UC Irvine
UC Irvine Previously Published Works

Title

Rainfall frequency analysis for ungauged regions using remotely sensed precipitation information

Permalink

https://escholarship.org/uc/item/7k06p11f

Authors

Faridzad, M
Yang, T
Hsu, K
et al.

Publication Date

2018-08-01

DOI

10.1016/j.jhydrol.2018.05.071

Research papers

Rainfall frequency analysis for ungauged regions using remotely sensed precipitation information

Mohammad Faridzad[a], Tiantian Yang[a,b,*], Kuolin Hsu[a], Soroosh Sorooshian[a], Chan Xiao[c]

[a] Department of Civil and Environmental Engineering, Center for Hydrometeorology and Remote Sensing [CHRS], University of California-Irvine, Irvine, CA, USA
[b] Deltares USA Inc., Silver Spring, MD, USA
[c] National Climate Center, China Meteorological Administration, Beijing, China

ABSTRACT

Rainfall frequency analysis, which is an important tool in hydrologic engineering, has been traditionally performed using information from gauge observations. This approach has proven to be a useful tool in planning and design for the regions where sufficient observational data are available. However, in many parts of the world where ground-based observations are sparse and limited in length, the effectiveness of statistical methods for such applications is highly limited. The sparse gauge networks over those regions, especially over remote areas and high-elevation regions, cannot represent the spatiotemporal variability of extreme rainfall events and hence preclude developing depth-duration-frequency curves (DDF) for rainfall frequency analysis. In this study, the PERSIANN-CDR dataset is used to propose a mechanism, by which satellite precipitation information could be used for rainfall frequency analysis and development of DDF curves. In the proposed framework, we first adjust the extreme precipitation time series estimated by PERSIANN-CDR using an elevation-based correction function, then use the adjusted dataset to develop DDF curves. As a proof of concept, we have implemented our proposed approach in 20 river basins in the United States with different climatic conditions and elevations. Bias adjustment results indicate that the correction model can significantly reduce the biases in PERSIANN-CDR estimates of annual maximum series, especially for high elevation regions. Comparison of the extracted DDF curves from both the original and adjusted PERSIANN-CDR data with the reported DDF curves from NOAA Atlas 14 shows that the extreme percentiles from the corrected PERSIANN-CDR are consistently closer to the gauge-based estimates at the tested basins. The median relative errors of the frequency estimates at the studied basins were less than 20% in most cases. Our proposed framework has the potential for constructing DDF curves for regions with limited or sparse gauge-based observations using remotely sensed precipitation information, and the spatiotemporal resolution of the adjusted PERSIANN-CDR data provides valuable information for various applications in remote and high elevation areas.

1. Introduction

Rainfall Frequency Analysis (RFA) is an important tool in hydrologic engineering (Bonnin et al., 2006; Hosking and Wallis, 2005; Stedinger, 1993). Depth-duration-frequency (DDF) curves, which link extreme rainfall depths to their probability of occurrence, are based on time series of extreme rainfall with different durations fitted with probability distribution functions. RFA has been traditionally performed using information from rain gauges. This approach has proven to be a useful tool in planning and design for regions where observational data is relatively abundant such as the United States or Europe. However, many parts of the world, particularly the developing

countries, do not have that advantage. In many developing countries, gauge observation networks over remote and mountainous regions are still sparse and limited in terms of duration.

With advances in tools and techniques for precipitation measurement using remotely sensed information, investigation of rainfall characteristics over remote and mountainous regions with limited gauge observations has become possible. In an effort to produce long and consistent climate records based on satellite observations, National Oceanic and Atmospheric Association (NOAA) under the Climate Data Record (CDR) program, in cooperation with the University of California, Irvine, developed a satellite precipitation product named the Precipitation Estimation from Remotely Sensed Information and

Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) (Ashouri et al., 2015). PERSIANN-CDR provides near-global (60°N to 60°S latitude and 0° to 360° longitude) precipitation information with 0.25° spatial and daily temporal resolution from 1983 to the present. Given its relatively high spatial resolution and long record, PERSI-ANN-CDR is a unique dataset for studying extreme precipitations and performing rainfall frequency analysis. The length of the PERSI-ANN-CDR dataset (34 + years) is particularly valuable for parts of the world that lack the gauge information for rainfall frequency analysis.

In recent years, several efforts have been made to develop DDF curves by employing remotely sensed precipitation information from weather radars and earth observing satellites (Eldardiry et al., 2015; Marra and Morin, 2015; Overeem et al., 2008; Overeem et al., 2009; Wright et al., 2013). For instance, Overeem et al. (2009) used an 11-year gauge-adjusted radar-rainfall dataset and performed a regional frequency analysis to extract DDF curves for the Netherlands. They found that radar data, despite being useful for real-time rainfall analysis, still suffer from serious limitations, such as significant errors in extreme rainfall estimates and shortness of data, that limit their usefulness for RFA. Thus, the application of radar data for rainfall frequency analysis is hampered by: (1) its relatively short length of record which leads to sampling issues during distribution fitting process and results in larger uncertainties of the frequency estimates especially for longer durations; and, (2) estimation uncertainties and heterogeneities due to the continuous development of radar quantitative precipitation estimation (QPE) instruments and methods (Allen and DeGaetano, 2005; Lombardo et al., 2006). Eldardiry et al. (2015) quantified the effects of each of these sources of uncertainty and attributed much of the quantile estimation uncertainty to the length of the dataset. However, the conditional bias intrinsic to the radar dataset was the main reason for the observed systematic underestimations in the rainfall frequency estimates. As compared to rain gauges and radar network, satellite QPE is able to provide global coverage and has been employed in a number of studies for rainfall frequency analysis (Awadallah et al., 2011; Endreny and Imbeah, 2009; Marra et al., 2017; Zhou et al., 2015). Yet, similar to radars, the application satellite QPEs for RFA is undermined by the data length issues and estimation uncertainties associated with each of the precipitation estimation products.

Among different remotely sensed precipitation datasets, PERSI-ANN-CDR is a viable candidate for extreme precipitation analysis given: (1) its high spatial and temporal resolution: when compared with the long-term Global Precipitation Climatology Project (GPCP) (Huffman et al., 1997) product which is monthly and 2.5° by 2.5°, PERSIANN-CDR has a higher temporal (daily) and spatial resolution (0.25° by 0.25°). The 2.5° spatial and monthly temporal resolution is not capable of capturing the spatial and temporal variability of the extreme precipitations especially over regions with complex topographic conditions, and (2) its long record: PERSIANN-CDR has relatively longer data record (34 + years and continually expanding) in comparison to TRMM 3b42 V7 (Huffman et al., 2007) with 20 + years of data, or CMORPH (Joyce et al., 2004) with 16 + years of record. Based on these strengths, Gado et al. (2017) employed the PERSIANN-CDR dataset to estimate extreme rainfall quantiles at two homogenous regions in the Western United States. They combined information from the PERSIANN-CDR pixels and nearby gauges in a homogenous region and used an innovative regional frequency analysis method to derive quantile estimates at ungauged locations.

The primary goal of this research is to evaluate the feasibility of using the PERSIANN-CDR dataset for rainfall frequency analysis by constructing the required DDF curves over regions with limited gauge information or mountainous areas. As a proof of concept, this study has been conducted over the United States, where longer gauge observations with sufficient spatial coverage exist. As some studies have reported, there are biases in the PERSIANN-CDR estimates which necessitate the application of bias-adjustment techniques to improve the accuracy of the PERSIANN-CDR estimates of extreme precipitations

(Miao et al., 2015; Duan et al., 2016; Shah and Mishra, 2016; Yang et al., 2016; Liu et al., 2017). This study was designed with the following objectives: (1) to propose an elevation-based bias correction model applicable to the PERSIANN-CDR dataset and to test it over a large number of river basins in the continental United States, and, (2) to demonstrate the usefulness of satellite-based precipitation data in rainfall frequency analysis and to use the derived frequency estimates to further verify the effectiveness of the proposed bias-correction model. In the proposed frequency analysis framework, only the PERS-IANN-CDR information is used to estimate extreme precipitation quantiles and no information from nearby gauges is incorporated in the development of DDF curves (Gado et al., 2017).

The rest of this paper is organized as follows: in Section 2, a detailed description of gauge and PERSIANN-CDR datasets used in the study is presented, followed by the specifications of the studied basins. The bias-adjustment approach, cross-validation techniques and the frequency analysis procedures pursued in the study are introduced in Section 3. Section 4 presents the results and discussion. The main findings and conclusions are summarized in Section 5.

2. Data

2.1. Gauge data

Global Historical Climatology Network (GHCN)-Daily is a quality controlled dataset that is used in this study. This dataset contains comprehensive information of daily summaries of more than 40 meteorological variables, including precipitation, temperature, snow depth, wind information, evaporation, etc. recorded by 100,000 land surface stations operated by 20 agencies around the world.

In this study, we select 20 basins located in the Eastern and Western United States (Fig. 1). The daily rainfall data from rain gauges with 34 + years of observation (1/1/1983–12/31/2015) were downloaded from National Oceanic and Atmospheric Association-National Climatic Data Center (NOAA-NCDC) database (https://www.ncdc.noaa.gov/ghcnd-data-access). A brief description of the selected basins with their hydrologic unit codes (HUC) and the number of gauges with 34 + years of data selected for this study are presented in Table 1. The selected basins incorporate a wide range of elevations, from 0 to 3700 m mean sea level, and diverse climatic conditions based on Köppen-Geiger climate classification system (Kottek et al., 2006).

2.2. PERSIANN-CDR data

PERSIANN-CDR is a retrospective multi-satellite precipitation dataset that provides near-global precipitation information, 60°N-60°S latitude and 0°–360° longitude, at 0.25° spatial resolution (around 25 km) and daily temporal resolution from 1 January 1983 to near present (Ashouri et al., 2015). The PERSIANN-CDR dataset was developed by the following steps. In the first step, the PERSIANN algorithm (Hsu et al., 1997) is implemented on the archive of Gridded Satellite (GridSat-B1) Infrared Data (Knapp et al., 2011) from Geostationary Earth Orbiting satellites (GEOs). The model is pre-trained using the National Center for Environmental Prediction (NCEP) Stage IV hourly precipitation data. Then the parameters of the model are kept fixed, and the model is run on the entire historical records of GridSat-B1 to estimate the historical precipitation at 3-hourly resolution. In the next step, the estimated rain-rates are resampled to 2.5° spatial resolution and bias-adjusted with GPCP product v2.2 (Adler et al., 2003) to keep it consistent with the GPCP monthly product. Finally, the PERSIANN-CDR dataset is obtained by accumulating the 3-hourly bias adjusted data. In this research, daily PERSIANN-CDR data for the selected basins for the time period of 1/1/1983 to 12/31/2015 was used.
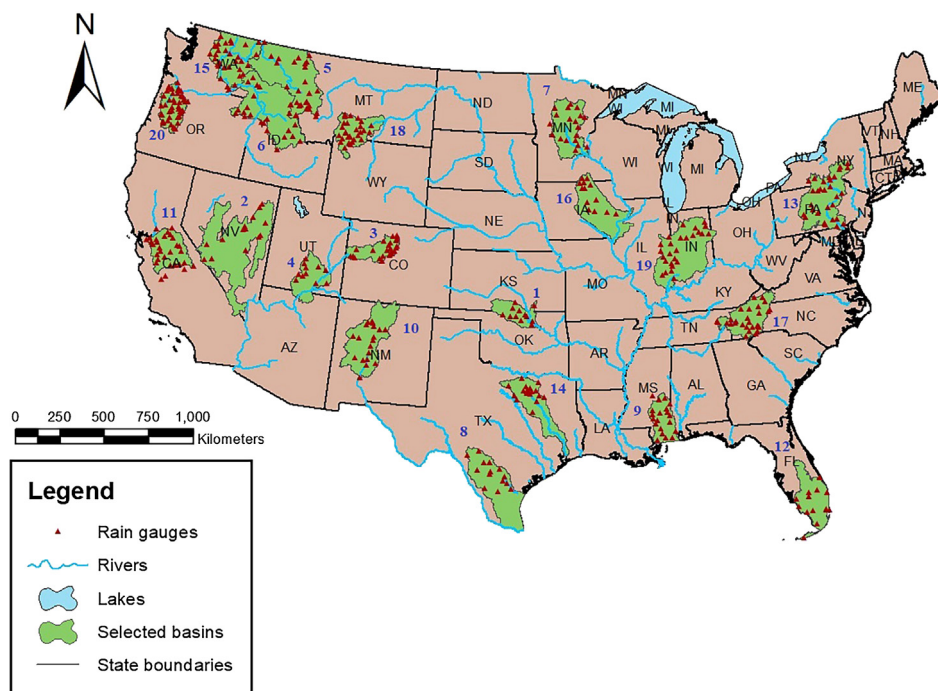
Fig. 1. The geographic location of the selected basins and gauges.

3. Methodology

3.1. Model description

In our proposed bias correction model, we first correct the PERSIANN-CDR estimates of annual maximum series with gauge data at pixels with available gauge records for the study period. The time series of annual maximum precipitation from both gauge network and PERSIANN-CDR for the corresponding pixels are extracted and sorted in an ascending order. For simplicity, we denote the gauge-based annual maximum series as "GM", and the PERSIANN-CDR annual maximum series as "PM" hereafter. A zero-intercept regression line is fitted to the scatterplot of GM and PM time series, with the corresponding PM values in the Y-axis and GM values in the X-axis (Fig. 2a). The slope of this regression line (called "Correction Factor" or CF hereafter) shows the deviation of PM with respect to ground truth (GM), and it indicates the level of correction required for correcting PM to GM. A CF value larger than one indicates an overestimation of the extreme precipitation by PERSIANN-CDR, and a CF smaller than one implies the underestimation. The larger the deviation of a CF value from the one to one case, the greater the correction required for the PM (Fig. 2a).

To investigate the orographic characteristics of bias at each basin, the CF values at individual gauges are plotted against the corresponding gauge elevations (Fig. 2b). The basin-scale plots are further merged to

Table 1
Summary of the features of the selected basins and gauges.

| Basin No.	Basin Name	HUC	State(s)	Area (sq. mi.)	No. of gauges	Gauge elevations (m)			Köppen-Geiger climate class	Climate Description
						Min	Max	Mean		
1	Arkansas-Keystone	1106	KS, OK	9750	12	253.0	454.2	347.7	Dfa, Cfa	Continental, temperate
2	Central Nevada Desert	1606	NV	47,100	14	1299.1	2448.5	1818.6	BSk, BWk, BWh, Dfb	Dry, continental
3	Colorado Headwaters	1401	CO	9730	25	1450.8	3474.7	2646.2	BSk, Dfb, Dfc	Dry, continental
4	Dirty Devil	1407	UT, AZ	13,500	11	1164.9	2987.0	2093.1	BSk, Bwk, Dfb	Dry, continental
5	Kootenai-Pend Oreille-Spokane	1701	MT, ID, WA	36,600	30	548.6	2514.6	1691.5	Dsb, Dfb	Continental
6	Lower Snake	1706	ID, WA, OR	11,800	23	328	2788.9	1523.3	Dfb, Dsb	Continental
7	Mississippi Headwaters	0701	MN	20,200	19	278.0	454.2	372.2	Dfb	Continental
8	Nueces-Southwestern Texas Coastal	1211	TX	29,000	15	43	625.4	217.8	Cfa	Temperate
9	Pascagoula	0317	MS	12,100	20	2.4	164.9	72.3	Cfa	Temperate
10	Rio Grande – Elephant Butte	1302	NM	26,900	19	1378.9	2621.3	1843.8	Dfb, BSk	Continental, dry
11	San Joaquin	1804	CA	15,600	30	3.7	2920	537.2	Csa, BSk	Temperate, dry
12	Southern Florida	0309	FL	18,700	14	0.9	10.7	4.3	Cfa, Aw	Temperate, tropical
13	Susquehanna	0205	PA, NY, MD	27,200	30	78.9	566.9	313.5	Dfa, Dfb	Continental
14	Trinity	1203	TX	18,000	24	103.9	339.9	208.7	Cfa	Temperate
15	Upper Columbia	1702	WA	22,600	32	113.1	1978.2	689.3	Dfb, Dsb, BSk	Continental, dry
16	Upper Mississippi-Iowa	0708	IA, IL, MN	22,800	25	195.1	390.1	319.0	Dfa, Dfb	Continental
17	Upper Tennessee	0601	TN, NC, GA, VA	17,200	28	230.1	1143.9	555.5	Cfa	Temperate
18	Upper Yellowstone	1007	MT, WY	14,400	30	944.3	2865.1	1906.2	Dfb, Dfc, BSk	Continental, dry
19	Wabash	0512	IN, IL, OH	32,600	26	134.7	284.1	202.2	Dfa	Continental
20	Willamette	1709	OR	11,400	46	6.4	1554.5	434.7	Csb, Dsb, Dsc	Temperate, continental
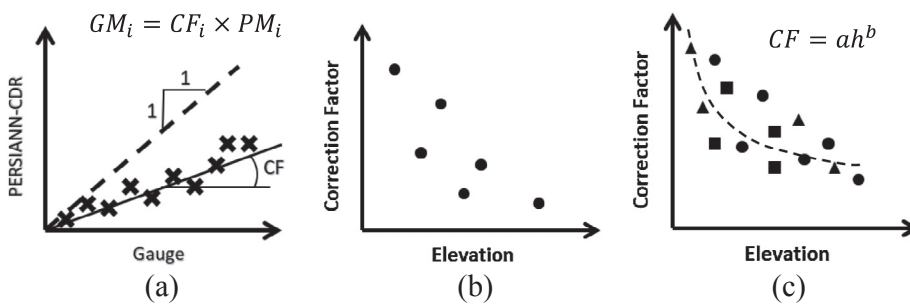
Fig. 2. Schematic view of the bias adjustment approach. (a) Estimation of correction factor at a single gauge with annual maximum series (AMS) of gauge and collocated PERSIANN-CDR pixel (b) Correction factor – elevation relationship at single basin scale (each point showing the CF for at a gauge) (c) Correction factor-elevation relationship for multiple basins and the schematic of the fitted exponential function (similar markers showing gauges from the same basin).

provide a more comprehensive view of the CF-elevation relationship (Fig. 2c). Following the approach mentioned above, an exponential function is fitted to the derived CF-elevation relationship at both individual basin and multi-basins scale as shown in Fig. 2c. We construct a correction function based on the CF-elevation relationship derived from 4 Western US basins and test its performance with different cross-validation and validation methods on other basins. The selected basins are San Joaquin River Basin (California), the Willamette River Basin (Oregon), the Upper Columbia River Basin (Washington) and the Colorado Headwaters (Colorado). These basins are selected since they provided bias-elevation information at different elevations and encompassed different climatic conditions, which are representative for building a robust and effective bias correction model applicable to other river bases in the United States. Finally, the correction model based on these four selected river basins is tested on the other 16 basins with different elevation ranges and climatic conditions in the Western and Eastern U. S.

3.2. Hold-out cross-validation

Hold-out cross-validation is implemented to examine how the performance of the correction model is influenced by the number of basins incorporated in the model calibration, and to investigate whether incorporating information from fewer basins could improve the PERSIANN-CDR estimates of AMS. The four basins used for training the correction function are divided into two groups. The basins are grouped in a way that information from different elevations and climates are included for each case. A correction function based on the gauge and the PERSIANN-CDR information from the basins in the first group is used to adjust the PM for the basins of the other group, and vice versa. In other words, an exponential regression function is fitted to the CF-elevation relationship from the two basins in the first group and is then used to adjust the AMS from the PERSIANN-CDR dataset for the basins in the second. The effectiveness of the bias-correction functions is assessed using the root mean squared error (RMSE) of the sorted AMS from the adjusted PERSIANN-CDR and that of gauge observations, at each of the gauge locations and basins.

3.3. Comparison with gauge interpolation

Besides comparing the original and corrected PERSIANN-CDR data using the approaches mentioned above, we also include a commonly used basin-scale interpolation method for analyzing extreme precipitation over remote and mountainous areas where the gauge network is insufficient or even non-existent (Chen et al., 2008; Doumounia et al., 2014).

3.3.1. Leave-one-out cross-validation

Precipitation intensity at an ungauged location is commonly estimated by interpolating observations from nearby gauges. Performance of the bias-adjusted PERSIANN-CDR dataset in estimating the annual maximum time series at an ungauged location is compared with the estimates from the interpolation method and the original

PERSIANN-CDR dataset. At each of the calibration basins, we leave one gauge out of the training phase, and the entire time series of precipitation at this particular gauge location is constructed with the linear interpolation of observations from the remaining gauges. Then, the annual maximum series at the location of the held out gauge is extracted from the interpolated time series. The CF-elevation relationship for the selected calibration basins is derived, and the CF corresponding to the elevation of the removed gauge is used to correct the PM time series at the PERSIANN-CDR pixel over the left-out gauge location. Finally, the interpolation-based annual maximum time series and the corrected PM are compared with the original GM. RMSE is used as the measure of the difference between the calculated time series and the GM. It is worth mentioning that we repeat this procedure for all the gauges at each calibration basin to investigate the robustness of our proposed correction method.

3.3.2. K-fold cross-validation

The leave-one-out cross-validation approach described in Section 3.3 evaluates the performance of the suggested bias-correction approach at a single gauge level. When there is a dense gauge network in a basin, interpolation of available gauge observations may result in better estimates of the AMS at an ungauged site. However, the gauge interpolated estimates could be less reliable when the region has limited or sparse gauge observations. Therefore, to find the breaking point where the corrected PERSIAN-CDR dataset starts to outperform the interpolation-based results, we carry out the k-fold cross-validation.

At each of the four basins used in the calibration process, different percentages (i.e., 10, 20, 30, 40, 50, 60, 70, and 80%) of gauges are randomly selected and left out. Then, the entire time series of precipitation for the locations of the removed gauges are constructed using the linear interpolation of the daily observations from the remaining gauges. The annual maximum series for the locations of the removed gauges are then extracted from the interpolated time series. Finally, we compare the corrected PM and the interpolation-based annual maximum time series at each gauge location with the GM for that location.

Since various combinations of gauges could be selected as test samples, results depend on the distribution of the remaining gauges and the distances between the held out and nearby gauges. To reduce the sensitivity of the results to the selection of gauges, we carry out 30 random selections of the hold-out gauges and consider each selection as an independent test. RMSE of the interpolation-based AMS is then compared with RMSE of the corrected PM for the selected gauges in each independent run. The average RMSE of the 30 independent runs is also calculated to have the overall error estimate for different hold-out scenarios (i.e. 10, 20, 30, 40, 50, 60, 70, and 80% of gauges being held out).

3.4. Satellite-based rainfall frequency analysis

National Oceanic and Atmospheric Administration (NOAA) Atlas 14 is a source of rainfall frequency estimates for the United States and its territories. NOAA Atlas 14 provides intensity-duration frequency (IDF) and depth-duration-frequency (DDF) curves for different regions based

on the regional frequency analysis approach (Bonnin et al., 2006). NOAA Atlas 14 IDF and DDF curves were developed using the best fit among different probability distributions, including the 3-parameter Generalized Extreme Value (GEV), the Generalized Normal, the Generalized Pareto, the Generalized logistic, the Pearson Type III distributions, the 4-parameter Kappa distribution; and the 5-parameter Wakeby distribution. At 80% of gauges and for sub-daily and daily durations, the GEV gave the best statistics among the 3-parameter distributions and its performance was comparable to that of 4 and 5 parameter distributions. Thus, the GEV was adopted across all gauges and durations (Bonnin et al., 2006).

The GEV distribution was firstly introduced by Jenkinson (1955), and it has been widely used for frequency analysis of extreme precipitation and was demonstrated superior over other probability distribution functions in terms of fitting the annual maxima time series (AMS) (Ben-Zvi, 2009; Bougadis and Adamowski, 2006; Fowler and Kilsby, 2003; Gellens, 2002; Norbiato et al., 2007; Villarini et al., 2011). The GEV distribution is a 3-parameter probability distribution that combines three extreme value distributions. The type of the distribution is characterized by the value of the shape parameter (ξ). Negative, zero, and positive values of the shape parameter determines the tail behavior of the distribution as short-tailed (Weibull), light-tailed (Gumbel) and heavy-tailed (Fréchet), respectively. The GEV cumulative distribution function is given by:

$$F(x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\} \text{for } \xi \neq 0 \tag{1}$$

$$F(x) = \exp\left\{-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \text{for } \xi = 0 \tag{2}$$

where ξ, μ, and σ are the shape, location, and scale parameters, respectively.

To fit the GEV distribution with the PERSIANN-CDR daily precipitation, we first adjust the data samples, in which the annual maximum series of PERSIANN-CDR for 2-day, 3-day, 4-day, 7-day, 10-day, 20-day, 30-day, 45-day and 60-day durations are corrected with gauge data using the same approach used for daily precipitation.

The GEV distribution is fitted to the annual maxima series of the corrected PERSIANN-CDR data for different durations using *gevfit* function from the Matlab Statistics and Machine Learning Toolbox (https://www.mathworks.com/help/stats/gevfit.html). Maximum likelihood estimation is used to estimate the parameters of the GEV distribution and the corresponding confidence intervals (Embrechts et al., 2013; Kotz and Nadarajah, 2000). The return level for each return period and duration is estimated using the inverse GEV function as in Eqs. (3) and (4):

$$X_T = \mu - \frac{\sigma}{\xi}\left\{1 - \left[-ln\left(1 - \frac{1}{T}\right)\right]^{-\xi}\right\} \text{for } \xi \neq 0 \tag{3}$$

$$X_T = \mu - \sigma \, ln\left[-ln\left(1 - \frac{1}{T}\right)\right] \text{for } \xi = 0 \tag{4}$$

where X_T is the return level (i.e., the rainfall depth that on average is exceeded once in T years), and $T = 1/(1 - F)$ is the return period. Using the return levels at different return periods and annual exceedance probabilities, the DDF curves are generated.

NOAA Atlas 14 (Bonnin et al., 2006) provides DDF curves with sub-daily, daily and multi-day durations. The DDF curves can be downloaded from the NOAA precipitation frequency data server (https://hdsc.nws.noaa.gov/hdsc/pfds/). Since the PERSIANN-CDR dataset gives precipitation estimates at daily time scale, the daily and multi-day durations were considered in generating the DDF curves. To remain consistent with NOAA frequency estimates, precipitation durations considered in this study are 1-day, 2-day, 3-day, 4-day, 7-day, 10-day, 20-day, 30-day, 45-day, and 60-day. It should be noted that the

durations considered here do not mean precipitation occurred during the entire period, but the sliding window gives the highest value of precipitation accumulation over the selected period. Lastly, we compare the return levels based on the corrected PERSIANN-CDR estimates with that of NOAA Atlas 14 at each duration and return period.

3.5. Uncertainty assessment

The confidence intervals of the return levels from the original and the adjusted PERSIANN-CDR datasets are estimated using a boot-strapping technique. We generate 1000 random samples with replacements from the original and adjust AMS at the target gauge locations. Then, the Maximum Likelihood estimation is used to calculate the parameters of the GEV distributions fitted to each of these random samples. Return levels for different durations are calculated using the inverse GEV function evaluated at different return periods. Finally, the 5th and 95th percentiles of the bootstrapped return levels at each duration and return period are taken as the 90 percent confidence intervals.

4. Results and discussion

4.1. Training basins

Fig. 3 shows the scatterplots of the PERSIANN-CDR and gauge AMS and the regression line equation at a number of gauges in the Willamette River Basin in the state of Oregon. In some of the gauge locations (Fig. 3a–d, and g–i), the original PERSIANN-CDR has a certain degree of underestimation or overestimation, while in some other gauge locations the PERSIANN-CDR estimates are in good agreement with gauge observations (Fig. 3e and f).

An important note here is that a PERSIANN-CDR pixel has an area about 625 km^2 which is much larger than the sampling area of a rain gauge. The value of a PERSIANN-CDR pixel represents the average precipitation within that pixel's spatial domain. In fact, even if the PERSIANN-CDR estimate at a pixel is completely accurate, its value tends to be smaller than the subpixel point measurements. In other words, by comparing a PERSIANN-CDR pixel with a point measurement, we are carrying out a "point-area" comparison. Therefore, by adjusting the PERSIANN-CDR pixels with point measurements, we are downscaling PERSIANN-CDR to point resolution. This implies that the adjusted dataset should be regarded as a point estimate, rather than an area estimate. Furthermore, there would be time discrepancies between the PERSIANN-CDR's daily interval, and the gauges' 24-hour intervals. Thus, the correction factor accounts for the influence of both "point-area" and time discrepancy issues.

At each of the four basins selected to build the correction model (i.e., the San Joaquin River Basin, the Willamette River Basin, the Upper Columbia River Basin, and the Colorado Headwaters), the CF at the basin scale is computed by fitting a zero-intercept regression line to the sorted AMS of all the available gauges and that of collocated PERSIANN-CDR pixels(Fig. 4). In general, the PERSIANN-CDR estimates of AMS tend to be lower than the AMS from gauge observation, with different levels of underestimation in the different basins. The AMS estimates from the original PERSIANN-CDR dataset show considerable underestimation at Colorado Headwaters and Upper Columbia River Basins with CFs equal to 0.46 and 0.57, respectively.

The scatterplots of CF and gauge elevation for different basins are shown in Fig. 5(a–d). In general, an exponential relationship exists between CFs and elevations at each of the four basins. Furthermore, when merging all the available gauge information from the selected river basins together, a comprehensive view of this relationship is demonstrated (Fig. 5e). As we can see from Fig. 5e, the CFs become smaller with increasing elevation of the gauges. This reduction in the CFs implies the underestimation of AMS at higher elevations.

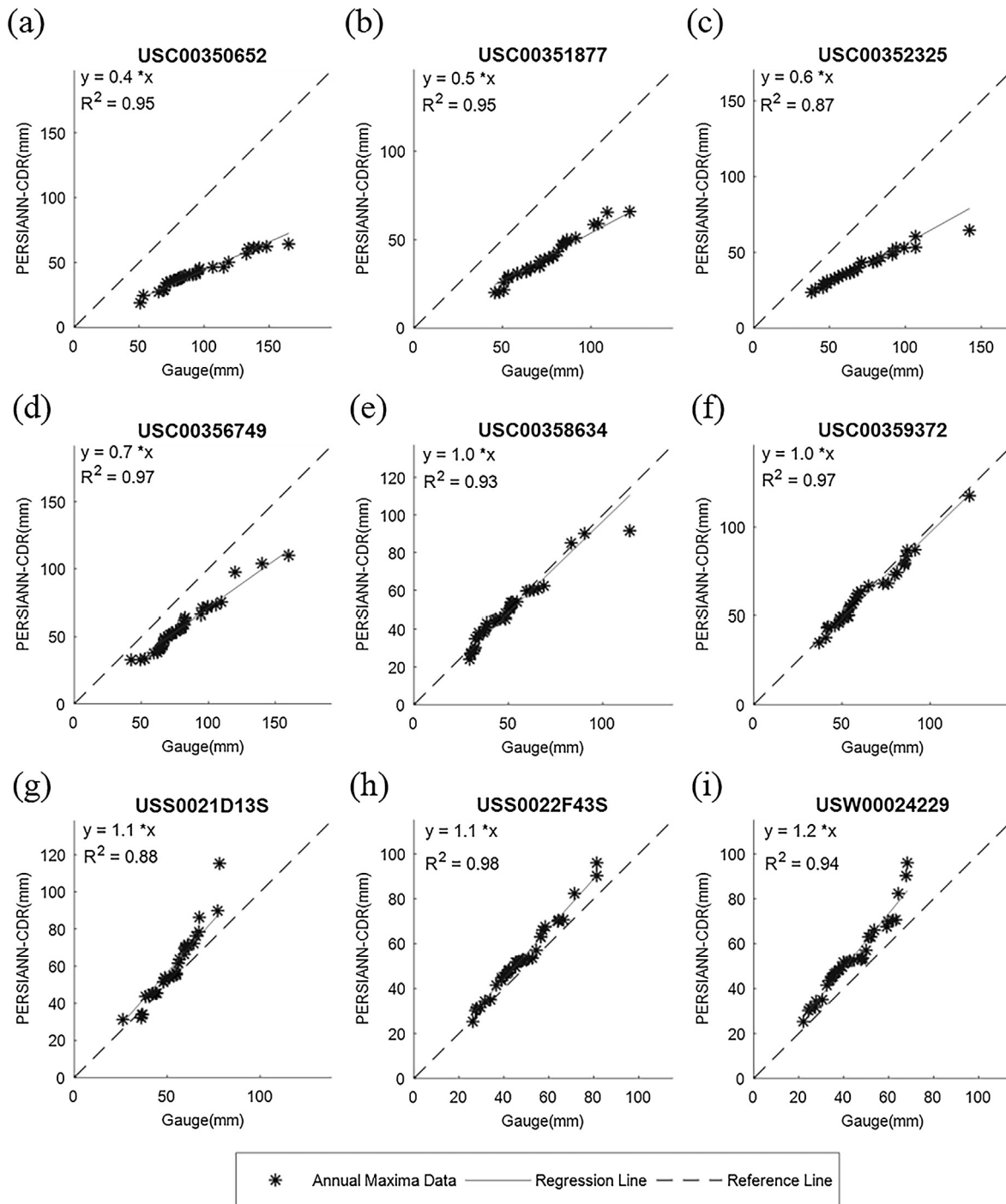Both IR-based (such as PERSIANN family) and Passive Microwave

Fig. 3. Scatterplots of gauge and PERSIANN-CDR AMS and the zero intercept regression lines at sample gauge locations in Willamette River Basin. Station names are written above the scatterplots.

based (such as TMPA (Huffman et al., 2007)) precipitation products have been reported to underestimate precipitation in high elevations (Hashemi et al., 2017). This underestimation has been related to several factors. Satellite-based precipitation products have difficulties in retrieving the solid form of precipitation (snow), which is the prevailing type of precipitation at high elevation regions and in the winter season (Hashemi et al., 2017). Moreover, since IR-based precipitation algorithms rely on the cloud top temperatures, they cannot fully detect the orographic enhancements in the liquid phase of precipitation in regions characterized by complex topographic conditions (Shige et al., 2013). In addition to the technical and methodological issues inherent to the satellite precipitation estimation methods, the spatial and temporal

inconsistencies between the satellite precipitation estimates and gauge observations at high elevation regions can be related to the poor sampling of gauges (Gebregiorgis and Hossain, 2014). For instance, Libertino et al. (2016) observed the lowest agreement in the timing of extreme events recorded by TRMM and gauge observations in sparsely gauged regions. Miao et al. (2015) also reported low spatial and temporal agreement in terms of extreme precipitation statistics between the PERSIANN-CDR estimates and gauge observations in regions with low density of gauges.

As shown in Fig. 5, we fit an exponential function to the scatterplots of CF for each gauge and its corresponding elevation. This function is used to correct the PERSIANN-CDR estimates of AMS at different basins
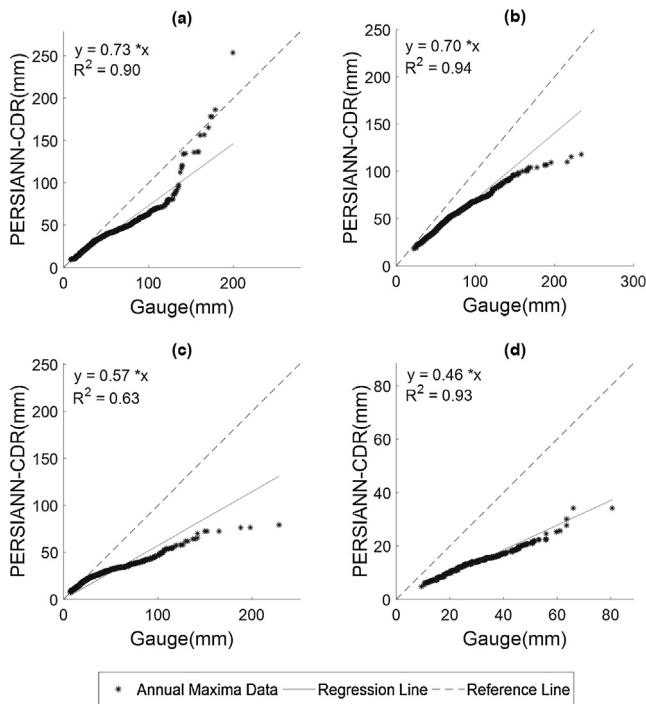
Fig. 4. CF at basin scale for the selected river basins in the Western United States, including (a) San Joaquin River Basin, (b) Willamette River Basin, (c) Upper Columbia River Basin, and (d) Colorado Headwaters River Basin.

in the Eastern and Western United States. However, since this correction model is based on only a few selected basins, it is necessary to be validated using different cross-validation techniques and then be tested on different basins over the continental United States.

4.2. Hold-out cross-validation results

We first carry out hold-out cross-validation on the four selected

river basins, in which a correction function based on CF-elevation relationship is built using the information from two of the river basins, Willamette and Upper Columbia river basins. The model is tested on the other two river basins (the San Joaquin and the Colorado Headwater River Basins), and vice versa. The goal is to examine the effects of limited gauge information and basin selection on the overall performance of the bias-correction approach.

The effect of bias-correction on the empirical CDF of the PERSIANN-CDR estimates at each of the calibration basins is shown in Fig. 6. At basin scale, the correction method shifts the empirical CDF of the AMS from the original PERSIANN-CDR towards the gauge-based empirical CDF. In the Willamette River Basin and the Colorado headwaters River Basin, the corrected CDF is close to that of the observation. In the San Joaquin River Basin, the extreme quantiles from the corrected data are closer to the observation. In the Upper Columbia River basin, the corrected PERSIANN-CDR gives better estimates of the largest extreme values compared to the original PERSIANN-CDR estimates. However, it results in an overestimation of the lower quantiles. This is consistent with the results shown in Fig. 4, where the regression-based estimates gave some overestimation for values lower than 55 mm.

The statistics of the hold-out cross-validation results at gauge scale are presented in Table 2. In most of the gauge locations (111 out of 127 gauges in different basins), the RMSE of the corrected PERSIANN-CDR is lower than that of the original PERSIANN-CDR. This implies the effectiveness of the proposed bias-adjustment approach in correcting the PM at pixel level even in basins with dense gauge networks. At 16 gauges, however, the correction method tends to deteriorate the original PERSIANN-CDR estimates. Among these gauges, 12 are located in low elevation regions (< 550 m from mean sea level), and more than half of them are associated with elevations less than 200 m from mean sea level. The Upper Columbia and the Willamette River Basins have a larger portion of these gauges with 7 and 5 unsuccessful corrections, respectively. The poor performance of the corrected PERSIANN-CDR at those gauge locations could be partly attributed to the complex topographic conditions of those basins, which pose some challenges for the PERSIANN algorithm to estimate precipitation accurately. However, as compared to the original PERSIANN-CDR, the proposed correction
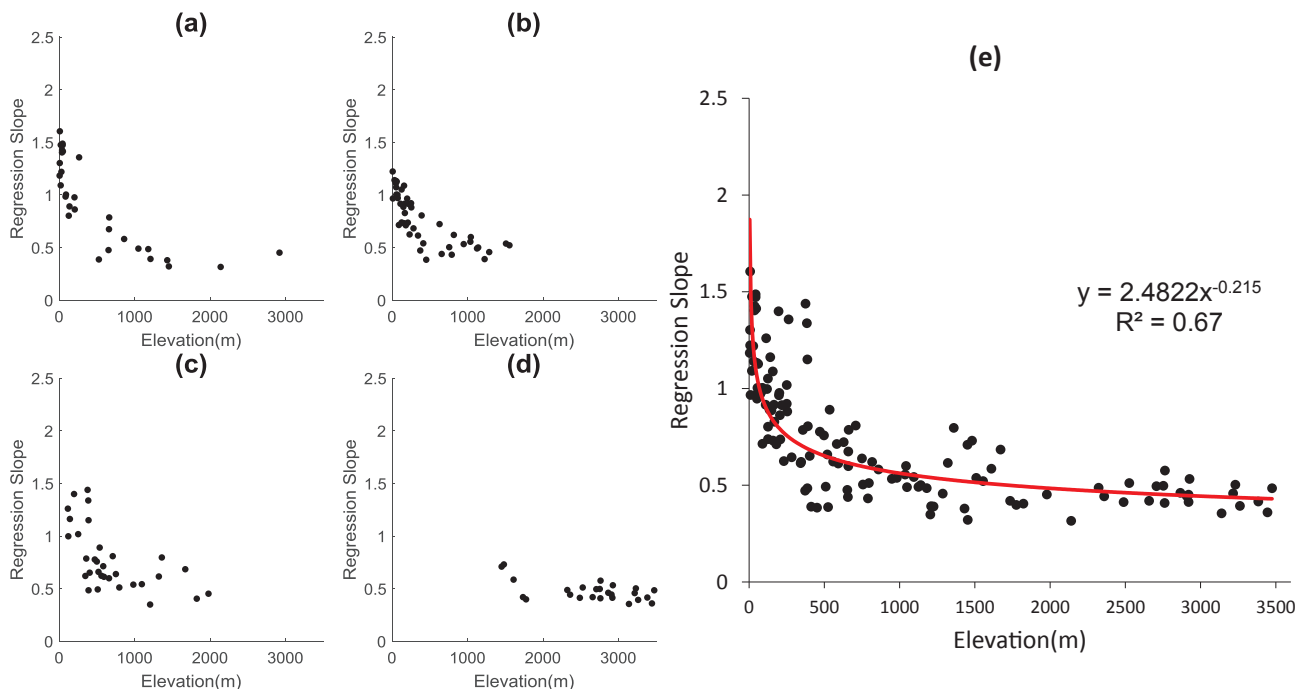

Fig. 5. CF and elevation relationship at basin and multi-basin scale for (a) San Joaquin River Basin, (b) Willamette River Basin, (c) Upper Columbia River Basin, (d) Colorado Headwaters River Basin, and (e) merging all gauge information for all the basins.
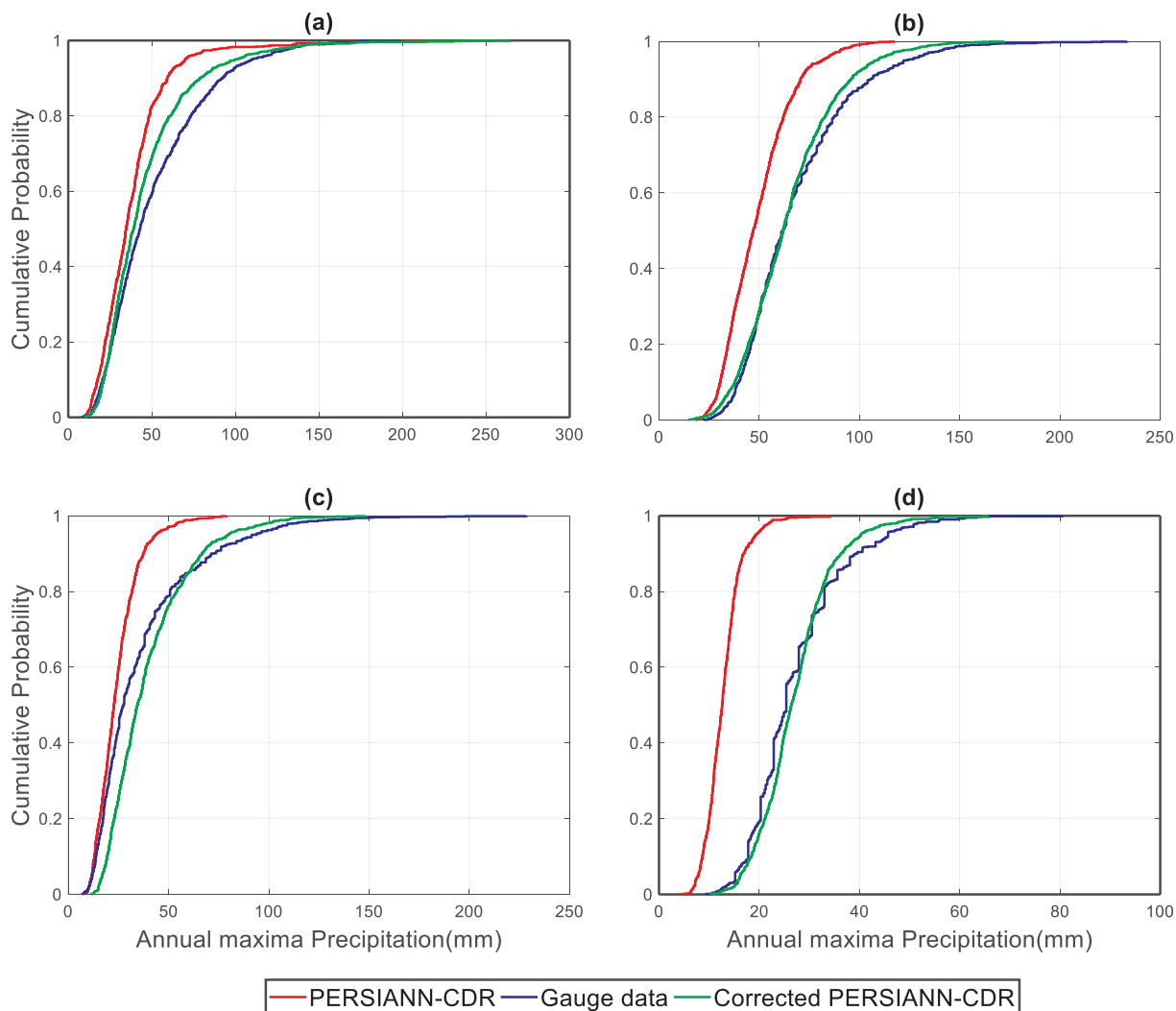
Fig. 6. Empirical CDF of the AMS from the gauge observations, the original and the corrected PERSIANN-CDR data at basin scale. (a) San Joaquin River Basin, (b) Willamette River Basin, (c) Upper Columbia River Basin, and (d) Colorado Headwaters River Basin.

approach works well for the majority of the gauges in the hold-out cross-validation as shown with the lower RMSE values in Table 2.

4.3. Leave-one-out cross-validation

In each basin, one gauge is left out at a time, and the time series of precipitation at that gauge location is constructed using linear interpolation. The annual maximum time series from the gauge interpolation and the corrected PERSIANN-CDR are compared with the gauge observations at the corresponding location (Table 3). As shown in Table 3, the RMSE values from the corrected PERSIANN-CDR are consistently lower than those of the original PERSIANN-CDR for all basins. When compared to the interpolation method, the corrected PERSIANN-CDR gives lower RMSE values at the San Joaquin, the Willamette, and the Colorado Headwaters River Basins. At the Upper Columbia River Basin, however, the leave-one-out cross-validation results suggest that the gauge interpolation performs better than the corrected PERSIANN-CDR data.

It is inferable from the gauge scale results that the correction model outperforms the interpolation method in most cases, even if only one of the gauges at a densely gauged basin is removed from the sample. The interpolation method also produces substantial errors at some gauge locations, particularly those locations where the interpolated gauge is relatively far from its surrounding gauges. It is possible that complex

topography leads to different precipitation characteristics between nearby gauges and results in uncertainties in the interpolated precipitation estimates.

4.4. K-fold cross-validation

The leave-one-out cross-validation results in the previous section demonstrate that the interpolation-based estimates of AMS achieved by removing one of the gauges may outperform the PERSIANN-CDR estimates at some gauges in a densely-gauged region (e.g., the Upper Columbia River Basin). In order to find the breakpoint where the corrected PERSIANN-CDR will outperform interpolation-based estimates at a basin scale, the k-fold cross validation is implemented. We randomly separate different fractions of all available gauges (0.1, 0.2…, 0.8 of the gauges) in a basin, and remove the selected gauges from the model training phase. This random selection and removal process is repeated 30 times for each fraction level. Then, the entire precipitation time series at those locations are constructed by the linear interpolation of observations from the remaining gauges in that basin. Then RMSE of AMS estimates from both corrected PERSIANN-CDR and gauge interpolation are computed at the removed gauges.

Fig. 7 shows the average RMSE values of AMS estimates from the interpolation method (blue line) and the corrected PERSIANN-CDR (red line) for different exclusion ratios in each river basin. The horizontal

Table 2
The RMSE(mm) of the original and corrected PERSIANN-CDR data during hold-out cross-validation at different gauge locations and basins.

San Joaquin		Willamette		Upper Columbia		Colorado Headwaters	
Corrected	Original	Corrected	Original	Corrected	Original	Corrected	Original
14.76	16.74	27.60	54.73	13.30	19.68	6.16	18.20
19.13	34.44	5.78	5.78	3.76	3.09	4.71	18.73
32.66	47.61	11.26	24.65	4.12	8.66	5.44	18.29
31.41	69.29	15.46	11.15	1.32	8.61	5.25	10.88
12.96	**4.26**	16.73	19.15	2.93	8.02	7.02	9.76
43.12	71.64	30.79	50.90	7.38	31.30	6.43	19.97
9.42	8.59	6.51	9.53	3.61	**1.91**	8.98	**6.50**
10.03	10.10	7.42	20.92	15.25	20.72	6.21	17.25
19.79	25.71	11.85	18.53	1.69	4.71	4.12	12.42
6.11	5.49	10.43	11.36	4.97	8.45	4.80	16.22
19.16	36.62	4.99	**4.21**	6.20	10.88	5.09	12.83
30.47	59.50	6.65	11.81	0.96	4.25	3.74	18.69
8.86	13.73	20.38	37.86	9.12	21.90	4.74	12.75
23.12	**10.42**	9.43	**7.62**	3.17	4.97	4.03	19.37
6.57	6.29	14.50	21.04	11.72	**6.27**	2.61	18.97
16.38	19.47	9.79	19.52	7.36	15.71	5.66	23.01
30.53	20.49	13.75	**8.71**	0.92	4.30	1.98	15.97
3.25	9.45	13.34	40.60	5.08	9.13	4.12	13.03
22.05	50.93	45.06	65.65	26.64	34.04	5.98	14.10
8.12	26.77	11.12	11.52	8.59	9.39	4.32	13.81
13.91	18.23	2.94	4.90	3.58	10.14	2.73	15.47
13.69	23.67	2.36	6.02	8.59	**4.84**	3.01	15.38
10.34	15.72	17.35	7.06	19.52	**8.05**	6.20	12.17
17.74	36.22	4.41	8.65	7.33	40.04	7.29	23.12
24.67	47.21	24.83	44.44	14.55	53.51	9.89	**5.98**
19.90	31.57	11.39	**4.63**	9.88	19.40	–	–
12.03	27.90	9.58	27.84	8.47	28.97	–	–
16.77	20.93	3.29	3.55	19.66	**14.44**	–	–
–	–	6.39	33.06	43.31	77.52	–	–
–	–	7.81	27.53	13.28	**7.61**	–	–
–	–	25.35	54.10	17.14	**10.74**	–	–
–	–	8.42	39.02	–	–	–	–
–	–	9.66	44.36	–	–	–	–
–	–	29.91	66.88	–	–	–	–
–	–	11.66	37.27	–	–	–	–
–	–	13.56	48.13	–	–	–	–
–	–	7.22	24.98	–	–	–	–
–	–	6.97	34.96	–	–	–	–
–	–	13.64	56.23	–	–	–	–
–	–	16.40	24.77	–	–	–	–
–	–	4.53	6.53	–	–	–	–
–	–	12.97	**10.91**	–	–	–	–
–	–	5.74	7.32	–	–	–	–

Note: The underlined values show the cases where the RMSE of the original PERSIANN-CDR is lower than that of the corrected PERSIANN-CDR.

Table 3
The RMSE (mm) from leave one out cross validation using gauge interpolation, original and corrected PERSIANN-CDR data.

Method	Basin Name			
	San Joaquin	Willamette	Upper Columbia	Colorado Headwaters
Interpolation	21.61	16.06	7.87	6.84
Corrected PERSIANN-CDR	17.34	12.77	**9.79**	5.22
Original PERSIANN-CDR	26.45	23.95	16.22	15.31

Note: The underlined value shows the case where the corrected PERSIANN-CDR method gave a higher RMSE value at basin scale than the interpolation method.

axis defines the number of the iteration, and the vertical axis presents the average RMSE value of the AMS estimates on the excluded gauge locations using the corrected PERSIANN-CDR and gauge interpolation. As the fraction of gauges being removed from the entire samples increases, the errors associated with the interpolation method become larger. In contrast, the errors produced with our proposed correction method remain consistently low for different basins over most of the test scenarios (i.e., different percentages of the gauge being removed). Moreover, the interpolation-based estimates result in large errors in some test scenarios and basins. For example, in the San Joaquin River Basin (Fig. 7a), substantial errors are observed in different scenarios and over several independent runs.

There are several reasons why errors from the interpolation method have large values in some of the iterations. Extreme precipitation events vary substantially in space and time. The annual maximum precipitation at different points of a basin could be results of various extreme events occurring in different times of the year. When interpolating the daily gauge observations in a region for constructing precipitation time series at an ungauged site, heavy precipitation observed at one or more gauge locations could be falsely extended to the locations that were less impacted by the storm. Similarly, by removing some of the gauges from the population, the extreme events impacting those locations may not be represented in the interpolated time series from the remaining sample and as a result the extreme event at that location would be missed. Both of these cases may result in considerable errors in the annual maximum series estimated from the interpolation method. Other factors that contribute to the significant interpolation errors include long distance of sample gauges from the target locations, substantial elevation differences between the target locations and sample
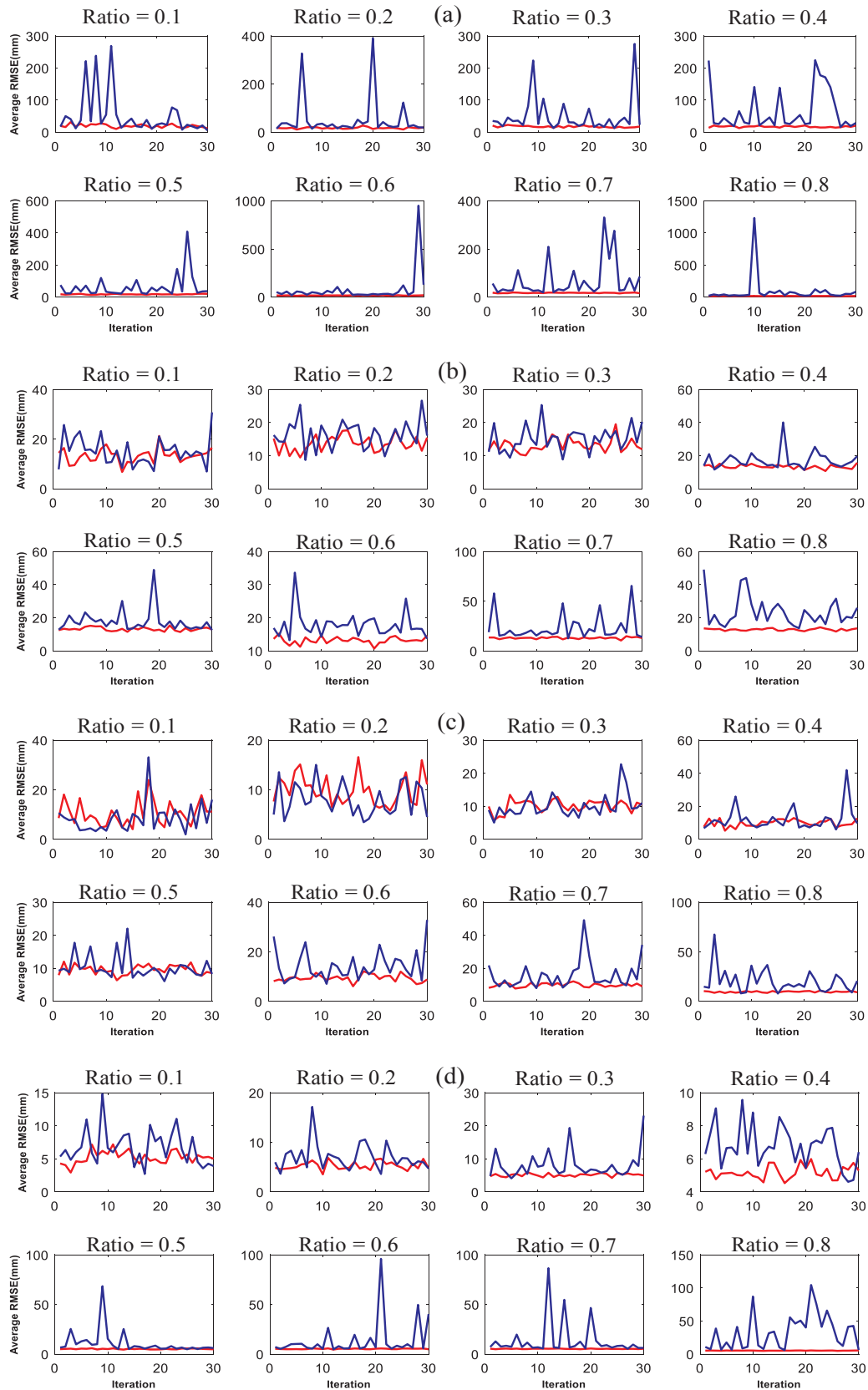
Fig. 7. Average RMSE of AMS estimates from the corrected PERSIANN-CDR (red lines) and the interpolation method, during the k-fold cross-validation (blue lines) for different exclusion ratios at: (a) the San Joaquin River Basin, (b) the Willamette River Basin, (c) the Upper Columbia River Basin, and (d) the Colorado Headwaters River Basin. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
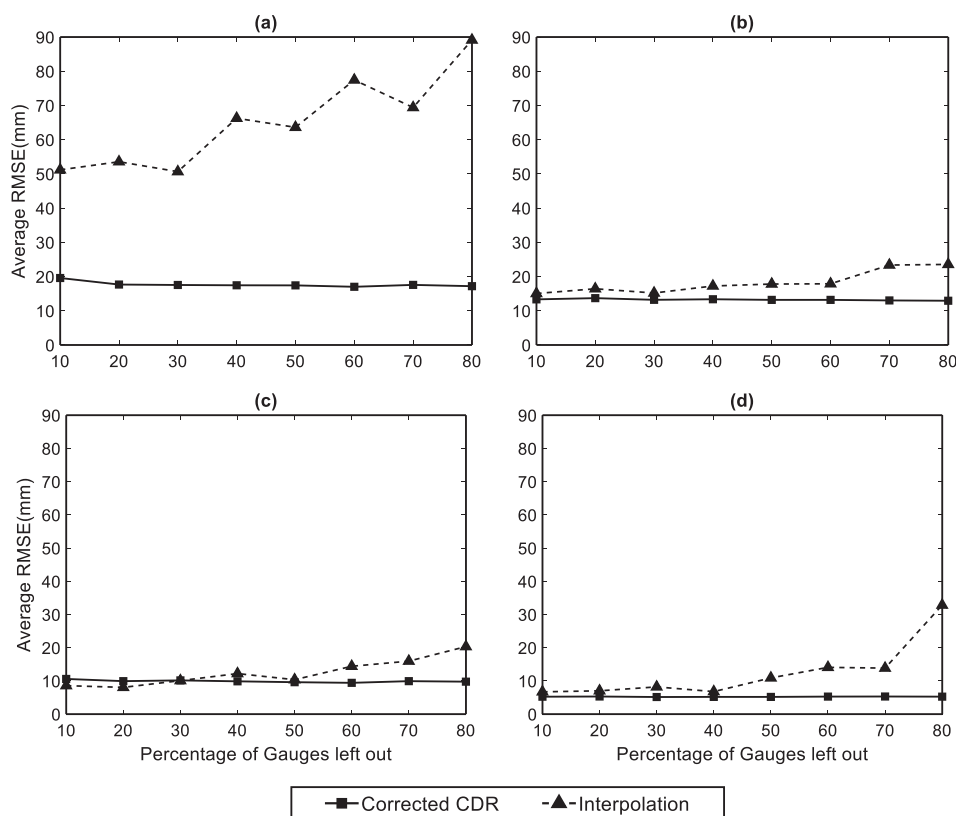
Fig. 8. Overall RMSE of AMS estimates from the corrected PERSIANN-CDR (solid line) and the interpolation method(dashed line) at different basins, including (a) the San Joaquin River Basin, (b) the Willamette River Basin, (c) the Upper Columbia River Basin, and (d) the Colorado Headwaters River Basin.

gauges, and the inability of the sample gauges to demonstrate the spatiotemporal variability of rainfall at target locations.

The overall errors from the corrected PERSIANN-CDR and the interpolation method at different exclusion ratios and basins are shown in Fig. 8. As the portion of gauges being left out increases, the RMSE produced by the interpolation method increases for all the basins, while the proposed correction method for PERSIANN-CDR shows stable errors over different ratios and basins with respect to the AMS results. In the San Joaquin River Basin (Fig. 8a), the Willamette River Basin (Fig. 8b), and the Colorado Headwaters River Basin (Fig. 8d), the corrected PERSIANN-CDR yields lower RMSE values than the gauge interpolation method throughout different ratios, suggesting the effectiveness of the proposed correction approach. In the Upper Columbia River Basin, the gauge interpolation method results in better estimates of the AMS at the ratios up to 30%. However, beyond the 30% threshold, the corrected PERSIANN-CDR produces more accurate estimates of the AMS. Therefore, 30% of total gauges is the breakpoint for the Upper Columbia River Basin in the context of interpolating point gauge information to spatial estimates. By comparing the statistics of corrected PERSIANN-CDR and the traditional interpolation method, it is observed that the proposed correction model generates more accurate estimates of the AMS than does the linear interpolation method. The superiority of the proposed bias-correction method becomes increasingly evident as the gauges become sparser.

4.5. Validation on the continental U.S.

In the previous sections, we demonstrated the effectiveness and robustness of the proposed correction model on the four representative river basins in the western U.S. In this section, we extensively validate the correction model on 16 additional river basins with different climates and topographic conditions across the continental United States (Table 1). The selected basins for validation cover all the climate classes

available in the United States based on the Köppen-Geiger climate classification system. In addition, these basins cover a broad range of elevations, from low-lying regions in the state of Florida to high elevation regions in the state of Utah. These basins are also associated with various dominant precipitation mechanisms (such as convective, orographic, and cyclonic) which could influence the performance of the satellite-based precipitation products (Hong et al., 2007; Liu and Zipser, 2009).

Table 4 presents the errors in AMS estimates from the original and the corrected PERSIANN-CDR data on the tested river basins. According to Table 4, in 15 out of the 16 basins, the correction model results in lower RMSE values compared to the original PERSIANN-CDR data. Significant improvements are observed at high elevation regions such as Dirty Devil, Rio Grande, and Upper Yellowstone river basins with 78.7%, 72.3%, 71.6% reduction in the RMSE of AMS, respectively. Also, the correction model considerably decreases the errors in the AMS estimates at mid-elevation regions, such as Mississippi headwaters, Upper Mississippi-Iowa, and Upper Tennessee River basins (Table 4). Among the low elevation regions, Nueces-Southwestern Texas Coastal and Trinity River basins were quite successful with respect to the error reduction by the correction model. However, Pascagoula River Basin is less successful (7.3% decrease in RMSE) and South Florida River Basin fails to improve (29.7% increase in RMSE). Both of these basins are located in the South Atlantic Gulf region, which is characterized by warm convective precipitation mechanisms. As a result of these convective systems, satellite precipitation products often fail to provide accurate estimates at these regions, as we see from the performance of raw data shown in Table 4. Both the Pascagoula and the South Florida river basins have high initial errors compared to the other river basins. Since the model is trained using the information from four river basins in the western United States with different hydroclimatic conditions, it is reasonable for it to not perform as well under temperate and tropical climatic conditions and for extremes caused by warm convective

Table 4
Bias-correction results in selected basins over the United States.

Basin No.	Basin Name	Average RMSE (mm)		Change (%)
		Original PERSIANN-CDR	Corrected PERSIANN-CDR	
1	Arkansas-Keystone	30.90	12.02	−61.1
2	Central Nevada Desert	14.16	6.91	−51.2
4	Dirty Devil	15.35	3.27	−78.7
5	Kootenai-Pend Oreille-Spokane	24.23	9.67	−60.1
6	Lower Snake	17.36	4.94	−71.6
7	Mississippi Headwaters	19.12	3.42	−82.1
8	Nueces-Southwestern Texas Coastal	28.68	17.43	−39.2
9	Pascagoula	39.20	36.32	−7.3
10	Rio Grande – Elephant Butte	15.67	4.34	−72.3
12	Southern Florida	35.40	45.92	+__29.7__
13	Susquehanna	22.61	10.64	−52.9
14	Trinity	35.62	21.87	−38.6
16	Upper Mississippi-Iowa	23.83	7.24	−69.6
17	Upper Tennessee	27.14	10.15	−62.6
18	Upper Yellowstone	18.68	5.31	−71.6
19	Wabash	23.36	11.46	−51.0

Note: The underlined value shows the case where the adjustment of PERSIANN-CDR data failed to improve the errors in the annual maximum series estimates.

systems.

Generally, satellite precipitation estimation algorithms perform poorly in estimating precipitation from shallow and warm convective clouds (Hong et al., 2007; Kubota et al., 2009; Liu and Zipser, 2009; Sorooshian et al., 2002). One reason behind this poor performance is that these algorithms relate heavy precipitations to deep convective clouds and subsequently underestimate heavy precipitations associated with shallow warm clouds (Hong et al., 2007; Liu and Zipser, 2009). Moreover, IR-based methods such as PERSIANN are based on cloud top temperature thresholds that are sometimes too cold for warm orographic clouds (Adler et al., 2003; Dinku et al., 2008). Finally, due to the contamination by the cold anvil cirrus clouds, IR-based precipitation estimates typically display 1-3hr phase shift compared to the maximum diurnal precipitation. These phase shifts influence the performance of IR-based methods in regions dominated by warm convective clouds (Sorooshian et al., 2002).

4.6. Multiday annual maximum series

In Fig. 9, we present the scatterplots of the CF-elevation for multi-day duration AMS at Colorado Headwaters (Fig. 9a–j) as an illustrative example. Fig. 9(a)–(j) suggest there is a similar CF-elevation behavior in multi-day AMS analysis for different durations. Fig. 9(k) presents the exponential regression functions fitted to each of the N-day maximum scatterplots. According to Fig. 9(k), as durations increase from 1 day to 60 days, the original PERSIANN-CDR estimates of the AMS become more accurate (i.e., closer to the CF = 1 line). This is because the PERSIANN-CDR dataset is bias-adjusted with GPCP dataset (Huffman et al., 1997) at a monthly scale and the values from the two datasets become closer to each other at longer durations. Therefore, at 30-days or 60-days analysis, the AMS estimates should be close to gauge observation. Although PERSIANN-CDR and gauge information are adjusted at a monthly scale, the monthly coefficients are applied to daily estimates (Ashouri et al., 2015). Therefore, the sub-monthly or daily estimates may not be compatible with gauge observations at the corresponding scale. Furthermore, the GPCP is a gauge interpolated dataset which its pixel values are essentially the average values of gauge observations within the large grid boundary. However, here we compare the PERSIANN-CDR estimates in pixel scale with the collocated gauge values which could differ substantially from the corresponding GPCP pixel values.

4.7. Depth-duration-frequency curves

Fig. 10 shows the DDF curves derived from the adjusted PERSIANN-CDR data, the frequency estimates from NOAA Atlas 14, and the 90% confidence intervals for a gauge location in Dirty Devil basin in the state of Utah (USC00420849). We present return levels for daily and multi-day durations given the daily resolution of the PERSIANN-CDR dataset. As shown in Fig. 10, the frequency estimates from the original PERSIANN-CDR data are outside the 90% confidence intervals of the NOAA Atlas 14 which suggests the necessity of bias adjustment prior to employing the data for frequency analysis. On the other hand, DDF curves from the adjusted PERSIANN-CDR data are well within the 90% confidence intervals of the NOAA Atlas 14 DDF curves. In most cases, the frequency estimates from the adjusted PERSIANN-CDR data are very close to the NOAA Atlas 14 estimates which are calculated by incorporating a large number of gauges and longer records of data for frequency estimation. Larger deviations from the gauge-based estimates are observed at longer return periods, and there is no clear trend in terms of overestimation or underestimation with respect to duration.

As shown in Fig. 10, the confidence intervals from the gauge-based and satellite-based DDF etimates become larger as the return periods increase. This higher uncertainty is because of the lower sample size at the tails of the distributions. Furthermore, the confidence intervals from the original and the adjusted PERSIANN-CDR datasets are relatively comparable given the similar lengths of the two datasets. However, the uncertainty bounds from the satellite-based DDF estimates are larger than those from the NOAA Atlas 14. One reason behind these larger confidence intervals is the shortness of the PERSIANN-CDR dataset when compared to the gauge information used for the development of NOAA Atlas 14 DDF curves. Another reason is the difference between the frequency analysis method implemented here and the method employed in the development of NOAA Atlas 14. NOAA uses the regional frequency analysis based on L-moments to estimate the frequency and intensity of extremes (Hosking and Wallis, 2005). The regional frequency analysis method is used in Atlas 14 in order to relieve the uncertainties arising from a low sample size (limited years of observations) during the GEV parameter estimation process. Although the regional frequency analysis method gives frequency estimates with lower uncertainties, it comes with the assumption of regional homogeneity in extreme rainfall characteristics which is not always a valid assumption. Here, the frequency analysis methods and uncertainties of the frequency estimates are outside of the scope of this study, and the DDF curves and their error analyses are the proof of concept.

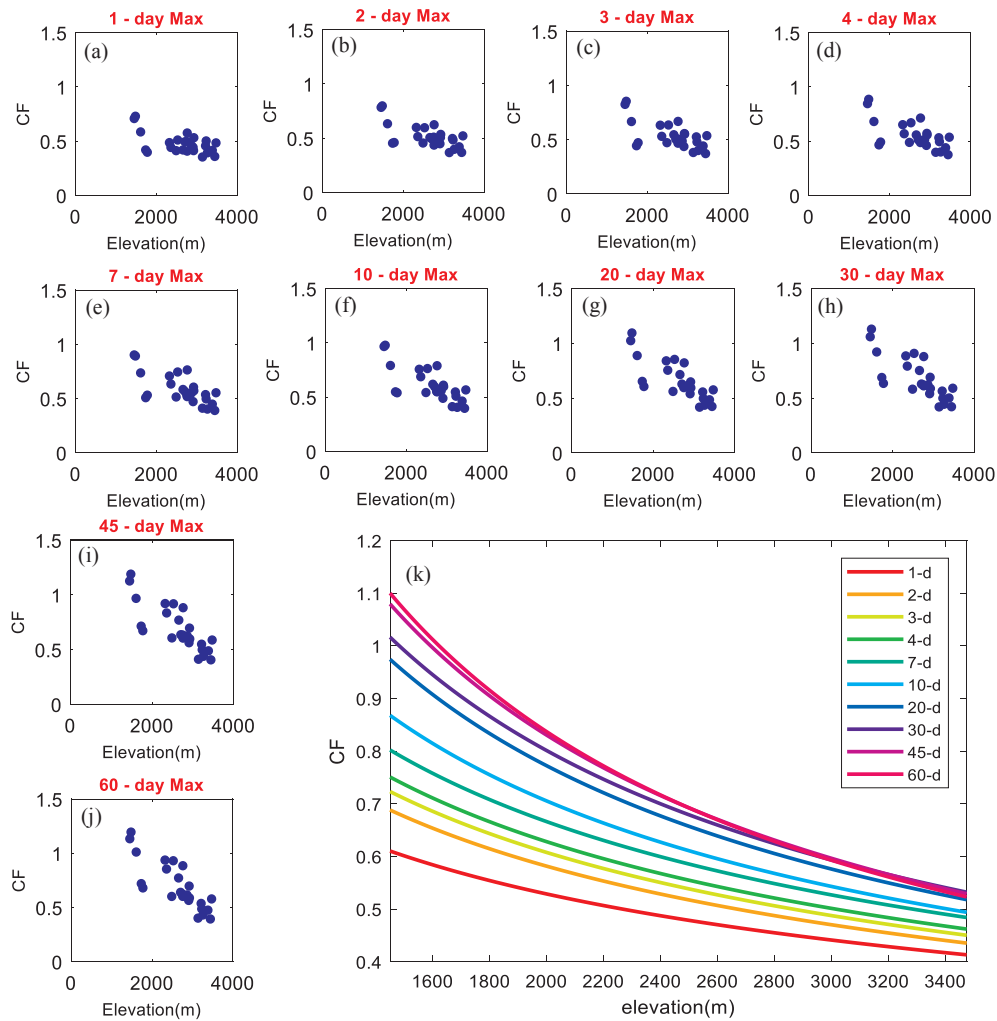As seen in Fig. 10, DDF curves from the original PERSIANN-CDR

Fig. 9. Scatterplots of CF-elevation for different durations (a–f) and the exponential regression fitted to the CF-elevation data for different durations (k) at Colorado Headwaters Basin.

suggest underestimation of the extreme precipitation quantiles for different durations. This is expected given the spatial resolution of this dataset. In fact, when considering the remotely sensed precipitation information, we should be aware that the pixel value represents a spatial average of precipitation within the extent of a pixel. In other words, the pixel value disregards the subpixel variability and even if the PERSIANN-CDR estimate at a pixel is completely accurate, its value tends to be smaller than the collocated point measurements. As a result, the extracted DDF curves from a satellite pixel tend to demonstrate lower return levels (Peleg et al., 2018b), as observed in Fig. 10.

It is worth noting that the estimated DDF curves from the PERSI-ANN-CDR data are not necessarily based on the liquid-phase precipitations and the extracted AMS may comprise snowfalls as well. This is because PERSIANN-CDR and many other satellite-based precipitation estimation algorithms do not distinguish between precipitation phases. In other words, the annual maximum time series extracted and used here may contain solid-phase precipitation extremes due to snowfalls. Although this study does not differentiate between solid and liquid phases of precipitation in order to obtain purely rain-based DDF curves, the current framework can be further modified to incorporate additional observations on solid precipitations. There are two approaches to achieve this goal. One approach for this would be to limit the analysis to warm seasons, but the definitions of warm season vary among various geographic locations. Another approach would be to distinguish snowfall from rainfall, but this would require snowfall and air temperature data that are not available everywhere. Future independent

research may improve upon the current study by including such additional information.

Fig. 11 displays the box plots of RMSE of the return level estimates from the original and corrected CDR for different durations and return periods at collocated PERSIANN-CDR pixels and gauges for different basins in the continental US. The corrected PERSIANN-CDR data was used to obtain frequency estimates at different gauge locations in the selected basins, and the results were compared with those from NOAA Atlas 14. Note that three out of the 16 basins (basins #5, #6 and #15) were located in the Pacific Northwest region and two basins (basins #8 and #14) were located in the state of Texas, all of which were not covered by or were being updated in the recent volumes of NOAA Atlas 14. Thus, the frequency estimates were only validated at the remaining 11 basins as shown in Fig. 11.

According to Fig. 11, the RMSE values for return level estimates corresponding to longer return periods are generally higher for both datasets. This is expected as the PERSIANN-CDR dataset is relatively shorter than the gauge information used for the development of NOAA Atlas 14. The shorter record will result in smaller samples, higher uncertainties, and larger deviations at the tails of the distribution. Over the tested basins, the frequency estimates from the corrected PERSI-ANN-CDR data have consistently lower median RMSE values than those from the original PERSIANN-CDR at different return periods. The RMSE values at the basins with higher elevations (such as Central Nevada, or Dirty Devil basins) were relatively lower than these at basins with lower elevations (such as Mississippi Headwaters, or Upper Mississippi-Iowa
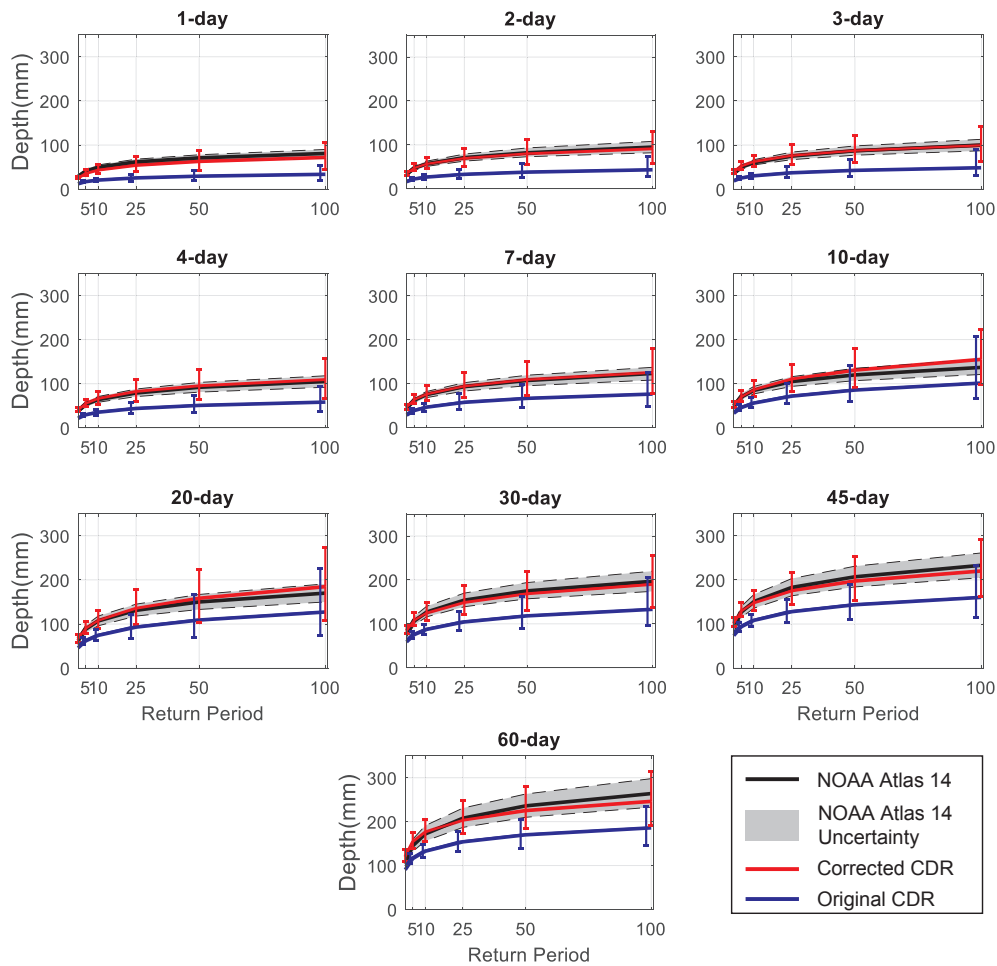
Fig. 10. Frequency estimates from corrected PERSIANN-CDR, original PERSIANN-CDR and NOAA Atlas 14 with 90% Confidence intervals at gauge USC00420849 in Dirty Devil basin, UT. The red vertical bars, blue vertical bars, and gray shaded area show the 90% Confidence intervals from corrected PERSIANN-CDR, original PERSIANN-CDR and NOAA Atlas 14, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

basins), which implies the suitability of the correction approach for high elevation regions. Furthermore, in most of the basins and at different return periods, the corrected dataset shows lower variability in RMSE of the frequency estimates. Corrected PERSIANN-CDR data also demonstrate superior performance in terms of median RMSE and variability of RMSE values at the gauges within the basins. The only case for which the corrected PERSIANN-CDR results in higher RMSE values at different return periods and durations is the South Florida basin, where it was previously shown that the correction model does not improve the AMS estimates due to the climate and the precipitation mechanism.

The relative errors of the frequency estimates are calculated to show the relative magnitude of the return level errors compared to the return levels from NOAA Atlas 14. The relative error here is the difference between the frequency estimates from PERSIANN-CDR (original and bias-adjusted) and NOAA Atlas 14, divided by the value from NOAA Atlas 14. Fig. 12 demonstrates the absolute value of the relative error for the frequency estimates at different durations and return periods from the corrected and original PERSIANN-CDR data. As shown, the relative errors from the corrected PERSIANN-CDR data have consistently lower median values, as well as, lower variability at different return periods in the tested basins. The median relative errors from the corrected data are less than 20% different from the return levels estimated by NOAA Atlas 14. Similar performance is observed when the relative errors of frequency estimates from the two datasets are compared with respect to the extreme precipitation duration. It is also noted

that the corrected PERSIANN-CDR dataset does not show a systematic increase or decrease in the relative errors of the frequency estimates, with respect to the duration. The relative errors of the return level estimates from the original PERSIANN-CDR data tend to decrease with increasing duration. This finding is consistent with our observations in Fig. 9 that revealed lower errors of the original PERSIANN-CDR data for longer duration extreme events. As with Fig. 11, the only case in which the corrected data resulted in higher RMSE values was the South Florida basin where the correction model did not improve the AMS estimates (Section 4.5).

5. Summary and conclusions

In this study, the application of the PERSIANN-CDR dataset for rainfall frequency analysis was investigated. A bias correction model was developed to further correct the PERSIANN-CDR estimates of annual maximum time series at the pixel scale. The proposed correction approach was implemented in two steps: (1) Bias correction factors at limited gauge locations were estimated using linear regression analysis between annual maximum series (AMS) of gauges and collocated pixels; and (2) The correction factors from the limited gauge locations were extended to other regions where gauge data were not available. The correction model was validated at 16 basins in the continental United States, covering various climates and elevations. Finally, depth-duration-frequency (DDF) curves were constructed by fitting the Generalized Extreme Value distribution to the AMS from the corrected data and
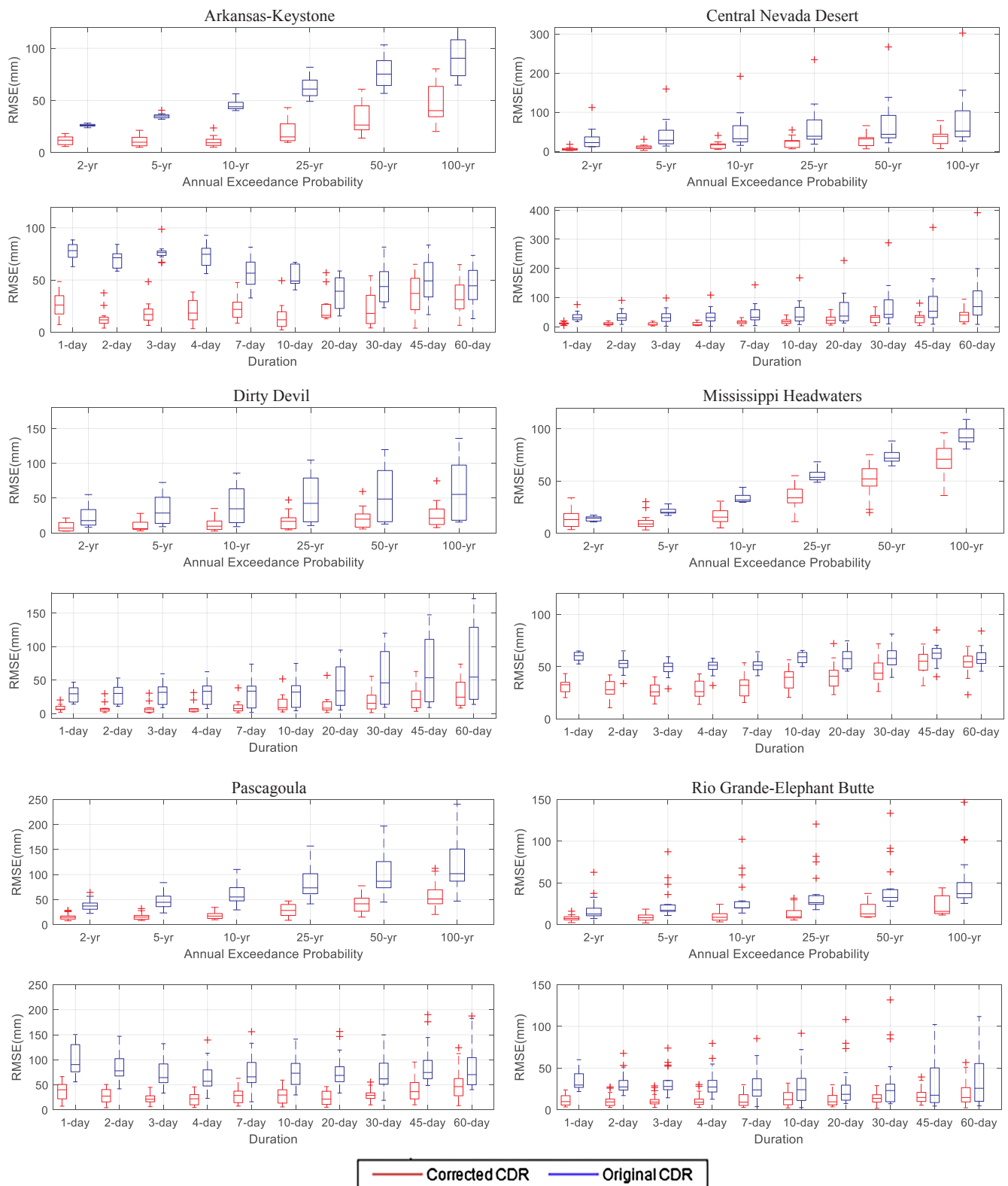
Fig. 11. RMSE of corrected (red) and original (blue) PERSIANN-CDR frequency estimates for the selected basins in the continental United States. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

estimating the quantiles of extreme precipitations. Below is a summary of our main findings:

1. The proposed bias correction approach has been demonstrated effective and robust in improving the accuracy of a remote sensing

precipitation estimation product (i.e., PERSIANN-CDR), especially in high-elevation river basins where gauge or radar networks are either limited or non-existent.

2. The hold-out cross-validation results indicated that the proposed bias correction model is capable of improving the AMS estimated by
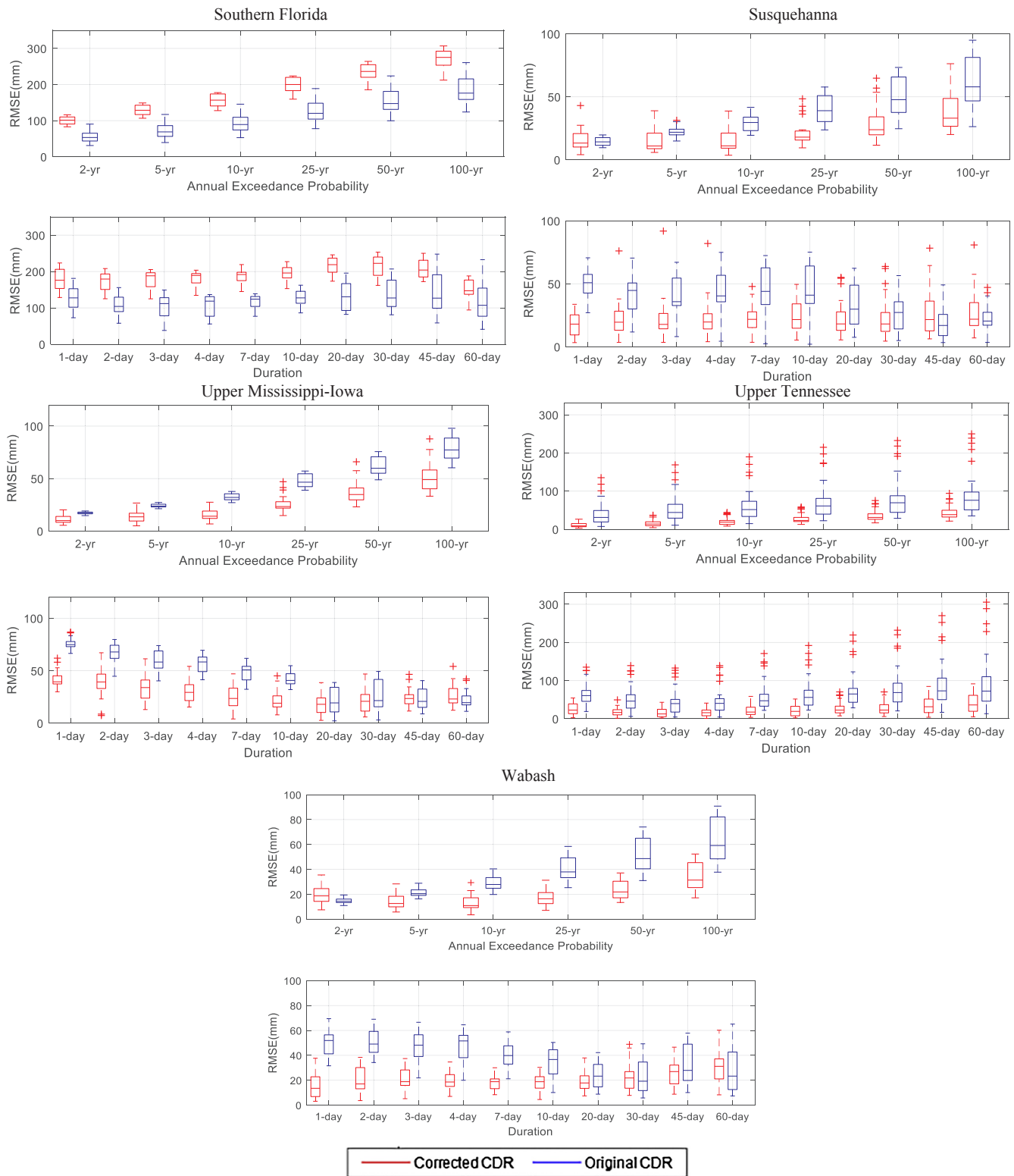
Fig. 11. (*continued*)

the PERSIANN-CDR dataset even in the case that limited gauge information was provided for the model calibration and the approach is generalizable to other locations with similar climates and elevations.

3. As shown by the leave-one-out cross-validation, the bias adjusted PERSIANN-CDR gave better estimates of the AMS for the ungauged sites at a majority of the basins even though these basins had dense

gauge networks.

4. Results from the k-fold cross-validation method suggested that the PERSIANN-CDR data, bias-corrected with the proposed correction approach, performs consistently better than the gauge interpolation method in estimating the AMS at a majority of regions with limited gauge observations. It was observed that the gauge interpolation may sometimes result in significant errors in AMS estimates,
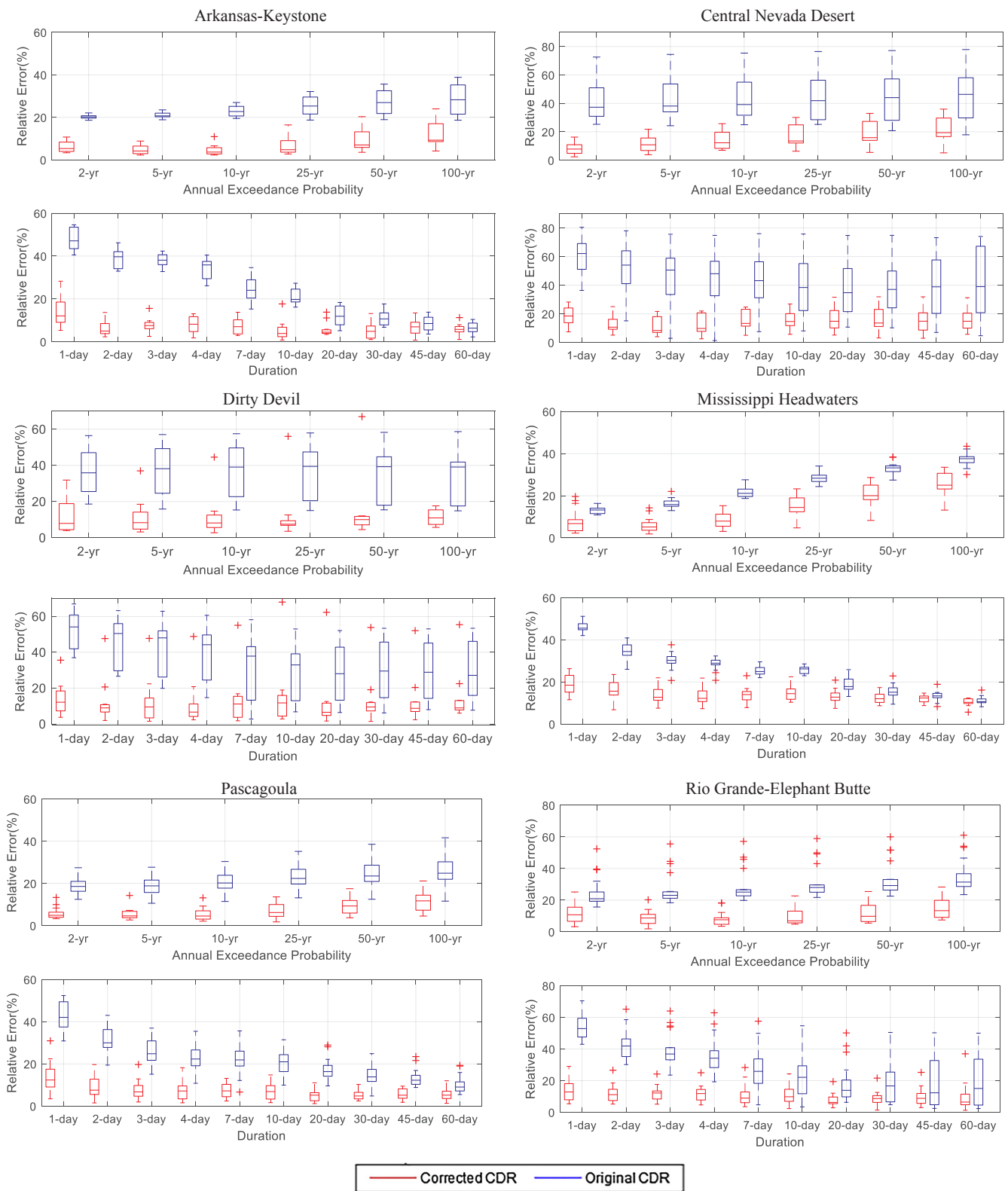
Fig. 12. Absolute relative error (%) of return level estimates for different durations from corrected (red) and original (blue) PERSIANN-CDR for the selected basins in the United States. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

especially in regions with complex topography.

5. The validation results over 16 basins across different climates and elevations indicated that the proposed correction method improves the PERSIANN-CDR estimates of AMS, especially in high elevation regions.

6. The bias-adjusted PERSIANN-CDR is further applied to derive the return levels for different return periods and durations. The frequency estimates from the corrected PERSIANN-CDR data are compared with those from the original PERSIANN-CDR and NOAA Atlas 14. Results revealed that the frequency estimates from the
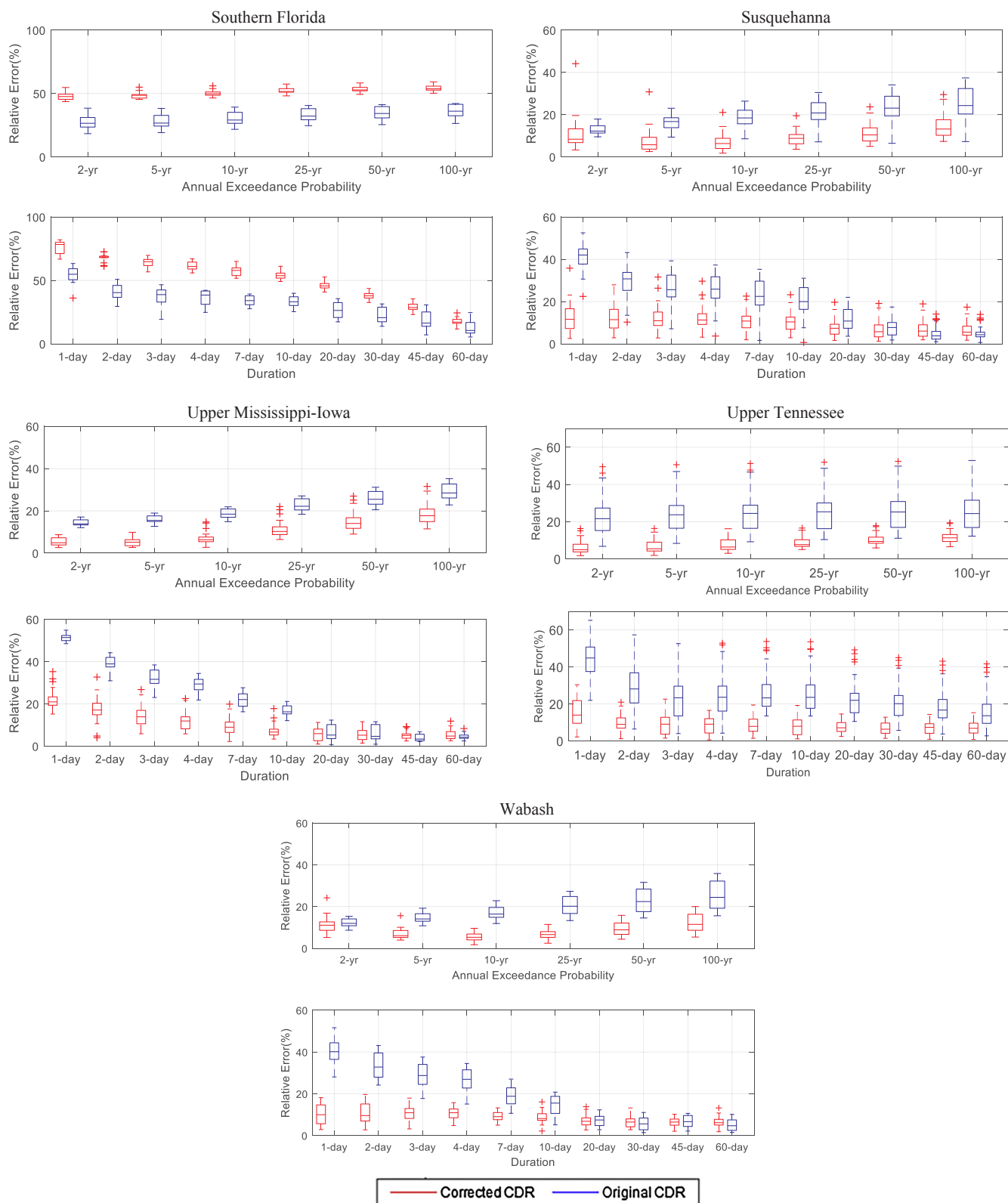
Fig. 12. (*continued*)

corrected dataset are consistently closer to the estimates from NOAA Atlas 14. They also lie within the uncertainty bounds of NOAA Atlas 14.

Thus, the PERSIANN-CDR dataset has the potential for being used in rainfall frequency analysis for the regions with limited ground-based observations. However, despite the promising results, there are still some limitations in this dataset and the proposed correction method for the application of frequency analysis. One of these limitations is the temporal resolution of the PERSIANN-CDR dataset. The daily temporal resolution limits the investigation of extreme events with shorter durations (e.g. 3-hourly or hourly). Another limitation is that the

frequency analysis here is conducted at the pixel scale using relatively limited samples. A sample of 33 annual maximum values is relatively limited for fitting a 3-parameter distribution. This would result in high uncertainties in estimating the parameters of the distribution and the return levels. One remedy to the sample size problem could be the application of regional frequency analysis methods to increase the sample size by incorporating information from the nearby locations with the same climatic conditions.

It is also important to note that given the rising global temperatures, rainfall intensities especially at shorter durations are expected to increase. Therefore, the increase in the global temperature could be used as an added factor to adjust historical design rainfall intensities for the warmer temperatures that lie ahead (Peleg et al., 2018a).

This work is part of an ongoing research and the presented approaches and results are intended as a proof of concept. Future research in this area may involve, bringing non-stationarities into the bias-adjustment framework (Tao et al., 2018), including covariates into the bias-adjustment framework which requires advanced optimization techniques (Yang et al., 2017), investigating the hydrological modeling applications of the corrected-PERSIANN-CDR data, and developing DDF curves for ungauged regions or areas not included in the current NOAA Atlas 14.

Acknowledgments

References

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., 2003. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). J. Hydrometeorol. 4 (6), 1147–1167.

Allen, R.J., DeGaetano, A.T., 2005. Considerations for the use of radar-derived precipitation estimates in determining return intervals for extreme areal precipitation amounts. J. Hydrol. 315 (1), 203–219.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P., 2015. PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull. Am. Meteorol. Soc. 96 (1), 69–83.

Awadallah, A.G., ElGamal, M., ElMostafa, A., ElBadry, H., 2011. Developing intensity-duration-frequency curves in scarce data region: an approach using regional analysis and satellite data. Engineering 3 (03), 215.

Ben-Zvi, A., 2009. Rainfall intensity–duration–frequency relationships derived from large partial duration series. J. Hydrol. 367 (1), 104–114.

Bonnin, G.M., Martin, D., Lin, B., Parzybok, T., Yekta, M., Riley, D., 2006. Precipitation-frequency atlas of the United States. NOAA Atlas 14 (2).

Bougadis, J., Adamowski, K., 2006. Scaling model of a rainfall intensity-duration-frequency relationship. Hydrol. Processes 20 (17), 3747–3757.

Chen, M., Shi, W., Xie, P., Silva, V., Kousky, V.E., Wayne Higgins, R., Janowiak, J.E., 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. J. Geophys. Res.: Atmos. 113 (D4).

Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S., Ropelewski, C., 2008. Validation of high-resolution satellite rainfall products over complex terrain. Int. J. Remote Sens. 29 (14), 4097–4110.

Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., Zougmore, F., 2014. Rainfall monitoring based on microwave links from cellular telecommunication networks: first results from a West African test bed. Geophys. Res. Lett. 41 (16), 6016–6022.

Duan, Z., Liu, J., Tuo, Y., Chiogna, G., Disse, M., 2016. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. Sci. Total Environ. 573, 1536–1553.

Eldardiry, H., Habib, E., Zhang, Y., 2015. On the use of radar-based quantitative precipitation estimates for precipitation frequency analysis. J. Hydrol. 531, 441–453.

Embrechts, P., Klüppelberg, C., Mikosch, T., 2013. Modelling Extremal Events: For Insurance and Finance. Springer Science & Business Media.

Endreny, T.A., Imbeah, N., 2009. Generating robust rainfall intensity–duration–frequency estimates with short-record satellite data. J. Hydrol. 371 (1), 182–191.

Fowler, H., Kilsby, C., 2003. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. Int. J. Climatol. 23 (11), 1313–1334.

Gado, T.A., Hsu, K., Sorooshian, S., 2017. Rainfall frequency analysis for ungauged sites using satellite precipitation products. J. Hydrol. 554, 646–655. http://dx.doi.org/10.1016/j.jhydrol.2017.09.043.

Gebregiorgis, A., Hossain, F., 2014. Making satellite precipitation data work for the developing world. IEEE Geosci. Remote Sens. Mag. 2 (2), 24–36.

Gellens, D., 2002. Combining regional approach and data extension procedure for assessing GEV distribution of extreme precipitation in Belgium. J. Hydrol. 268 (1), 113–126.

Hashemi, H., Nordin, M., Lakshmi, V., Huffman, G.J., Knight, R., 2017. Bias correction of long-term satellite monthly precipitation product (TRMM 3B43) over the conterminous United States. J. Hydrometeorol. 18 (9), 2491–2509.

Hong, Y., Gochis, D., Cheng, J.T., Hsu, K.L., Sorooshian, S., 2007. Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. J. Hydrometeorol. 8 (3), 469–482.

Hosking, J.R.M., Wallis, J.R., 2005. Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press.

Hsu, K.L., Gao, X., Sorooshian, S., Gupta, H.V., 1997. Precipitation estimation from remotely sensed information using artificial neural networks. J. Appl. Meteorol. 36 (9), 1176–1190. http://dx.doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Huffman, G.J., Adler, R.F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., Schneider, U., 1997. The global precipitation climatology project (GPCP) combined precipitation dataset. Bull. Am. Meteorol. Soc. 78 (1), 5–20.

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., Stocker, E.F., 2007. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8 (1), 38–55.

Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q. J. R. Meteorol. Soc. 81 (348), 158–171.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5 (3), 487–503.

Knapp, K.R., Ansari, S., Bain, C.L., Bourassa, M.A., Dickinson, M.J., Funk, C., Helms, C.N., Hennon, C.C., Holmes, C.D., Huffman, G.J., 2011. Globally gridded satellite observations for climate studies. Bull. Am. Meteorol. Soc. 92 (7), 893–907.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. Meteorol. Z. 15 (3), 259–263.

Kotz, S., Nadarajah, S., 2000. Extreme Value Distributions: Theory and Applications. World Scientific.

Kubota, T., Ushio, T., Shige, S., Kida, S., Kachi, M., Okamoto, K.I., 2009. Verification of high-resolution satellite-based rainfall estimates around Japan using a gauge-calibrated ground-radar dataset. J. Meteorol. Soc. Jpn. 87, 203–222 Ser. II.

Libertino, A., Sharma, A., Lakshmi, V., Claps, P., 2016. A global assessment of the timing of extreme rainfall from TRMM and GPM for improving hydrologic design. Environ. Res. Lett. 11 (5).

Liu, X.M., Yang, T.T., Hsu, K.L., Liu, C.M., Sorooshian, S., 2017. Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau. Hydrol. Earth Syst. Sci. 21 (1), 169–181.

Liu, C., Zipser, E.J., 2009. "Warm rain" in the tropics: seasonal and regional distributions based on 9 yr of TRMM data. J. Clim. 22 (3), 767–779.

Lombardo, F., Napolitano, F., Russo, F., Scialanga, G., Baldini, L., Gorgucci, E., 2006. Rainfall estimation and ground clutter rejection with dual polarization weather radar. Adv. Geosci. 7, 127–130.

Marra, F., Morin, E., 2015. Use of radar QPE for the derivation of intensity-duration-Frequency curves in a range of climatic regimes. J. Hydrol. 531, 427–440.

Marra, F., Morin, E., Peleg, N., Mei, Y., Anagnostou, E.N., 2017. Intensity–duration–frequency curves from remote sensing rainfall estimates: comparing satellite and weather radar over the eastern Mediterranean. Hydrol. Earth Syst. Sci. 21 (5), 2389.

Miao, C., Ashouri, H., Hsu, K.-L., Sorooshian, S., Duan, Q., 2015. Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China. J. Hydrometeorol. 16 (3), 1387–1396.

Norbiato, D., Borga, M., Sangati, M., Zanon, F., 2007. Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood. J. Hydrol. 345 (3), 149–166.

Overeem, A., Buishand, A., Holleman, I., 2008. Rainfall depth-duration-frequency curves and their uncertainties. J. Hydrol. 348 (1–2), 124–134.

Overeem, A., Buishand, T.A., Holleman, I., 2009. Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. Water Resour. Res. 45.

Peleg, N., Marra, F., Fatichi, S., Molnar, P., Morin, E., Sharma, A., Burlando, P., 2018a. (2018a) Intensification of convective rain cells at warmer temperatures observed from high-resolution weather radar data. J. Hydrometeorol.

Peleg, N., Marra, F., Fatichi, S., Paschalis, A., Molnar, P., Burlando, P., 2018b. Spatial variability of extreme rainfall at radar subpixel scale. J. Hydrol. 556, 922–933.

Shah, H.L., Mishra, V., 2016. Uncertainty and bias in satellite-based precipitation estimates over indian subcontinental basins: implications for real-time streamflow simulation and flood prediction*. J. Hydrometeorol. 17 (2), 615–636.

Shige, S., Kida, S., Ashiwake, H., Kubota, T., Aonashi, K., 2013. Improvement of TMI rain retrievals in mountainous areas. J. Appl. Meteorol. Climatol. 52 (1), 242–254.

Sorooshian, S., Gao, X., Hsu, K., Maddox, R., Hong, Y., Gupta, H., Imam, B., 2002. Diurnal variability of tropical rainfall retrieved from combined GOES and TRMM satellite information. J. Clim. 15 (9), 983–1001.

Stedinger, J.R., 1993. Frequency analysis of extreme events. Handb. Hydrol. 18.

Tao, Y., Yang, T., Faridzad, M., Jiang, L., He, X., Zhang, X., 2018. Non-stationary bias correction of monthly CMIP5 temperature projections over China using a residual-based bagging tree model. Int. J. Climatol. 38 (1), 467–482.

Villarini, G., Smith, J.A., Baeck, M.L., Vitolo, R., Stephenson, D.B., Krajewski, W.F., 2011. On the frequency of heavy rainfall for the Midwest of the United States. J. Hydrol. 400 (1), 103–120.

Wright, D.B., Smith, J.A., Villarini, G., Baeck, M.L., 2013. Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. J. Hydrol. 488, 150–165.

Yang, T., Asanjan, A.A., Faridzad, M., Hayatbini, N., Gao, X., Sorooshian, S., 2017. An enhanced artificial neural network with a shuffled complex evolutionary global optimization with principal component analysis. Inf. Sci. 418, 302–316.

Yang, X., Yong, B., Hong, Y., Chen, S., Zhang, X., 2016. Error analysis of multi-satellite precipitation estimates with an independent raingauge observation network over a medium-sized humid basin. Hydrol. Sci. J. 61 (10), 1813–1830.

Zhou, Y.P., Lau, W.K.M., Huffman, G.J., 2015. Mapping TRMM TMPA into Average Recurrence Interval for Monitoring Extreme Precipitation Events. J. Appl. Meteorol. Climatol. 54 (5), 979–995.