month)	Temporal correlation (−)	Spatial correlation (−)	RE (%)	AE (mm per month)
	Gauge	38.6	291.9				
Reanalysis	CFSR	37.0	222.9	0.87	0.85	−0.6	−1.6
	ERA-Interim	39.6	257.0	0.86	0.79	+14.6	+0.9
	ERA5	45.5 (40.1)	283.6 (241.5)	0.88 (0.87)	0.78 (0.65)	+23.7 (+14.3)	+6.9 (+1.9)
	JRA-55	36.8	190.0	0.83	0.65	+19.7	−2.0
Interpolated	CPC-Unified	23.0	158.9	0.81	0.73	−33.7	−15.6
	CRU TS4.01	26.0	144.6	0.82	0.69	−21.6	−12.6
	PREC/L	11.4	98.9	0.68	0.52	−61.8	−27.2
	GPCCv8	40.2	281.5	0.91	0.89	+5.1	+1.6
Multi-source	MSWEP V2	21.6 (19.8)	161.2 (141.9)	0.92 (0.87)	0.34 (0.17)	−29.3 (−34.7)	−17 (−19.9)
	HydroGFD2.0	40.8	290.9	0.85	0.80	+15.6	+2.2
	CHIRPS2.0	30.1 (29.3)	187.8 (182.1)	0.86 (0.83)	0.79 (0.58)	−12.2 (−13.8)	−8.5 (−11.1)
	PERSIANN-CDR	29.6 (29.5)	213.7 (202.8)	0.88 (0.87)	0.73 (0.68)	−7.5 (−5.0)	−9.0 (−9.5)
	SM2RAIN-CCI[a]	16.8 (15.6)	118.8 (118.1)	0.82 (0.80)	0.68 (0.73)	−11.3 (2.2)	−3.8 (−1.4)

Note: Values in parenthesis denote statistics based on the dataset's original resolution.
[a]Note that SM2RAIN-CCI metrics are afflicted by missing values.

elevation grid cells (Figure S3). The overall similar results indicate no or little importance of elevation for the seasonal performance patterns.

5.2 | All-basin summary evaluation of precipitation products

Results for the statistical analysis are presented as basin averages in Table 2. In terms of mean values, ERA-Interim, CFSR, and GPCCv8 agree best with the observed gauge values, while reanalyses perform generally well. The spread is large especially between the interpolated datasets, which is likely due to the different selection of gauges included in each individual dataset. Even though HydroGFD2.0 uses monthly anomalies from CPC-Unified, we find different mean values here because the climatology is derived from CHPclim, which includes more stations (Funk *et al.*, 2015).

As seen in Figure 2, most products underestimate precipitation over the basin. The underestimation of precipitation in CHIRPS2.0 could stem from the low density of employed stations. The underestimation found in PERSIANN-CDR might be related to the bias-adjustment, which is based on the GPCP dataset with a rather coarse 2.5° resolution. Such underestimation over mountainous regions in Iran has also been reported for other precipitation datasets which include satellite estimates (Moazami *et al.*, 2016; Alijanian *et al.*, 2017; Katiraie-Boroujerdy *et al.*, 2017). It might be associated with infrared sensors

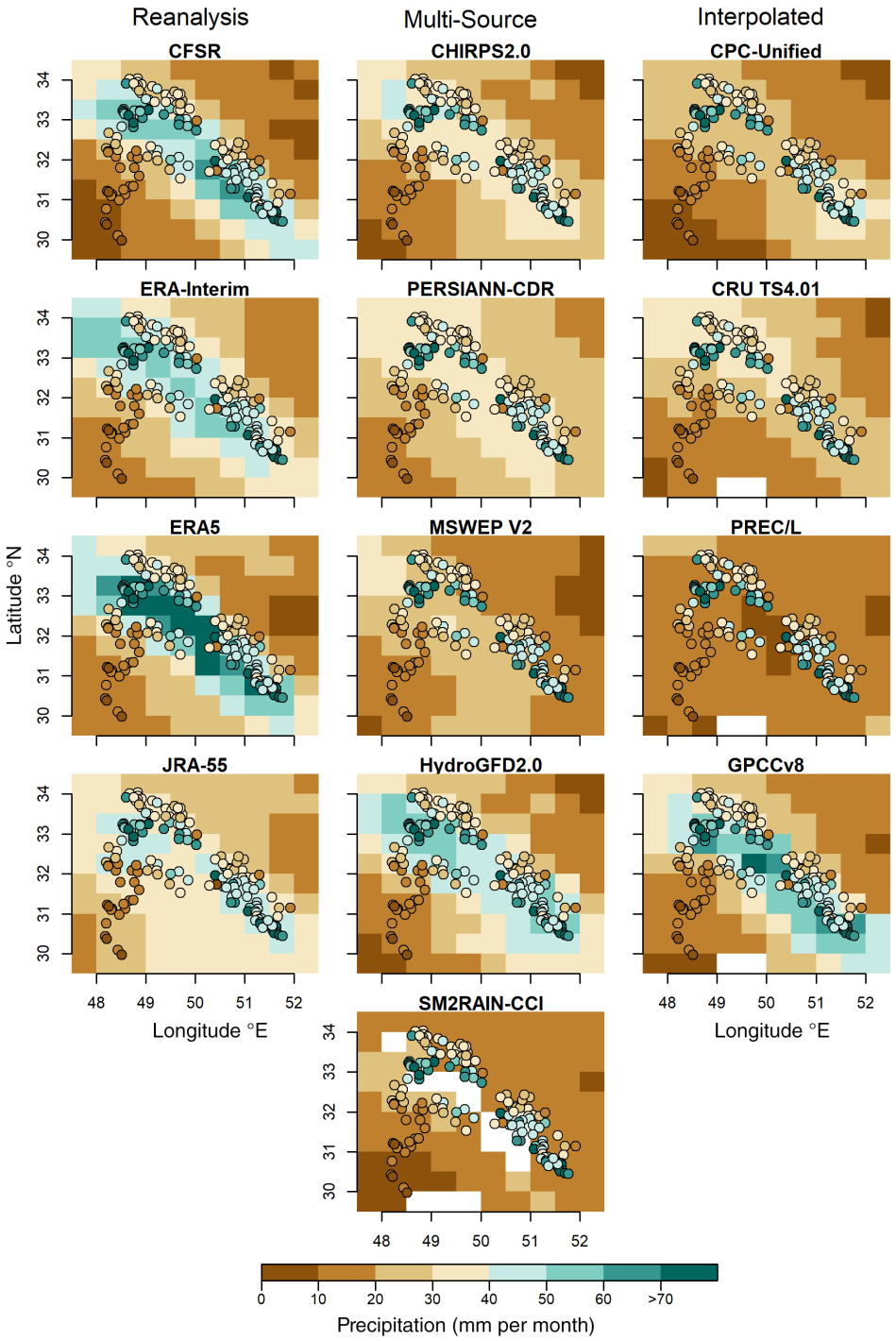

FIGURE 3 Mean precipitation (mm per month) over the basin for the 2000–2015 period. Background colours indicate precipitation of the respective gridded product while the circles denote observations

having difficulties to detect warm cloud process especially over mountains (Dinku *et al.*, 2018; O and Kirstetter, 2018). The underestimation in SM2RAIN-CCI could be related to the fact that once the surface soil is saturated after intense or long-lasting precipitation events, no additional increase is possible such that no additional precipitation can be deduced from the soil moisture data (Brocca *et al.*, 2013). According to Alijanian *et al.*, 2017, MSWEP V1 highly overestimated precipitation values over different climatic zones of Iran while this study interestingly indicates contrary results for MSWEP V2. This might have to do with a reduction of Iranian bias correction factor due to suspected issues with observed runoff data (Beck *et al.*, 2019).

Furthermore, for maximum precipitation, HydroGFD2.0, GPCCv8, and ERA5 agree best with in situ measurements whereas the other interpolated datasets, as well as MSWEP V2 and SM2RAIN-CCI, show large underestimations. Considering the mean relative error over Karun basin, the CFSR reanalysis shows the best agreement with gauge data, followed by GPCCv8. Merged datasets outperform some reanalyses which may be attributed to a

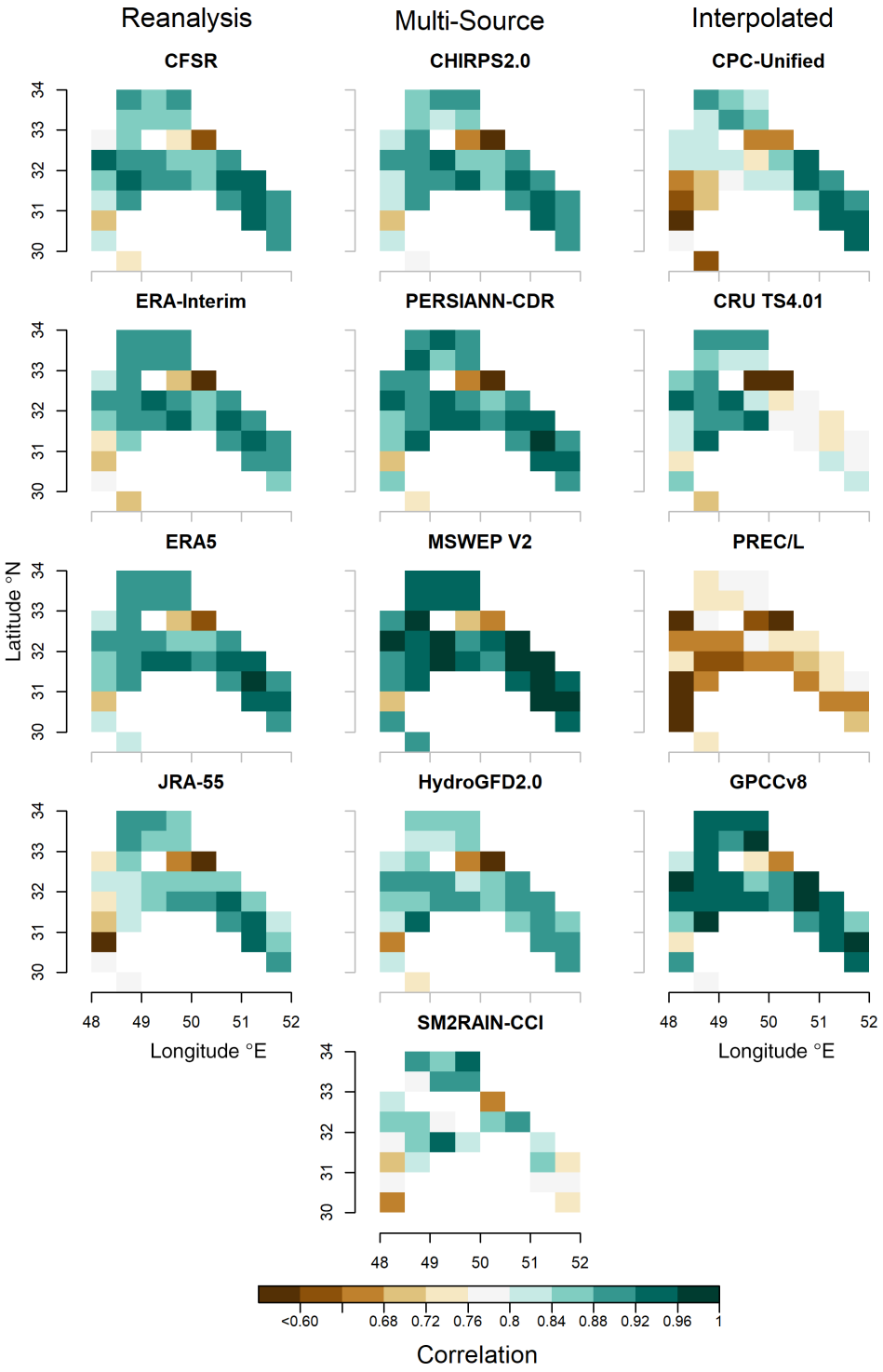

FIGURE 4 Temporal correlation between different global precipitation datasets and local gauge data

large overestimation of these reanalyses at low-altitude regions (Figure S3).

The spatial correlation results indicate the best agreement of GPCCv8 and CFSR with gauge data, but also reanalysis datasets (except JRA-55) show similarly high correlations whereas MSWEP V2 and PREC/L, show particularly poor performance. We also explore the performance of the products in their original spatial resolution (shown in Table 2 with parenthesis). Overall, we find a minor role of the spatial resolution, except for spatial correlation where results seem to be more sensitive.

5.3 | Spatial evaluation of precipitation datasets

In addition to the all-basin evaluation of the considered datasets in the previous section, in this section, we

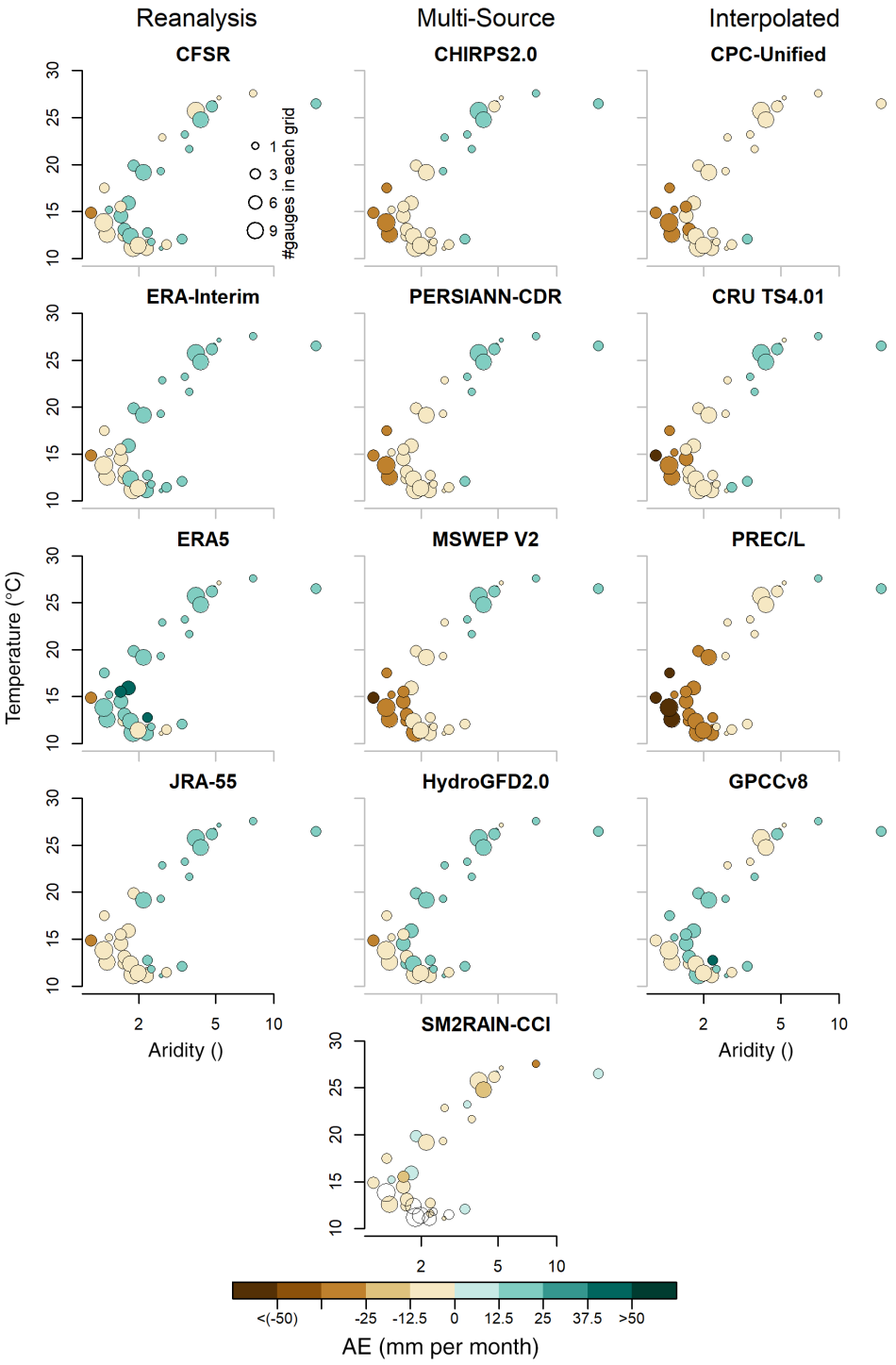

FIGURE 5 Bias values (mm per month) of datasets and the number of gauges within each grid cell (size of the circles) in the different climatic zone of the basin (circles with no colour means missing data)

analyse the spatial variability of datasets versus gauge measurements. Figure 3 presents dataset versus gauge-based mean precipitation across the Karun basin. The above-mentioned underestimation mostly occurs in mountainous regions. In general, reanalysis datasets along with GPCCv8 and HydroGFD2.0 agree best with gauge-derived rainfall patterns over the basin, confirming results from Table 2. Aside from the overall biases, we find that MSWEP V2 and PREC/L have most difficulties in capturing the spatial precipitation patterns, including the contrast between low areas and the mountains (relative error of products in each station are presented in Figure S4).

In Figure 4, we analyse the spatial pattern of the temporal correlations evaluated in Table 2. We find overall highest correlations for GPCCv8 and MSWEP V2. For the latter, this was reported earlier by Alijanian *et al.* (2017), who attributed this fact to either applying high weights to gauge observations in MSWEP, or the inclusion of reanalysis precipitation data. Most datasets show high

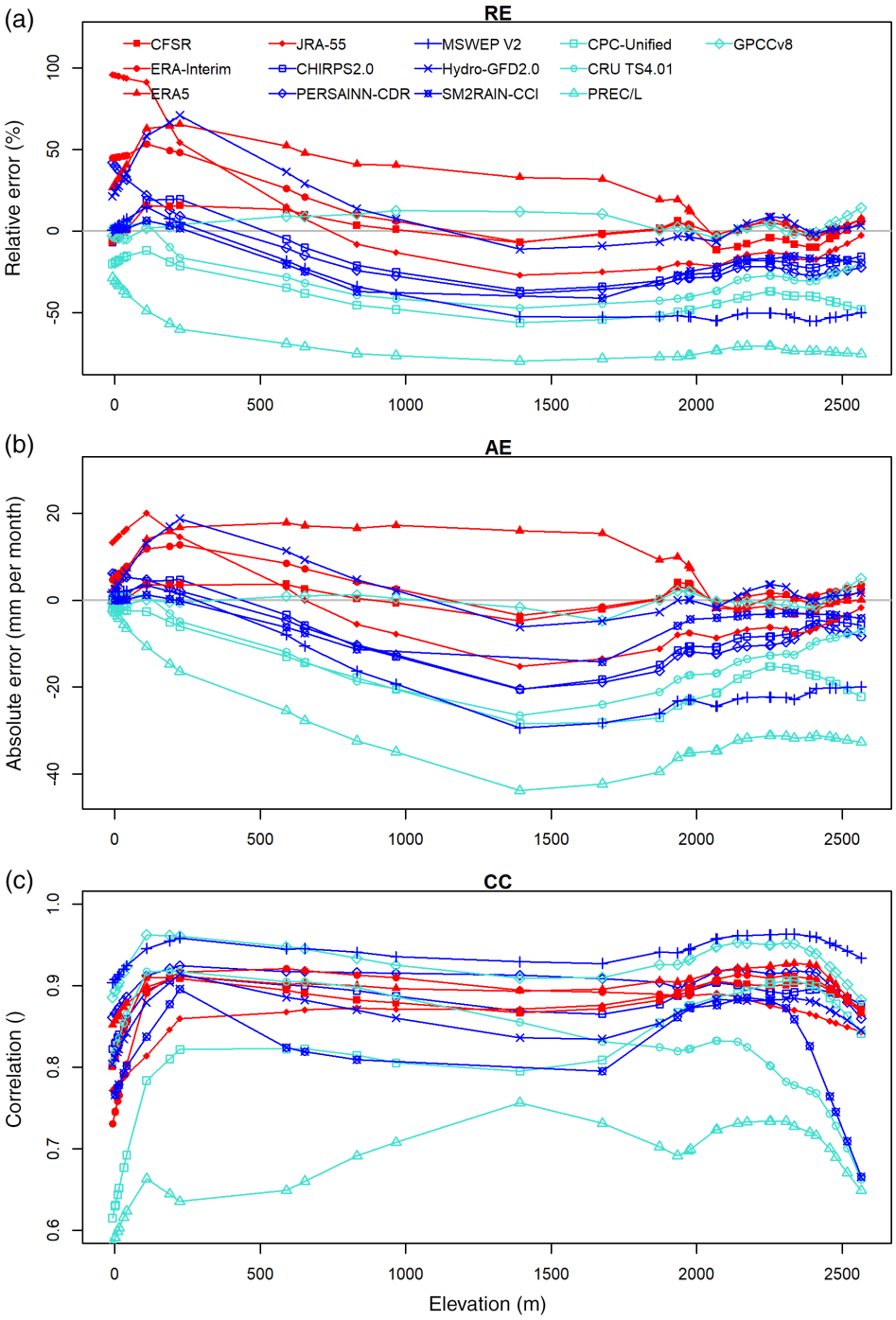

FIGURE 6 (a) Absolute error, (b) relative error and (c) temporal correlation as a function of elevation. Symbols indicate the elevations of the grid cells. Lines are smoothed with a LOWESS filter

spatial variability in the correlation results across the basin, except for GPCCv8 and MSWEP V2 where correlation is generally high, and for PREC/L where correlation is low everywhere.

5.4 | Performance in different climates and elevations

Figure 5 illustrates the absolute error of the considered precipitation datasets with respect to climate in the considered grid cells. Further, the number of gauges located within each grid cell is reflected by the size of the points; smallest circles indicate one gauge per grid cell, while largest circles refer to nine gauges. Performance of the datasets in each grid cell is shown through colour coding. We find no systematic difference in the results indicated by small versus large circles. This suggests that the different number of gauges in each grid cell has only a negligible impact on our conclusions.

The previously mentioned underestimation in CHIRPS2.0, PERSIANN-CDR and interpolated gauge

(CPC-Unified and CRU TS4.01) datasets, as well as MSWEP V2, is mostly found in regions with comparatively cold and wet climates; regions with elevations above 2,500 m as opposed to the low altitude surfaces. GPCCv8 and ERA5 generally show more similar performance across the different climates. Most datasets agree better with gauge observations in dry and warm climate. This underlines the importance of considering the orographic and geographic effects on the quality of precipitation data (Hu *et al*., 2016). Similar results are found for relative errors of the datasets (Figure S5); while the climate sensitivity of correlation performance interestingly is overall lower (Figure S6).

The elevation effect on the agreement between gauge data and precipitation products is analysed in Figure 6, using each grid cell's elevation according to a 0.5° digital elevation map (Amante and Eakins, 2009). Overall, the performance of the considered products varies similarly to altitude. Reanalysis datasets tend to overestimate precipitation at low altitudes while they show no or little bias at higher altitudes, despite the challenging topographic complexity associated with high elevation regions (Isotta *et al*., 2015). All other datasets show largest biases at medium altitudes and slightly improved results for the lowest and highest elevation regions. Interestingly, the temporal correlation analysis (Figure 6c) reveals opposite results; highest correlations are found at medium altitudes and performances are degraded in the lowest and highest grid cells. These performance differences are controlled by altitude-dependent (a) topographic complexity and (b) climate (Figure S7).

Figure 7 illustrates which dataset agrees best with gauge observations in each grid cell and climate. For this purpose, a ranking of datasets is computed in each grid cell according to (a) temporal correlation and (b) absolute error. Then, the dataset with the lowest sum of the two ranks is selected as the best dataset for a particular grid cell. GPCCv8 performs best in 15 grid cells, thereby clearly outperforming all other datasets. Among the remaining datasets, CFSR and ERA5 stand out with the best agreement against gauge data in three grid cells each. The outstanding performance of GPCCv8 in this context is probably due to the fact that this dataset employs some of the gauge data used as reference in this study. While this is true for other (interpolated) datasets, they probably could use the gauge information more efficiently in their derivation procedure.

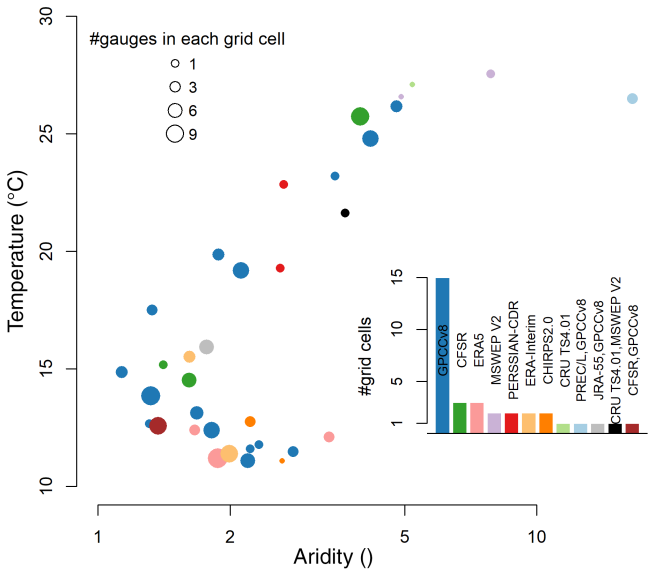

FIGURE 7 Datasets showing best agreement with gauge data in each grid with respect to correlation and absolute error. The number of gauges within each grid cell is denoted by the size of the points

6 | CONCLUSIONS

Known discrepancies between state-of-the-art precipitation datasets have motivated us to assess the accuracy of a great variety of state-of-the-art precipitation datasets against gauge observations in a comparatively observation-sparse region, Karun basin in Iran. Thereby, we analysed the spatiotemporal variability and find characteristic strengths and weaknesses of each dataset, while no single dataset is superior in all respects.

The overall best agreement with observations was found for the GPCCv8 dataset. However, GPCCv8 is likely biased to better results as it might include a large part of the gauge data used for the evaluation. Therefore, the result rather points out that a comprehensive gauge selection is most important for the quality of any large-scale precipitation dataset.

While merged products include gauge data for calibration, reanalyses are independent of the gauge data since they do not assimilate surface gauge precipitation. Among the latter, ERA5, ERA-Interim, and CFSR outperform JRA-55 over Karun basin. Given this rather independent nature of the reanalyses, the results do show value in such data, particularly in regions where no gauge data are available.

Comparing older, established datasets with more recent products we find mixed results. While ERA5 shows overall improved agreement with gauge measurements over ERA-Interim, MSWEP V2 shows significant biases even though with the opposite sign as the previous version (Alijanian *et al*., 2017). Further, MSWEP V2 and HydroGFD2.0 as the most recent of the considered products are not outperforming previously released datasets.

A caveat for our analysis is potential errors in gauge measurements. Especially at higher altitudes, precipitation under-catch in such measurements is a known issue (Mekonnen *et al*., 2015). Therefore, the underestimations we find in most considered datasets are likely even more considerable than shown here. In addition, it should be noted that the gauge measurements used in this study have (partly) been employed in the derivation of (some of) the considered datasets. Despite this, we find no consistently improved agreement between the gauge measurements and datasets that use more of them versus datasets that do not use them.

The identification of such important shortcomings of state-of-the-art datasets highlights potential avenues for future development. This way, the present study contributes to more reliable, high-quality precipitation datasets which are key to hydro-climatological monitoring and modelling, especially given the potential increase of related extremes in the context of climate change.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers for their valuable comments. Further, we acknowledge IRIMO (Islamic Republic of Iran Meteorological Organization) for providing gauge observations. A.F. acknowledges financial support from Ministry of Science, Research and Technology I.R. of Iran, and also the support in the form of hosting and supervision provided by the Swedish Meteorological and Hydrological Institute in Norrköping, Sweden, as well as the Max Planck Institute for Biogeochemistry in Jena, Germany. R.O. and S.O acknowledge funding support by the German Research Foundation (Emmy Noether grant number 391059971). Moreover, we thank Hylke Beck, Maryam Dehghani, and Mohammadali Alijanian for their comments and suggestions.

ORCID

Ali Fallah https://orcid.org/0000-0001-8273-4638
Sungmin O https://orcid.org/0000-0002-7364-2122

REFERENCES

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P. and Nelkin, E. (2003) The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *Journal of Hydrometeorology*, 4(6), 1147–1167. https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

AghaKouchak, A., Farahmand, A., Melton, F.S., Teixeira, J., Anderson, M.C., Wardlow, B.D. and Hain, C.R. (2015) Remote sensing of drought: progress, challenges and opportunities. *Reviews of Geophysics*, 53(2), 452–480. https://doi.org/10.1002/2014rg000456.

Alijanian, M., Rakhshandehroo, G.R., Mishra, A.K. and Dehghani, M. (2017) Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran. *International Journal of Climatology*, 37(14), 4896–4914. https://doi.org/10.1002/joc.5131.

Alijanian, M., Reza Rakhshandehroo, G., Mishra, A. and Dehghani, M. (2019) Evaluation of remotely sensed precipitation estimates using PERSIANN-CDR and MSWEP for spatio-temporal drought assessment over Iran. *Journal of Hydrology*, 579, 124189. https://doi.org/10.1016/j.jhydrol.2019.124189.

Amante, C. and Eakins, B.W. (2009) ETOPO1 1 Arc-Minute Global Relief Model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA. https://doi.org/10.7289/V5C8276M.

Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J.C.M., Hasan, A. and Pineda, L. (2019) Global catchment modelling using World-Wide HYPE (WWH), open data and stepwise parameter estimation. *Hydrology and Earth System Sciences Discussions*, 2019, 1–34. https://doi.org/10.5194/hess-2019-111.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R. and Prat, O.P. (2015) PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96(1), 69–83. https://doi.org/10.1175/bams-d-13-00068.1.

Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Blyth, E., Bousserez, N., Boussetta, S., Brown, A., Buizza, R., Buontempo, C., Chevallier, F., Choulga, M., Cloke, H., Cronin, M.F., Dahoui, M., De Rosnay, P., Dirmeyer, P.A., Drusch, M., Dutra, E., Ek, M.B., Gentine, P., Hewitt, H., Keeley, S.P.E., Kerr, Y., Kumar, S., Lupu, C., Mahfouf, J.-F., McNorton, J., Mecklenburg, S., Mogensen, K., Muñoz-Sabater, J., Orth, R., Rabier, F., Reichle, R., Ruston, B., Pappenberger, F., Sandu, I., Seneviratne, S.I., Tietsche, S., Trigo, I.F., Uijlenhoet, R., Wedi, N., Woolway, R.I. and Zeng, X. (2018) Satellite and in situ observations for advancing global earth surface modelling: a review. *Remote Sensing*, 10(12), 2038. https://doi.org/10.3390/rs10122038.

Beaulieu, C., Seidou, O., Ouarda, T.B.M.J., Zhang, X., Boulet, G. and Yagouti, A. (2008) Intercomparison of homogenization techniques for precipitation data. *Water Resources Research*, 44, W02425. https://doi.org/10.1029/2006wr005615.

Beck, H.E., van Dijk, A.I., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B. and de Roo, A. (2017a) MSWEP: 3-hourly 0.25 global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21(1), 589. https://doi.org/10.5194/hess-21-589-2017.

Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A.I.J.M., Weedon, G.P., Brocca, L., Pappenberger, F., Huffman, G.J. and Wood, E.F. (2017b) Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, 21(12), 6201–6217. https://doi.org/10.5194/hess-21-6201-2017.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., van Dijk, A.I.J.M., McVicar, T.R. and Adler, R.F. (2019) MSWEP V2

Global 3-hourly 0.1° precipitation: methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, 100(3), 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U. and Ziese, M. (2013) A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, 5(1), 71–99. https://doi.org/10.5194/essd-5-71-2013.

Berg, P., Donnelly, C. and Gustafsson, D. (2018) Near-real-time adjusted reanalysis forcing data for hydrology. *Hydrology and Earth System Sciences*, 22(2), 989–1000. https://doi.org/10.5194/hess-22-989-2018.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W. and Levizzani, V. (2014) Soil as a natural rain gauge: estimating global rainfall from satellite soil moisture data. *Journal of Geophysical Research. Atmospheres*, 119, 5128–5141. https://doi.org/10.1002/2014JD021489.

Brocca, L., Moramarco, T., Melone, F. and Wagner, W. (2013) A new method for rainfall estimation through soil moisture observations. *Geophysical Research Letters*, 40(5), 853–858. https://doi.org/10.1002/grl.50173.

Budyko, M.I. (1974) In: Miller, D.H. (Ed.) *Climate and Life (English edition)*. New York: Academic Press, 508 p. Available at: https://trove.nla.gov.au/version/26211710.

Chen, C.-T. and Knutson, T. (2008) On the verification and comparison of extreme rainfall indices from climate models. *Journal of Climate*, 21(7), 1605–1621. https://doi.org/10.1175/2007JCLI1494.1.

Chen, M., Shi, W., Xie, P., Silva, V.B.S., Kousky, V.E., Wayne Higgins, R. and Janowiak, J.E. (2008) Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research. Atmospheres*, 113, D04110. https://doi.org/10.1029/2007jd009132.

Chen, M., Xie, P., Janowiak, J.E. and Arkin, P.A. (2002) Global land precipitation: a 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology*, 3(3), 249–266. https://doi.org/10.1175/1525-7541(2002)003<0249:glpaym>2.0.co;2.

Ciabatta, L., Massari, C., Brocca, L., Gruber, A., Reimer, C., Hahn, S., Paulik, C., Dorigo, W., Kidd, R. and Wagner, W. (2018) SM2RAIN-CCI: a new global long-term rainfall data set derived from ESA CCI soil moisture. *Earth System Science Data*, 10(1), 267–280. https://doi.org/10.5194/essd-10-267-2018.

Copernicus Climate Change Service. (2017) *ERA5: fifth generation of ECMWF atmospheric reanalyses of the global climate*. Copernicus Climate Change Service Climate Data Store (CDS). Available at: https://cds.climate.copernicus.eu/cdsapp#!/home [Accessed 13 December 2019].

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. https://doi.org/10.1002/qj.828.

Dezfooli, D., Abdollahi, B., Hosseini-Moghari, S.-M. and Ebrahimi, K. (2018) A comparison between high-resolution satellite precipitation estimates and gauge measured data: case study of Gorganrood basin, Iran. *Journal of Water Supply Research and Technology*, 67(3), 236–251. https://doi.org/10.2166/aqua.2018.062.

Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S.J. and Ropelewski, C.F. (2008) Validation of high-resolution satellite rainfall products over complex terrain. *International Journal of Remote Sensing*, 29(14), 4097–4110. https://doi.org/10.1080/01431160701772526.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H. and Ceccato, P. (2018) Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society*, 144(S1), 292–312. https://doi.org/10.1002/qj.3244.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J. (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2(1), 150066. https://doi.org/10.1038/sdata.2015.66.

Gebregiorgis, A.S., Kirstetter, P.-E., Hong, Y.E., Carr, N.J., Gourley, J.J., Petersen, W. and Zheng, Y. (2017) Understanding overland multisensor satellite precipitation error in TMPA-RT products. *Journal of Hydrometeorology*, 18(2), 285–306. https://doi.org/10.1175/JHM-D-15-0207.1.

Ghajarnia, N., Liaghat, A. and Arasteh, P.D. (2015) Comparison and evaluation of high-resolution precipitation estimation products in Urmia Basin-Iran. *Atmospheric Research*, 158, 50–65. https://doi.org/10.1016/j.atmosres.2015.02.010.

González-Rouco, J.F., Jiménez, J.L., Quesada, V. and Valero, F. (2001) Quality control and homogeneity of precipitation data in the southwest of Europe. *Journal of Climate*, 14(5), 964–978. https://doi.org/10.1175/1520-0442(2001)014<0964:qcahop>2.0.co;2.

Hameed, M., Ahmadalipour, A. & Moradkhani, H. (2020) Drought and food security in the middle east: An analytical framework. *Agricultural and Forest Meteorology*, 281, 107816. https://doi.org/10.1016/j.agrformet.2019.107816.

Harris, I., Jones, P.D., Osborn, T.J. and Lister, D.H. (2014) Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. *International Journal of Climatology*, 34 (3), 623–642. https://doi.org/10.1002/joc.3711.

Hosseini-Moghari, S.-M., Araghinejad, S. and Ebrahimi, K. (2018) Spatio-temporal evaluation of global gridded precipitation datasets across Iran. *Hydrological Sciences Journal*, 63(11), 1669–1688. https://doi.org/10.1080/02626667.2018.1524986.

Hu, Z., Hu, Q., Zhang, C., Chen, X. and Li, Q. (2016) Evaluation of reanalysis, spatially interpolated and satellite remotely sensed precipitation data sets in central Asia. *Journal of Geophysical Research. Atmospheres*, 121(10), 5648–5663. https://doi.org/10.1002/2016JD024781.

Isotta, F.A., Vogel, R. and Frei, C. (2015) Evaluation of European regional reanalyses and downscalings for precipitation in the Alpine region. *Meteorologische Zeitschrift*, 24(1), 15–37. https://doi.org/10.1127/metz/2014/0584.

Japanese 55-year Reanalysis. (2013) JRA-55: monthly means and variances. https://doi.org/10.5065/D60G3H5B

Jones, P.W. (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, 127(9), 2204–2210. https://doi.org/10.1175/1520-0493(1999)127<2204:fasocr>2.0.co;2.

Katiraie-Boroujerdy, P.-S., Akbari Asanjan, A., Chavoshian, A., Hsu, K.I. and Sorooshian, S. (2019) Assessment of seven CMIP5 model precipitation extremes over Iran based on a satellite-based climate data set. *International Journal of Climatology*, 39 (8), 3505–3522. https://doi.org/10.1002/joc.6035.

Katiraie-Boroujerdy, P.-S., Akbari Asanjan, A., Hsu, K.-I. and Sorooshian, S. (2017) Intercomparison of PERSIANN-CDR and TRMM-3B42V7 precipitation estimates at monthly and daily time scales. *Atmospheric Research*, 193, 36–49. https://doi.org/10.1016/j.atmosres.2017.04.005.

Katiraie-Boroujerdy, P.-S., Nasrollahi, N., Hsu, K.-I. and Sorooshian, S. (2013) Evaluation of satellite-based precipitation estimation over Iran. *Journal of Arid Environments*, 97, 205–219. https://doi.org/10.1016/j.jaridenv.2013.05.013.

Khodadoust Siuki, S., Saghafian, B. and Moazami, S. (2017) Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products. *International Journal of Remote Sensing*, 38(2), 558–571. https://doi.org/10.1080/01431161.2016.1268735.

Kirtman, B., Power, S.B., Adedoyin, J.A., Boer, G.J., Bojariu, R., Camilloni, I., Doblas-Reyes, F.J., Fiore, A.M., Kimoto, M., Meehl, G.A., Prather, M., Sarr, A., Schär, C., Sutton, R., van Oldenborgh, G.J., Vecchi, G. and Wang, H.J. (2013) Near-term climate change: projections and predictability. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (Eds.) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge/New York: Cambridge University Press, pp. 953–1028. https://doi.org/10.1017/CBO9781107415324.023.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K. and Takahashi, K. (2015) The JRA-55 reanalysis: general specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Series II*, 93(1), 5–48. https://doi.org/10.2151/jmsj.2015-001.

Koster, R.D., Brocca, L., Crow, W.T., Burgin, M.S. and De Lannoy, G.J.M. (2016) Precipitation estimation using L-band and C-band soil moisture retrievals. *Water Resources Research*, 52(9), 7213–7225. https://doi.org/10.1002/2016wr019024.

Mekonnen, G.B., Matula, S., Doležal, F. and Fišák, J. (2015) Adjustment to rainfall measurement under catch with a tipping-bucket rain gauge using ground-level manual gauges. *Meteorology and Atmospheric Physics*, 127(3), 241–256. https://doi.org/10.1007/s00703-014-0355-z.

Moazami, S., Golian, S., Hong, Y., Sheng, C. & Kavianpour, M. R. (2016) Comprehensive evaluation of four high-resolution satellite precipitation products under diverse climate conditions in Iran. *Hydrological Sciences Journal*, 61(2), 420–440. https://doi.org/10.1080/02626667.2014.987675.

Moazami, S., Golian, S., Kavianpour, M.R. and Hong, Y. (2014) Uncertainty analysis of bias from satellite rainfall estimates using copula method. *Atmospheric Research*, 137, 145–166. https://doi.org/10.1016/j.atmosres.2013.08.016.

Naumann, G., Dutra, E., Barbosa, P., Pappenberger, F., Wetterhall, F. and Vogt, J.V. (2014) Comparison of drought indicators derived from multiple data sets over Africa. *Hydrology and Earth System Sciences*, 18(5), 1625–1640. https://doi.org/10.5194/hess-18-1625-2014.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O.B., Déqué, M., Fernandez, J., Hänsler, A., Meijgaard, E.v., Samuelsson, P., Sylla, M.B. and Sushama, L. (2012) Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *Journal of Climate*, 25(18), 6057–6078. https://doi.org/10.1175/JCLI-D-11-00375.1.

O, S. and Kirstetter, P.-E. (2018) Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. *Quarterly Journal of the Royal Meteorological Society*, 144(S1), 270–281. https://doi.org/10.1002/qj.3218.

Orth, R. and Destouni, G. (2018) Drought reduces blue-water fluxes more strongly than green-water fluxes in Europe. *Nature Communications*, 9, 3602. https://doi.org/10.1038/s41467-018-06013-7.

Parker, W.S. (2016) Reanalyses and observations: What's the difference? *Bulletin of the American Meteorological Society*, 97(9), 1565–1572. https://doi.org/10.1175/BAMS-D-14-00226.1.

Saeidizand, R., Sabetghadam, S., Tarnavsky, E. and Pierleoni, A. (2018) Evaluation of CHIRPS rainfall estimates over Iran. *Quarterly Journal of the Royal Meteorological Society*, 144(S1), 282–291. https://doi.org/10.1002/qj.3342.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-y., Juang, H.-M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P.V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Dool, H.v.d., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G. and Goldberg, M. (2010) The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8), 1015–1058. https://doi.org/10.1175/2010bams3001.1.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M. and Becker, E. (2012) NCEP Climate Forecast System version 2 (CFSv2) monthly products. https://doi.org/10.5065/D69021ZF.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., Dool, H.v.d., Zhang, Q., Wang, W., Chen, M. and Becker, E. (2014) The NCEP Climate Forecast System version 2. *Journal of Climate*, 27(6), 2185–2208. https://doi.org/10.1175/jcli-d-12-00823.1.

Salmani-Dehaghi, N. and Samani, N. (2019) Spatiotemporal assessment of the PERSIANN family of satellite precipitation data over Fars Province, Iran. *Theoretical and Applied Climatology*, 138(3), 1333–1357. https://doi.org/10.1007/s00704-019-02872-2.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A. and Ziese, M. (2018) GPCC Full Data Monthly Product version 2018 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historical data. https://doi.org/10.5676/DWD_GPCC/FD_M_V2018_050

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M. and Rudolf, B. (2014) GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, 115(1), 15–40. https://doi.org/10.1007/s00704-013-0860-x.

Schulzweida, U. (2019, February 6) CDO User Guide (Version 1.9.6). Zenodo. https://doi.org/10.5281/zenodo.2558193.

Sharifi, E., Steinacker, R. and Saghafian, B. (2016) Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: preliminary results. *Remote Sensing*, 8(2), 135. https://doi.org/10.3390/rs8020135.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S. and Hsu, K.-L. (2018) A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1), 79–107. https://doi.org/10.1002/2017rg000574.

Vaghefi, S.A., Keykhai, M., Jahanbakhshi, F., Sheikholeslami, J., Ahmadi, A., Yang, H. and Abbaspour, K.C. (2019) The future of extreme climate in Iran. *Scientific Reports*, 9(1), 1464. https://doi.org/10.1038/s41598-018-38071-8.

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., García-Vera, M.A. and Stepanek, P. (2010) A complete daily precipitation database for northeast Spain: reconstruction, quality control, and homogeneity. *International Journal of Climatology*, 30 (8), 1146–1163. https://doi.org/10.1002/joc.1850.

Weedon, G.P., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., Adam, J.C., Bellouin, N., Boucher, O. and Best, M. (2011) Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology*, 12 (5), 823–848. https://doi.org/10.1175/2011jhm1369.

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y. and Liu, C. (2007) A gauge-based analysis of daily precipitation over East Asia. *Journal of Hydrometeorology*, 8(3), 607–626. https://doi.org/10.1175/JHM583.1.

Yamamoto, J.K. (2000) An alternative measure of the reliability of ordinary kriging estimates. *Mathematical Geology*, 32(4), 489–509. https://doi.org/10.1023/a:1007577916868.

Zhan, W., Guan, K., Sheffield, J. and Wood, E.F. (2016) Depiction of drought over sub-Saharan Africa using reanalyses precipitation data sets. *Journal of Geophysical Research. Atmospheres*, 121, 10555–10574. https://doi.org/10.1002/2016JD024858.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Fallah A, Rakhshandehroo GR, Berg P, O S, Orth R. Evaluation of precipitation datasets against local observations in southwestern Iran. *Int J Climatol*. 2020;40:4102–4116. https://doi.org/10.1002/joc.6445