

Water Resources Research

RESEARCH ARTICLE

10.1029/2019WR026444

Key Points:

- A generalized three-cornered hat (TCH) method is developed to compare the random error of different precipitation products
- Weighting multiple precipitation data based on the inverse of TCH estimated error-covariance can help improve global precipitation estimation
- The weighted precipitation data substantially reduces the random errors and outperforms GPM IMERG and two other merging methods

Supporting Information:

- Supporting Information S1

Correspondence to:

X. Zhang and C. Hu,
zhangxiangsw@whu.edu.cn;
huchl@cug.edu.cn

Citation:

Xu, L., Chen, N., Moradkhani, H., Zhang, X., & Hu, C. (2020). Improving global monthly and daily precipitation estimation by fusing gauge observations, remote sensing, and reanalysis data sets. *Water Resources Research*, 56, e2019WR026444. <https://doi.org/10.1029/2019WR026444>

Received 1 OCT 2019

Accepted 26 FEB 2020

Accepted article online 28 FEB 2020

Improving Global Monthly and Daily Precipitation Estimation by Fusing Gauge Observations, Remote Sensing, and Reanalysis Data Sets

Lei Xu^{1,2} , Nengcheng Chen^{1,3} , Hamid Moradkhani² , Xiang Zhang¹ , and Chuli Hu^{4,5}

¹State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China, ²Center for Complex Hydrosystems Research, Department of Civil, Construction, and Environmental Engineering, The University of Alabama, Tuscaloosa, AL, USA, ³Collaborative Innovation Center of Geospatial Technology, Wuhan, China, ⁴School of Geography and Information Engineering, China University of Geoscience, Wuhan, China, ⁵National Engineering Research Center for Geographic Information System, China University of Geoscience, Wuhan, China

Abstract Precipitation estimation at a global scale is essential for global water cycle simulation and water resources management. The precipitation estimation from gauge-based, satellite retrieval, and reanalysis data sets has heterogeneous uncertainties for different areas at global land. Here, the 13 monthly precipitation data sets and the 11 daily precipitation data sets are analyzed to examine the relative uncertainty of individual data based on the developed generalized three-cornered hat (TCH) method. The generalized TCH method can be used to evaluate the uncertainty of multiple (>3) precipitation products in an iterative optimization process. A weighting scheme is designed to merge the individual precipitation data sets to generate a new weighted precipitation using the inverse error variance-covariance matrix of TCH estimated uncertainty. The weighted precipitation is then validated using gauged data with the minimal uncertainty among all the individual products. The merged results indicate the superiority of the weighted precipitation with substantially reduced random errors over individual data sets and a state-of-the-art multisatellite merged product, namely, the Integrated Multi-Satellite Retrievals for Global Precipitation Measurement at validated areas. The weighted data set can largely reproduce the interannual and seasonal variations of regional precipitation. The TCH-based merging results outperform two other mean-based merging methods at both monthly and daily scales. Overall, the merging scheme based on the generalized TCH method is effective to produce a new precipitation data set integrating information from multiple products for hydrometeorological applications.

1. Introduction

Precipitation is one of the major variables influencing global water cycle (Hou et al., 2014; Legates & Willmott, 1990; Schneider et al., 2014). Precipitation estimation at a regional or global scale has important applications for hydrological, meteorological, agricultural, and socioeconomic sectors, such as drought and flood early warning, water management, crop water demand, and electricity generation (Jalota et al., 2006; Khajehei et al., 2017; Siddique-E-Akbor et al., 2014; Xu et al., 2018; Xu et al., 2019). However, it is challenging to estimate global precipitation accurately at daily, monthly, and yearly scales because the uncertainties in instruments, retrieval methods, and numerical models are still high (Sun et al., 2018), especially at high spatial resolution (e.g., 5 km). There are currently three ways to estimate global precipitation: rain gauges, satellite retrieval, and reanalysis. A number of global precipitation products at different spatial and temporal scales are currently available based on these ways. However, there is not a unique best data set yet. These precipitation products may have different accuracy at different areas (Xu et al., 2019). How to quantify and reduce the uncertainty of different precipitation products is helpful for global hydrometeorological applications.

A direct method to observe precipitation is to use rain gauges. Rain gauge measures the depth of rainfall accumulated over time. Different types of rain gauges have their advantages and disadvantages (Sun et al., 2018). Rain gauge measurement is considered as the most accurate way to estimate precipitation relative to satellite retrieval and model simulation. However, the number of globally available rain gauges is sparse and irregular (Becker et al., 2013; Harris et al., 2014; Schneider et al., 2016), with more rain gauges

in Europe and North America and less in other continents. There are also many missing values in the global rain gauge records. The density and number of rain gauges strongly affect the accuracy of gauge-based precipitation data sets. Interpolation techniques and weighting schemes also influence the precipitation estimation at gridded data. These factors contribute to the uncertainties of gauge-based precipitation products, especially at ungauged and sparsely gauged areas.

The emerging satellite sensors provide great opportunities to observe Earth's precipitation distribution from onboard instruments (Hou et al., 2014). These sensors include visible/infrared (VIS/IR) sensors, active microwave (MW), and passive MW (PMW) sensors. VIS/IR methods relate bright and cold clouds to convection and use the cloud top temperature to estimate precipitation on the ground. PMW can penetrate clouds and is sensitive to precipitation-sized particles. PMW sensors, such as the Special Sensor Microwave/Imager (SSM/I) (Hollinger et al., 1987), Tropical Rainfall Measuring Mission (TRMM) MW Imager (TMI) (Kummerow et al., 1998), advanced MW sounding unit (AMSU), and the Advanced MW Scanning Radiometer for the Earth Observing System (AMSR-E) (Kawanishi et al., 2003), enable the spatiotemporal observations of global precipitation. Active MW observations are also used in precipitation estimation such as the TRMM mission (Kummerow et al., 1998). The ongoing Global Precipitation Measurement (GPM) mission (Hou et al., 2014) uses a constellation of MW sensors to obtain unified precipitation retrieval and more accurate instantaneous precipitation estimates. Numerous precipitation retrieval algorithms have been developed, of which the Goddard Profiling algorithm (Kummerow et al., 2001) is mostly used. Compared to gauge-based measurement, satellite estimates have broader spatial extent, capable of monitoring large-scale rainfall events. The accuracy of precipitation estimation using satellite sensors is dependent on instruments, retrieval algorithms, and observed quality.

The reanalysis data are produced based on data assimilation algorithms, observations, and physically based numerical models. Various observed data are assimilated into process models to generate physically consistent, spatiotemporally homogeneous, and synthesized climatic variables. Several reanalysis data sets are developed by individual research centers worldwide, such as the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA) (Dee et al., 2011), the Japanese 55-year Reanalysis (JRA55) (Kobayashi et al., 2015), and the Modern-Era Retrospective analysis for Research and Applications (MERRA) (Gelaro et al., 2017; Rienecker et al., 2011). The accuracy of precipitation reanalysis data is dependent on observations, assimilation methods, model structure, and parameterization process. With the development of data assimilation algorithms (Abbaszadeh et al., 2019; Pathiraja et al., 2018; Reichle et al., 2008), increasing climatic observations and model optimization, precipitation estimation is gradually improved in reanalysis data. The reanalysis data are widely used in near real-time hydrometeorological applications (Bastola & Misra, 2014; Serreze et al., 2003).

The gridded precipitation data are commonly validated using rain gauges. It is challenging to validate current precipitation products at a global scale using dense rain gauge networks because the available stations are sparse and irregular, especially at mountainous and remote areas. The fraction of global land grids that covers at least four rain gauges on average for a 0.5° grid is less than 4% in the Global Precipitation Climatology Centre (GPCC) data in January 1983 and 2016 (Figure 1). On the other hand, using point observations to represent grid values may suffer from representativeness error (O'Carroll et al., 2008; Swinbank et al., 2012), which is dominant in validation error. Therefore, it is currently impossible to validate precipitation products globally using available rain gauges.

Recently, there are some validation researches that are not dependent on gauge stations (Gruber et al., 2016; Liu et al., 2013; McColl et al., 2014; Rieckh & Anthes, 2018). The three-cornered hat (TCH) method and the triple collocation (TC) method are two commonly used ways to evaluate the uncertainty of remote sensing or hydrological products without using ground gauges (Ferreira et al., 2016; Long et al., 2014; Massari et al., 2017). TCH is a difference-based method to estimate the uncertainty of a specific variable by removing the common signal of true observations. It requires three independent data sets to obtain the uncertainty when TCH was first proposed. Later, an expanded research loosened the limitation and considered the correlation between data sets by minimizing the global correlation (Premoli & Tavella, 1993). TCH was first proposed to measure the instability of clocks, and little research has examined its validity in precipitation assessment (Ferreira et al., 2016; Long et al., 2014). TC is a method to estimate the error variance of a geophysical variable using three collocated data sets (Stoffelen, 1998). The TC method assumes error

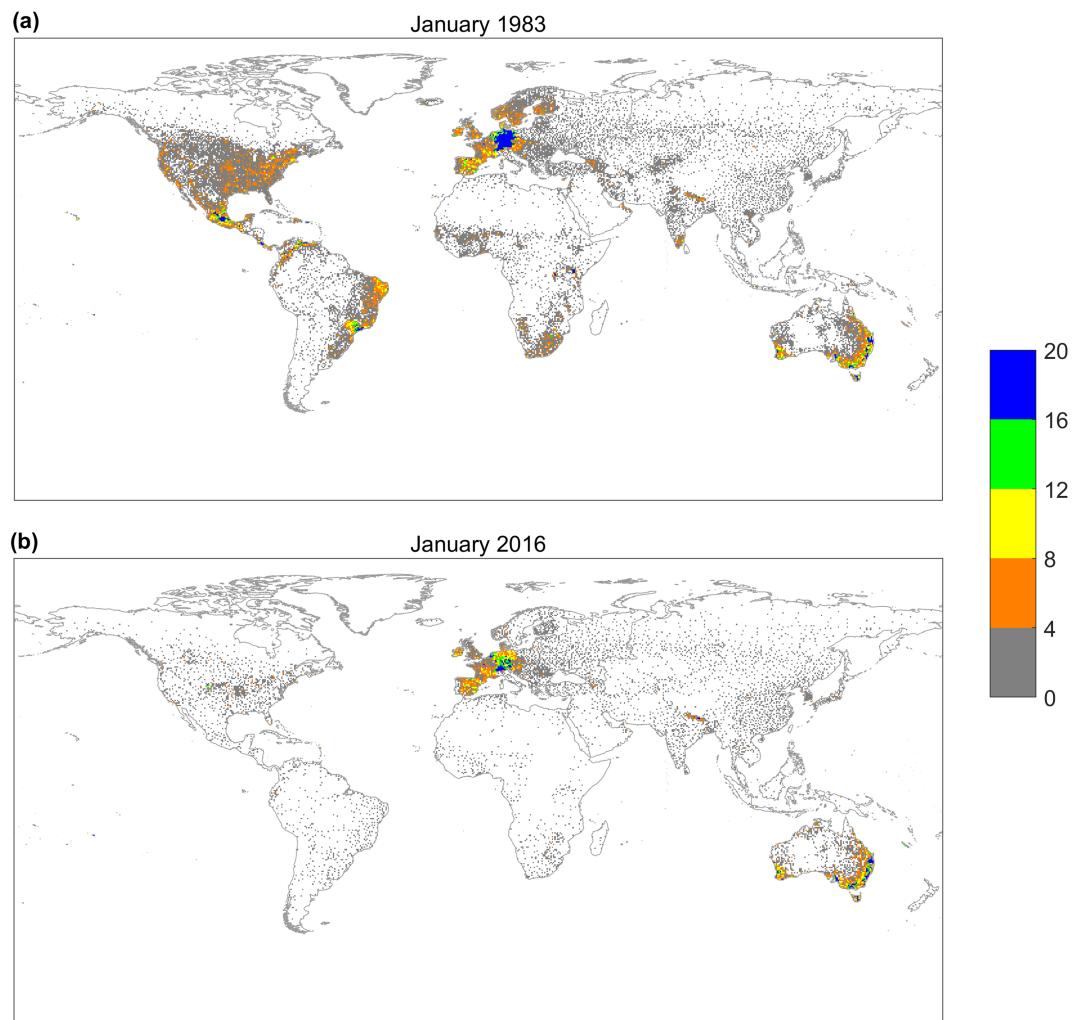


Figure 1. The available rain gauge stations in the GPCC data at (a) January 1983 and (b) January 2016. The numbers in the legend denotes the amount of stations in a 0.5° grid.

stationarity, error orthogonality, and zero-error cross-correlation between the collocated data (Gruber et al., 2016). Although some studies proposed some adaptations by violating some assumptions, the original TC method obeying these assumptions is mostly used and seems more robust than those adaptations.

The accuracy of available precipitation products varies with locations, topography, and climate regimes. It seems promising to combine the strengths of multiple precipitation products to produce a merged data set with higher accuracy. A recent study (Beck et al., 2017) integrates seven different precipitation products by a weighted average to generate a new data set named Multi-Source Weighted-Ensemble Precipitation. The weights are determined by the density of rain gauge networks at gauged areas and by the comparative performance using surrounding rain gauges at ungauged areas. A further validation indicates the good performance of Multi-Source Weighted-Ensemble Precipitation products (Beck et al., 2019). Therefore, it is a feasible way to merge multisource precipitation data sets to improve precipitation accuracy. Despite the merging study, more available precipitation products can be added and the weighting scheme can be adapted to relax the dependence of gauges when determining weights because ungauged or sparsely gauged areas may be dominant in substantial parts of global land and the representativeness error may not be well quantified.

Here, the original TCH method is extended to multiple kinds of data to evaluate the uncertainty of 13 monthly gauge-based, satellite-estimated, and reanalysis precipitation products and 11 daily products spanning from 2003 to 2016. The uncertainty refers to the random error (second-order error) in this work.

The inverse of estimated error variance-covariance is then regarded as weights to merge these precipitation products into a new precipitation data set. The TC method has been used in weighting multiple soil moisture data (Gruber et al., 2017); however, no previous research is found in weighting using TCH approach. The difference between TC and TCH methods lies in the modeling of error correlation. The TC method may fail in the presence of error cross-correlation between multiple data (e.g., precipitation), while TCH can reduce the error correlation in an optimization process. A preliminary validation based on the generalized TCH method indicates an advantage of the fused data set over the individual precipitation products involved. This merging scheme can be regarded as an easy but effective way to combine multiple precipitation products to improve precipitation estimation at a global scale.

2. Data

2.1. Gauge-Based Data Sets

The GPCC collected rain gauge data from all over the world, resulting in a database with more than 85,000 stations worldwide and surpassing 200-year coverage in time (Becker et al., 2013; Schneider et al., 2016). The GPCC database integrates rain gauge records from several sources, including the near real-time collection from the Global Telecommunication System and offline data collection from World Meteorological Organization members, international regional projects, Climate Research Unit (CRU), the United Nations Food and Agriculture Organization, and the Global Historical Climatology Network (GHCN) at National Climatic Data Center, USA. GPCC provides 0.5°, 1°, and 2.5° precipitation data from 1901 to present, with the monitoring and first-guess products keeping up-to-date and full data reanalysis left behind. GPCC is regarded as the global precipitation data set with maximum gauges versus other data sets. A daily version (Ziese et al., 2018) of the GPCC data are also released at 0.5°, 1°, and 2.5° resolutions from 1982 to 2016.

The CRU TimeSeries (CRU TS) is a monthly 0.5° gauge-based climatic data set across world's land areas excluding Antarctica (Harris et al., 2014), developed by the University of East Anglia. The rain gauges used in the CRU TS come from the monthly climate data worldwide of World Meteorological Organization and the additional monthly data from several countries. There are two parallel versions of this data: version 3.xx and 4.xx. CRU TS version 3 uses the IDL routines "triangulate" and "trigrid" to effect triangulated linear interpolation, and version 4 uses the angular distance weighting for gridding the monthly anomalies. The use of angular distance weighting enables the total control of selecting station observations for gridding. CRU TS version 3 is not planned to be continued, and the current version 4 data is CRU TS v4.02. The monthly CRU TS v4.02 precipitation data and its stations are collected in this study.

The Climate Prediction Center (CPC) provides a unified gauge-based global daily precipitation data set from 1979 to present at a 0.5° resolution (Xie et al., 2007; Xie et al., 2010). CPC unified precipitation integrates data from in situ rain gauges from the Global Telecommunication System, satellite estimates from CPC Morphing technique (CMORPH), and numerical model-based precipitation from National Oceanic and Atmospheric Administration (NOAA) National Environmental Prediction Center Global Forecast System by optimal interpolation algorithm. The CPC unified precipitation from 1979 to 2005 is the retrospective version with more than 30,000 gauges involved, and the data from 2006 to present is the real-time version with approximately 17,000 gauges used.

The University of Delaware (UDEL) established a global monthly 0.5° precipitation data set from 1900 to near present (Willmott & Matsuura, 1995) using rain gauges from the GHCN version 2, a version of the Daily GHCN (GHCN-Daily) and station records from other countries and organizations. The latest version of UDEL data extends to year 2017. The climatologically aided interpolation and an enhanced distance-weighting method with the spherical version of Shepard's algorithm were used to perform spatial interpolation in this data.

2.2. Satellite-Estimated Data Sets

The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Nguyen et al., 2019; Sorooshian et al., 2014) is a global daily precipitation data set covering 60°N to 60°S at a 0.25° resolution from 2000 to present, developed by the Center for Hydrometeorology and Remote Sensing at the University of California, Irvine. PERSIANN (Nguyen et al., 2019; Sorooshian et al., 2014) is produced using neural network function classification/approximation procedures to

compute an estimate of rainfall rate based on geostationary IR brightness temperature image. Later, the daytime visible imagery is added into the PERSIANN system.

CMORPH (Joyce et al., 2004) is a global daily 0.25° precipitation data set covering 60°S to 60°N . CMORPH estimates precipitation from low orbiter satellite MW observations and propagates the precipitation information to geostationary satellite IR data by motion vectors. CMORPH integrates precipitation estimates from the PMWs aboard the Defense Meteorological Satellite Program 13, 14, and 15 (SSM/I), the NOAA-15, 16, 17, and 18 (AMSU-B), and AMSR-E and TMI aboard the National Aeronautics and Space Administration's (NASA) Aqua and TRMM spacecraft, respectively.

The TRMM mission (Kummerow et al., 1998) was launched in 1997 by NASA and the Japan Aerospace Exploration Agency to estimate rainfall, and the onboard instruments were turned off on June 2015. Satellite sensors onboard TRMM include the VIS/IR radiometer, TMI, and precipitation radar. The 3-hr daily and monthly precipitation data are issued from 1998 to present at a 0.25° spatial resolution covering 50°S to 50°N . The TRMM Multi-Satellite Precipitation Analysis (Huffman et al., 2010) product is a research version of TRMM data set by combining precipitation estimates from multiple satellite sensors, including TMI, SSM/I, AMSR-E, AMSU-B, the MW Humidity Sounders, and sensors onboard the European Meteorological Operational satellite. Although the TRMM Multi-Satellite Precipitation Analysis product is planned to be terminated in December 2019, the ongoing GPM mission is a strong successor of TRMM. The TRMM 3B42RT daily precipitation data are collected here for comparison.

The Integrated Multi-satellitE Retrievals for GPM (IMERG) data (Hou et al., 2014; Huffman et al., 2015) is a multisatellite precipitation data set obtained from multiple satellite MW precipitation estimates, MW-calibrated IR satellite estimates, precipitation gauge analyses, and other precipitation estimates globally. First, the MW precipitation estimates from a number of PMW sensors are intercalibrated based on quantile-quantile matching. Then, the MW-calibrated IR precipitation estimates are used to fill the holes in the PMW constellation. Finally, the gauged data from GPCC are incorporated to control bias. Therefore, the IMERG data can be considered as a kind of integrated precipitation product. The monthly and daily GPM IMERG data are used to compare with the TCH merged data set.

The Global Satellite Mapping of Precipitation (GSMaP) product (Kubota et al., 2007; Ushio et al., 2003) aims to produce a high-precision, high-resolution global precipitation map using satellite data, developed by Core Research for Evolutional Science and Technology of the Japan Science and Technology Agency during 2002–2007. GS MaP is produced based on the combined MW-IR algorithm using GPM-Core Microwave Imager, TRMM TMI, the Global Change Observation Mission-Water/the Advanced Microwave Scanning Radiometer 2, Defense Meteorological Satellite Program series Special Sensor Microwave Imager/Sounder, NOAA series AMSU, Meteorological Operational satellite series AMSU, and Geostationary IR developed by the GS MaP project. The hourly 0.1° GS MaP data include GS MaP Reanalysis Products (GS MaP_RNL), GS MaP Microwave-IR Combined Product (GS MaP_MVK), GS MaP Near-Real-Time (GS MaP_NRT), a real-time version (GS MaP_NOW), and a nowcast data set by RIKEN institute (GS MaP_RNC).

The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) (Funk et al., 2015) is a quasi-global precipitation data set spanning 50°S to 50°N at a 0.05° resolution, developed by the University of California Santa Barbara. The CHIRPS is generated based on a 0.05° climatology that incorporates satellite information to represent sparsely gauged locations and by incorporating daily, pentadal, and monthly 1981 to present 0.05° IR cold cloud duration-based precipitation estimates and a blending procedure considering the spatial correlation structure of cold cloud duration estimates to assign interpolation weights. The 0.05° CHIRPS version 2.0 data is collected and resampled to 0.5° in this study.

2.3. Reanalysis Data Sets

ERA-Interim (Dee et al., 2011) is a global reanalysis data set based on four-dimensional variational data assimilation algorithm spanning from 1979 to present, developed by ECMWF. ERA-Interim was conducted to overcome some shortages in ERA-40 (Uppala et al., 2005), such as the humidity analysis scheme and bias adjustments for IR radiance. The spatial resolution of the ERA-Interim is approximately 80 km (T255 spectral) on 60 vertical levels from the surface up to 0.1 hPa. The 3-hourly, 6-hourly, daily and monthly data are released at the ECMWF public data sets web interface.

Table 1
A Summary of Global Monthly and Daily Precipitation Data Sets Used in this Study

Type	Data	Resolution	Frequency	Coverage	Period
Gauge-based	CRU TS	0.5°	Monthly	Global land	1901–2017
	GPCC	0.5°	Monthly	Global land	1891–2016
	GPCC-daily	0.5°	Daily	Global land	1982–2016
	CPC-Unified	0.5°	Daily	Global land	1979 to present
	UDEL	0.5°	Monthly	Global land	1900–2017
Satellite-based	PERSIANN	0.25°	1, 3, 6 h/Daily	60°S to 60°N	2000 to present
	CMORPH	0.25°/8 km	30 min/3 h/daily	60°S to 60°N	2002 to present
	TRMM 3B42RT	0.25°	3 h/Daily	50°S to 50°N	2000 to present
	GSMaP	0.1°	1 h/daily	60°S to 60°N	2000 to present
Reanalysis	CHIRPS	0.05°	Daily/monthly	50°S to 50°N	1981 to present
	ERA-Interim	0.75°	6 h/monthly	Global	1979 to present
	JRA55	0.5625°	3, 6 h/monthly	Global	1958 to present
	MERRA-2	0.5° × 0.625°	Daily/monthly	Global	1980 to present
	WFDEI	0.5°	Daily	Global	1979–2016

JRA55 (Kobayashi et al., 2015) is a daily 3-hourly and 6-hourly global reanalysis data extending back to 1958 using four-dimensional variational data assimilation with variational bias correction for satellite radiances at a TL319L60 spatial resolution, developed by the Japan Meteorological Agency. JRA-55 overcomes some drawbacks in the first Japanese reanalysis, namely, JRA-25, including higher spatial resolution, a new radiation scheme, and introduction of greenhouse gases with time-varying concentrations.

The MERRA version 2 (MERRA-2) (Gelaro et al., 2017) is a global reanalysis data spanning from 1980 to present, produced by NASA Global Modeling and Assimilation Office. It was conducted to replace the original MERRA data set because of the advances in the assimilation system. The hyperspectral radiance and MW observations and GPS-Radio Occultation data sets are assimilated into MERRA-2 system. NASA's ozone profile observations are incorporated into MERRA-2. MERRA-2 is regarded the first long-term global reanalysis to assimilate aerosols and their interactions with other physical processes in the climate system.

The Water and Global Change Forcing Data (WFD) methodology applied to ERA-Interim (WFDEI) meteorological forcing data set (Weedon et al., 2014) is generated using the same methodology as the WFD by making use of the ERA-Interim reanalysis data. Some improvements are shown in precipitation, wind speed, and downward shortwave fluxes over the WFD. The WFDEI data are issued at a 3-hourly and daily temporal resolution and 0.5° spatial resolution from 1979 to 2016 at a global scale.

Table 1 lists a summary of the data used in this study, including the gauge-based, satellite-estimated, and reanalysis precipitation data sets. These data are collected from 2003 to 2016 and are regressed to 0.5° spatial resolution. The data set that has spatial resolution integer multiples of 0.5° (e.g., 0.25°) is matched to 0.5° by averaging the grid cells within a 0.5° grid. The data set that does not have spatial resolution integer multiples of 0.5° is regressed to 0.5° by bilinear interpolation. There are a total of 13 kinds of monthly data sets and 11 daily data sets. Among the satellite-based products, CHIRPS incorporates numerous gauge stations, and the rest of them are pure satellite-estimated data.

3. Methods

3.1. Overall Framework

A new framework is designed to merge gauge-based, satellite-estimated, and reanalysis precipitation products (Figure 2). The uncertainties of individual precipitation data are first estimated using the generalized TCH method. A rescaling procedure is needed to ensure that different

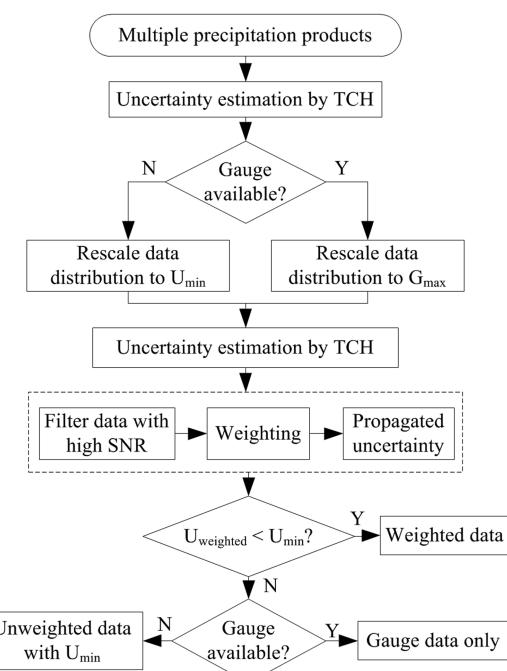


Figure 2. The designed framework for merging multiple precipitation products.

precipitation data have the same climatology and distribution before weighting. If there are gauges available in a grid cell, all of its precipitation data are rescaled to the distribution of the gauged data using empirical quantile mapping. Quantile mapping can correct all the moments of the precipitation time series between two data sets, including the first-order and second-order errors. We performed quantile mapping on the whole precipitation time series. This processing procedure is consistent for all the individual precipitation products before weighting. The quantile mapping unification is used only to generate the new merged data set together with the TCH method. Therefore, the validation is performed on the individual precipitation data without unification and on the merged data. This is reasonable because the individual precipitation data used for validation are not unified, which ensures a consistent comparison between these precipitation products. If gauges are not available, the precipitation data are rescaled to the data with the minimum uncertainty. In this case, the TCH-based merging method may not intrinsically converge to the best minimum system bias in ungauged areas. The TCH method is utilized again to estimate the uncertainties of the rescaled data. Then, a filtering procedure is conducted to filter out the data with low signal-to-noise ratio (SNR). The filtered data are weighted based on the inverse of the estimated variance-covariance matrix to produce a weighted data set. The detailed descriptions of TCH algorithm, data filtering, and weighting are introduced in the following subsections.

The monthly and daily data sets are merged individually to test the performance of the TCH merging method. The aggregation of daily merged data may not equal to the merged monthly data set if no further adjustment procedures are applied. A further adjustment by adjusting the aggregated daily data to monthly data is commonly used to ensure the consistency between the two in operational released precipitation product. This procedure should also be conducted to ensure the consistency between the merged daily and monthly data sets in our work. We adjust the daily merged data to the merged monthly data by a multiplicative factor. In this way, the consistency can be assured between daily and monthly merged precipitation data sets.

A test experiment is performed to examine the performance of TCH method without using gauge-based data. The satellite (PERSIANN, CMORPH, TRMM 3B42RT, and GSMap) and reanalysis data (ERA-Interim, JRA55, MERRA-2, and WFDEI) are merged into a new data set and is validated using gauged data. The CHIRPS data are not used in the merging process because it incorporates many in situ gauges. The validation data are chosen from GPCC, CRU TS, UDEL, CPC, and CHIRPS based on the minimal uncertainty rule over global gauged areas. The gauged areas are determined based on the available gauges from GPCC, CRU TS, and CPC as the gauges in the UDEL and CHIRPS data are not available. If the available gauges in GPCC, CRU TS, and CPC are equal or larger than 1 in a grid, this grid is regarded as gauged areas. The gauged areas are selected as the validation areas if the one of the gauged data (GPCC, CRU TS, UDEL, CPC, and CHIRPS) has the minimal uncertainty among all the 13 kinds of precipitation data (Table 1). This selection ensures that the validation data have the minimal uncertainty (Figures S1 and S2 in the supporting information) among all the individual precipitation products with gauge coverage.

The spatial patterns of the difference of precipitation trends between involved precipitation products and two references (CPC and GPCC) are demonstrated to examine how much the individual and weighted precipitation data sets can reproduce the regional precipitation trends. The trends are calculated using a linear fitting method based on the monthly precipitation products. GPCC is selected as the reference of precipitation in parts of Europe and CPC is chosen as the representative in parts of North America because the two kinds of data have dense ground stations in corresponding areas. It should be noted that both densely and sparsely gauged areas are demonstrated for spatial representation. If the trend difference is near zero, the data set is assumed to be much consistent with the reference.

3.2. Generalized TCH Method

TCH (Premoli & Tavella, 1993) is a method to derive the uncertainty (random error) of a specific variable without the use of true value over an area if at least three reference data sets are available. Although the TCH method is proposed years ago, it is commonly used to measure the instability of clocks. Until recently, it is examined in climate and hydrological applications (Ferreira et al., 2016; Long et al., 2014). To date, no research is found in weighting using this method. Therefore, this study can be regarded as a preliminary attempt to combine multiple precipitation products based on TCH. The successful application of this method will be promising for data fusion and uncertainty reduction. The TCH method is commonly applied using

three data sets, and it is extended to arbitrary number of data (≥ 3) here, that is, the generalized TCH. Although Ferreira et al. (2016) demonstrated the principle of TCH, only the fixed four kinds of data were used and Allan's variance was obtained instead of sample variance. Here the implementation of the generalized TCH method is different from Ferreira et al. (2016) and is extended to incorporate multiple data sets based on sample variance. Consider the precipitation series to be $\{P_i, i = 1, 2, \dots, N\}$; they can be generally expressed as

$$P_i = P_{\text{true}} + \varepsilon_i, \quad (1)$$

where P_{true} denotes the true precipitation in a specific area and is unknown; ε_i is the measurement error. Choosing an arbitrary data set as the reference, the differences of $N - 1$ precipitation products and the reference data can be calculated.

$$D_{iN} = P_i - P_N = \varepsilon_i - \varepsilon_N, \quad i = 1, \dots, N - 1, \quad (2)$$

where P_N is the reference series and the TCH result is independent of the selection of reference data (Ferreira et al., 2016). These $N - 1$ differences can be concatenated into an $M \times (N - 1)$ matrix

$$D = [D_{1N} \ D_{2N} \ \dots \ D_{(N-1)N}], \quad (3)$$

where the rows of D denote the differenced time series with M length. The variance/covariance of D is then obtained by

$$S = \left[\frac{1}{M-1} \right] \left[(D - \bar{D})^T (D - \bar{D}) \right], \quad (4)$$

where \bar{D} represents the average of D . The covariance of D can be related to S by

$$S = K^T \cdot R \cdot K, \quad K = \begin{bmatrix} I \\ -u^T \end{bmatrix}, \quad (5)$$

where R is the $N \times N$ covariance matrix of $\{\varepsilon_i, i = 1, 2, \dots, N\}$ and u represents the vector $[1 \dots 1]^T$. For a detailed deduction process, one can refer to Premoli and Tavella (1993). The covariance matrix of D is regarded as the Allan covariance when TCH was initially proposed to evaluate the instability of clocks, which requires a filtering operation on the original time series before differencing. However, the covariance of D , that is, the R matrix, does not necessarily have to be the Allan covariance but can be the common sample variance, as used in some studies (Liu et al., 2013; Rieckh & Anthes, 2018). Equation 5 can be reformatted as

$$S = [I - u] [\hat{R} rr^T r_{NN}] \begin{bmatrix} I \\ -u^T \end{bmatrix}, \quad (6)$$

where \hat{R} is the $(N - 1) \times (N - 1)$ matrix and r is the $(N - 1)$ vector $[r_{1N} \ r_{2N} \ \dots \ r_{N-1,N}]$ grouping covariance estimates with the N th time series and r_{NN} denotes the variance of N th time series. This partitioning can help solve the underdetermined problem in equation 5 by isolating the N free parameters (r and r_{NN}). When the free parameters are determined, the unknown elements in \hat{R} is obtained by

$$\hat{R} = S - r_{NN} [uu^T] + ur^T + ru^T. \quad (7)$$

A solution to N free parameters requires a suitable objective function, ensuring the positive definiteness of R . Tavella and Premoli (1994) suggested a constraint $\det(R) > 0$ to keep the validity of the free parameters in the solution domain. This constraint may not be sufficient to ensure the unique solution of R . Then, an optimal criterion was proposed by Premoli and Tavella (1993) to minimize the global correlation between the noises of time series while maintaining the positive definiteness of R . A suitable objective function is suggested by

Galindo and Palacio (1999) based on the Kuhn-Tucker theorem to minimize the quadratic mean of covariances

$$F(\mathbf{r}, \mathbf{r}_{NN}) = \sum_{i < j} \frac{r_{ij}^2}{L^2}, \quad (8)$$

subjecting to a constraint

$$G(\mathbf{r}, \mathbf{r}_{NN}) = -\frac{\mathbf{r}_{NN} - [\mathbf{r} - \mathbf{r}_{NN}\mathbf{u}]^T \cdot \mathbf{S}^{-1} \cdot [\mathbf{r} - \mathbf{r}_{NN}\mathbf{u}]}{L} < 0, \quad (9)$$

where $L = \sqrt[n-1]{\det(\mathbf{S})}$. The initial conditions are selected as follows to meet the constraint (Torcaso et al., 1998).

$$r_{iN}^{(0)} = 0, i < N \text{ and } r_{NN}^{(0)} = (2 \cdot \mathbf{u}^T \cdot \mathbf{S}^{-1} \cdot \mathbf{u})^{-1}. \quad (10)$$

Once the free parameters are determined, the unknown elements in R can be solved using equation 7.

The gauge-based, satellite-estimated, and reanalysis data are collected for each grid. The truth is estimated using all these data sets but is unknown. In high-latitude areas, only gauge-based and reanalysis data are used in the analysis due to a limited coverage of satellite data. The 13 monthly precipitation products and the 11 daily products from a variety of sources are used to generate a weighted precipitation data based on the generalized TCH method and are validated based on gauged areas. Apart from the sensor information content, retrieval, and interpolation algorithms, the differences of climatology in different precipitation products may also affect the relative uncertainty and the merging result. The climatology of these precipitation products is unified by quantile mapping. The experiments are conducted from 2003 to 2016 as an illustration. This time period is chosen because all the precipitation data cover this time range. Earlier time periods can also be selected using part of these data.

3.3. Data Filtering and Weighting

The precipitation data with high SNR should be merged, and the noisy data with low SNR should be excluded. The variance of a precipitation product can be decomposed by the sum of the variance of the true signal and the variance of random noise according to equation 1.

$$\sigma_i^2 = \sigma_\Theta^2 + \sigma_{\varepsilon_i}^2 + 2\sigma_{\Theta\varepsilon_i}, \quad (11)$$

where σ_Θ^2 and $\sigma_{\varepsilon_i}^2$ are the variances for the true signal and the noise, respectively. $\sigma_{\Theta\varepsilon_i}$ is the error orthogonality. Here the error is assumed to be orthogonal, that is, $\sigma_{\Theta\varepsilon_i} = 0$. Although this assumption may not hold, the impact of the non-orthogonality is damped by rescaling or even compensated by the potential similar magnitudes of error orthogonality among multiple data sets (Yilmaz & Crow, 2014).

The SNR is defined as the ratio of the variance of the signal to the variance of the noise.

$$\text{SNR} = \frac{\sigma_\Theta^2}{\sigma_{\varepsilon_i}^2}. \quad (12)$$

The linear correlation coefficient between the underlying true signal and the precipitation product is proposed by McColl et al. (2014):

$$R_i^2 = \frac{\sigma_\Theta^2}{\sigma_\Theta^2 + \sigma_{\varepsilon_i}^2} = \frac{\text{SNR}_i}{1 + \text{SNR}_i}, \quad (13)$$

where R_i^2 is the coefficient of determination.

The SNR can be expressed in its decibel units (dB):

$$\text{SNR[dB]} = 10\log(\text{SNR}) = 10\log\left(\frac{\sigma_\Theta^2}{\sigma_\epsilon^2}\right). \quad (14)$$

When the signal variance equals to the noise variance, R_i^2 equals to 0.5 and SNR[dB] is zero. A 3-dB SNR means the signal variance is two times of the noise variance. The precipitation data set with high noise usually has a low SNR. Therefore, the SNR threshold 0.5 is used to filter out the data with high noise in the merging process.

The weights used to merge the individual precipitation products are set as the inverse of the error variance-covariance of TCH estimated uncertainty by the considering of error correlation. According to the Gauss-Markov theorem, a weighted average that has the minimal variance is expressed as

$$W = C^{-1}, \quad (15)$$

$$P_{\text{weighted}} = (J^T W J)^{-1} (J^T W P), \quad (16)$$

where C denotes the error variance-covariance matrix obtained by TCH and W is the weight matrix. J is the design matrix and is a vector of ones, that is, $[1, \dots, 1]^T$.

According to equation 11, the covariance of two precipitation data sets is expressed as

$$\sigma_{ij} = \sigma_\Theta^2 + \sigma_{\Theta\epsilon_i} + \sigma_{\Theta\epsilon_j} + \sigma_{\epsilon_i\epsilon_j}, \quad (17)$$

where $\sigma_{\epsilon_i\epsilon_j}$ is the error cross-correlation. Ignoring this may cause an underestimation of error covariance. Equation 8 aims to minimize the quadratic mean of covariances, thus minimizing the error cross-correlation term in some degree. In the merging process, the correlation of error variance needs to be considered to calculate the propagated uncertainty. As a number of precipitation data sets are involved, their errors are inevitably correlated due to the possibility of the commonly used gauges in gauged data, the same satellite sensor types for precipitation retrieval, and the same data assimilation algorithms in reanalysis data. The weight matrix in equation 15 takes the error correlation into consideration to merge the individual data sets into a new data with minimal propagated error variance.

3.4. Evaluation Metrics

Four statistical measures, namely, Pearson's correlation coefficient (PCC), bias, root mean square error (RMSE), and standard deviation (SD), are used to evaluate the performance of TCH weighting using gauged data. PCC measures the linear correlation between two variables.

$$\text{PCC} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (18)$$

where x and y are individual time series; \bar{x} is the arithmetic average of x and is given by $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, and \bar{y} is the arithmetic average of y and is given by $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

Bias is the difference of the expected values between the truth and estimated one.

$$\text{Bias} = \bar{y}_i - \hat{y}_i, \quad (19)$$

where \bar{y}_i and \hat{y}_i denote the mean of y_i (the truth) and \hat{y}_i (estimated value), respectively.

The RMSE measures the difference between forecasted value and true value.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (20)$$

where y_i is the true value or observation and \hat{y}_i is the forecasted value; n is the sample size. The SD statistic measures the dispersion of a data set.

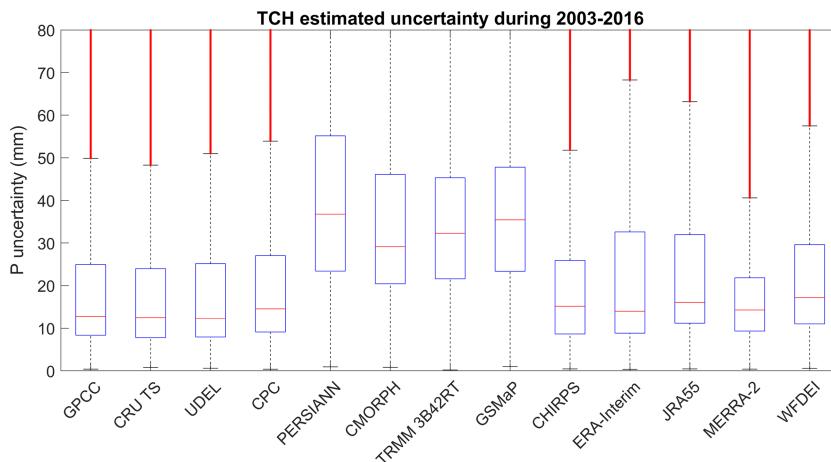


Figure 3. The uncertainty of individual precipitation products based on the generalized TCH method during 2003–2016. The red horizontal line inside each box denotes the median, and the upper and lower margins of the box represent the upper quartile (Q_3) and lower quartile (Q_1), respectively. The upper whisker indicates the upper limit ($Q_3 + 1.5[Q_3 - Q_1]$), and the lower whisker indicates the lower limit ($Q_1 - 1.5[Q_3 - Q_1]$). The red dots outside the whiskers represent outliers.

$$SD = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{n}}. \quad (21)$$

4. Results

4.1. Individual Precipitation Data Uncertainty and Weights

Figure 3 demonstrates the distribution of uncertainties for the raw individual precipitation products (no rescaling) estimated using the generalized TCH method during 2003–2016. The uncertainties of individual precipitation products are estimated based on all the gauge-based, satellite-based, and reanalysis data. The relative uncertainty is different in the involved precipitation products. PERSIANN, CMORPH, TRMM 3B42RT, and GSMap have much higher uncertainty than the others (Figure 3). GPCC, CRU TS, UDEL, and MERRA-2 have lower uncertainty relative to other individual precipitation data. As CHIRPS are corrected by gauges, the other four satellite-based products have larger random error than gauge-based and reanalysis data. It should be noted that the uncertainty estimation is referred as random error, as the first-order systematic bias cannot be obtained without the knowledge of the truth. Therefore, the estimated uncertainty based on TCH should be less than the uncertainty estimated by the truth (first-order plus second-order errors).

A spatial map of the TCH estimated uncertainty during 2003–2016 is shown in Figure 4. The uncertainty is generally larger in the tropical areas than extratropical places (Figure 4), where the precipitation climatology has more rainfall in the former than the latter. Substantial uncertainty is seen in Malay Archipelago in almost all the individual precipitation data. Large uncertainty exists in the northern areas of South America in all the precipitation data. In Australia, GPCC, UDEL, CPC, CHIRPS, and MERRA-2 have smaller uncertainty than other individual precipitation data. CRU TS and CHIRPS improve the precipitation estimation in Central Africa over other precipitation products.

The weights of individual precipitation data in the merging process during 2003–2016 are displayed in Figure 5. The weights can be used to infer the relative uncertainty of individual precipitation data, similar to that in Figure 4. In North America, MERRA-2, CHIRPS, and ERA-Interim have larger weights in parts of this continent than other precipitation data. In South America, CHIRPS and MERRA-2 are assigned larger weights. In Africa, CHIRPS shows better performance than the others over many areas, followed by CRU TS. MERRA-2, GPCC, CHIRPS, and CRU TS have larger weights in parts of Europe and Asia than other data sets. MERRA-2, GPCC, and CHIRPS exhibit higher weights in Australia than other data sets.

4.2. TCH Weighted Precipitation

A total of 4,858 gauged grids are used for validation of monthly data, of which 17% are GPCC, 12% for CRU TS, 9% for UDEL, 7% for CPC, and 55% for CHIRPS. The correlation coefficient, Bias, RMSE, and SD metrics

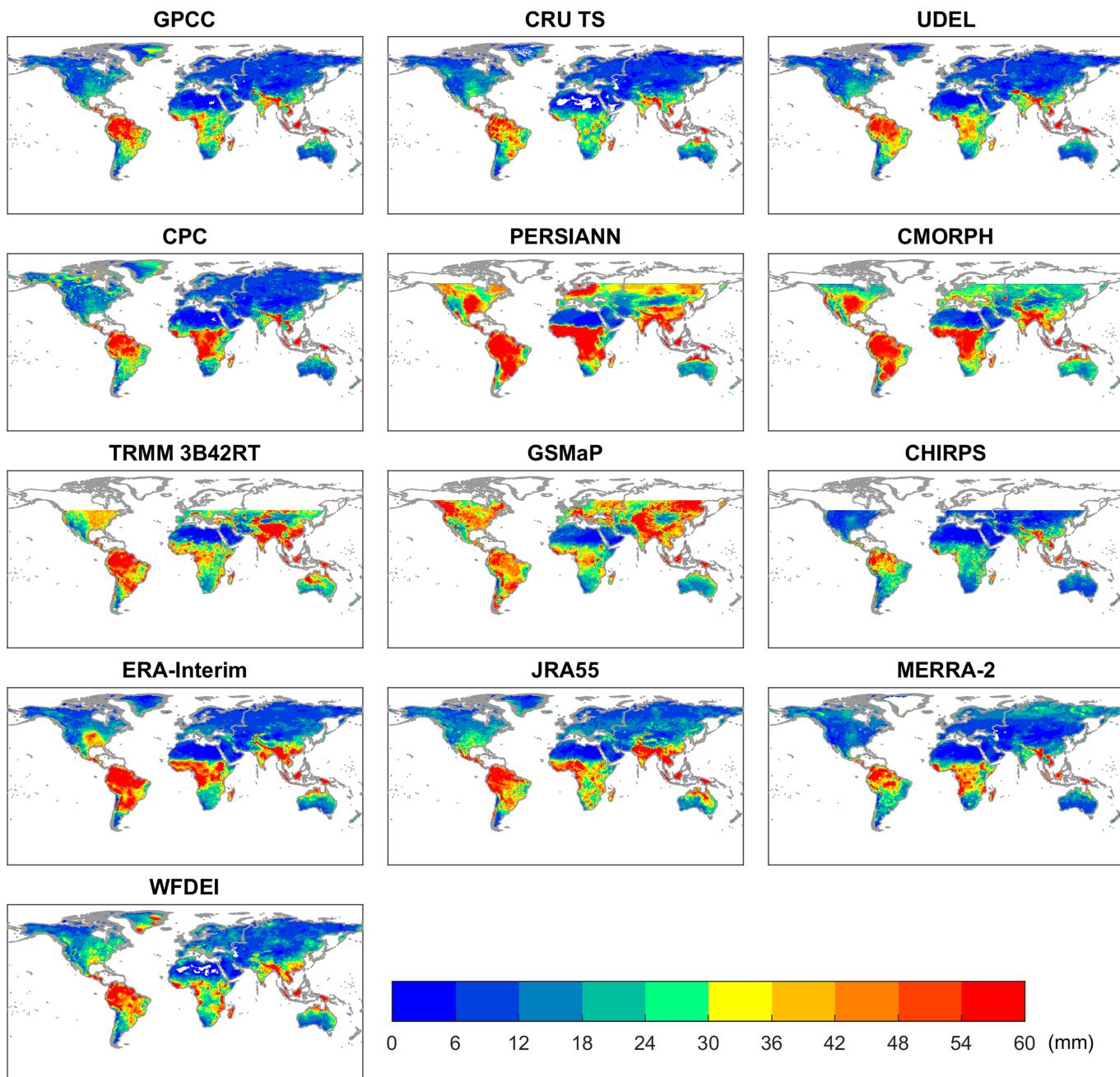


Figure 4. The spatial map of the uncertainty of individual precipitation products during 2003–2016 based on the generalized TCH method.

are shown in Figure 6 based on gauged validation. ERA-Interim, JRA55, MERRA-2, and WFDEI exhibit better overall performance than other individual precipitation products in terms of these statistics, of which MERRA-2 has the minimum RMSE. TCH achieves comparable performance to the best involved individual data sets during this time, with a median PCC of 0.94 and a median RMSE of 15 mm, slightly better than MERRA-2 in terms of Bias, RMSE, and SD. The reason that TCH does not show a substantial improvement is due to several facts. First, the satellite-based data sets have substantial uncertainties compared to reanalysis data and may not contribute much to the merging performance. On the other hand, the TCH method requires relatively independent data sets to achieve stable results. However, no gauged data sets are used. Therefore, the errors in the involved data may be strongly correlated, which influences the merging results. If all the 13 kinds of data are included, the merging results exhibit

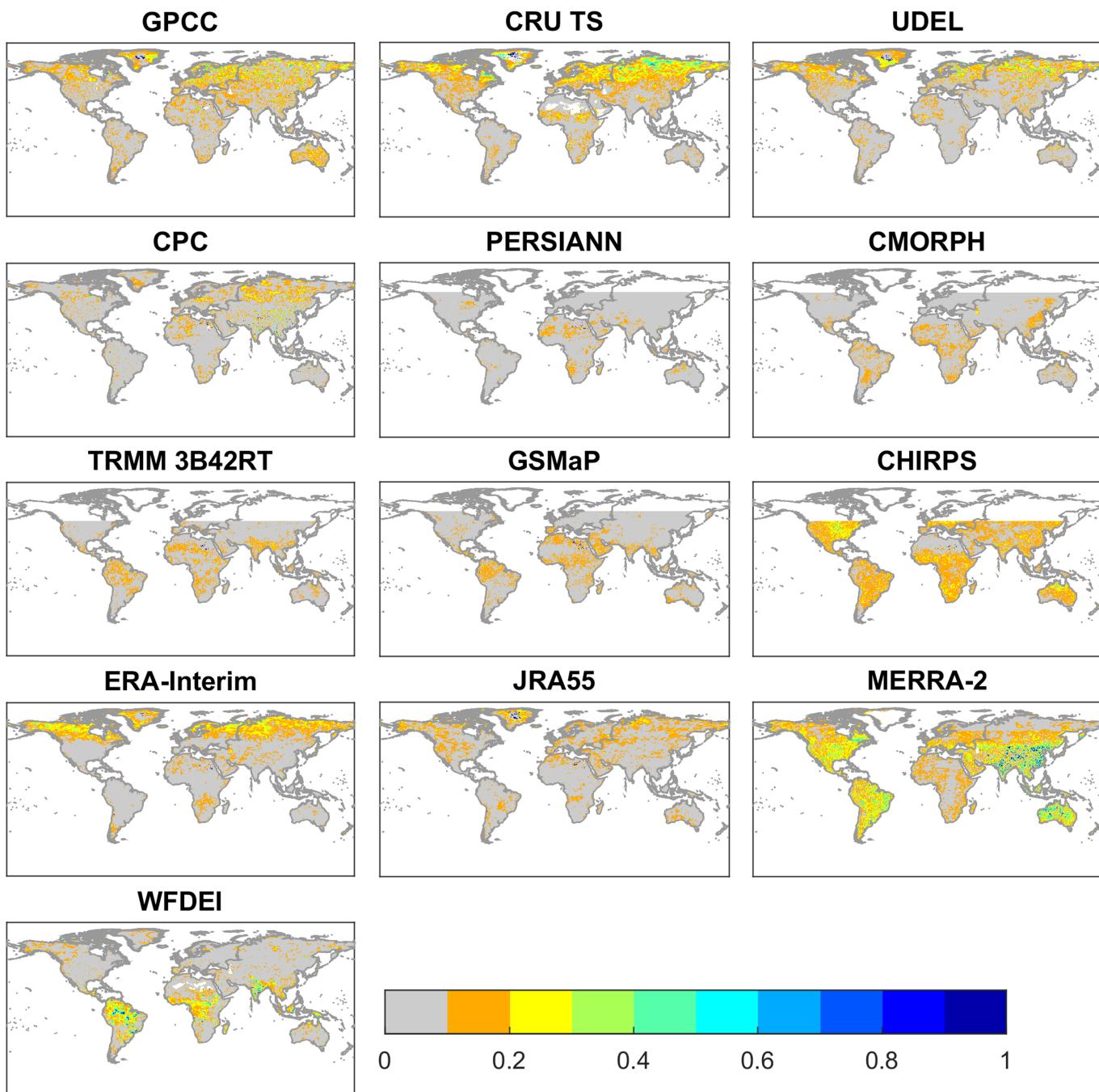


Figure 5. The weights of TCH based on all the available monthly precipitation products during 2003–2016.

substantial improvement than MERRA-2 (Figures 7 and S3 to S5). It should be noted that the gauged data are used for validation and thus the calculated statistics for gauged data are only demonstrated (Figures S3 to S5) for visualization (not for comparison).

The spatial maps of the RMSE of individual and weighted precipitation data sets in parts of Europe and Australia are shown in Figures 8 and 9, respectively. The TCH weighted precipitation indicates a comparable performance with ERA-Interim and MERRA-2 in parts of Europe and Australia, much better than the satellite-based data in PERSIANN, CMORPH, TRMM 3B42RT, and GSMaP. The weighted data set exhibits similar performance with MERRA-2, with equal or lower random error in the validated areas.

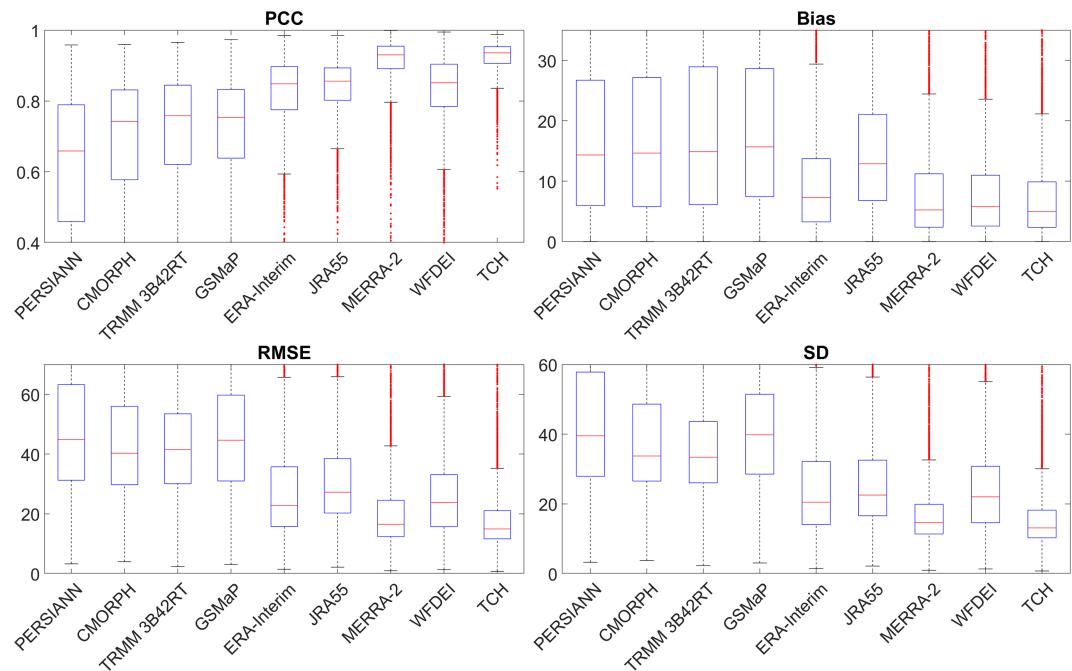


Figure 6. The performance of TCH weighted monthly precipitation (without gauged data) using gauged areas for validation. Four metrics are calculated: (top left) PCC, (top right) Bias (mm), (bottom left) RMSE (mm), and (bottom right) SD (mm). The bias statistic is shown in its absolute value. The boxplots are truncated to better demonstrate their differences.

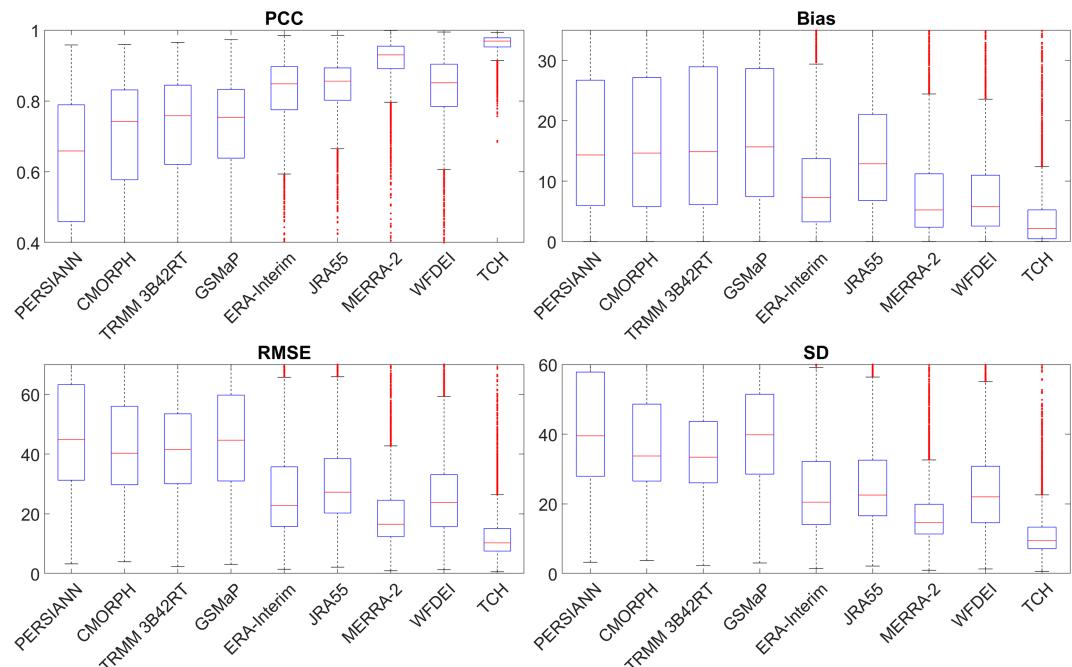


Figure 7. The performance of TCH weighted monthly precipitation (with gauged data) using gauged areas for validation. Four metrics are calculated: (top left) PCC, (top right) Bias (mm), (bottom left) RMSE (mm), and (bottom right) SD (mm). The bias statistic is shown in its absolute value. The boxplots are truncated to better demonstrate their differences.

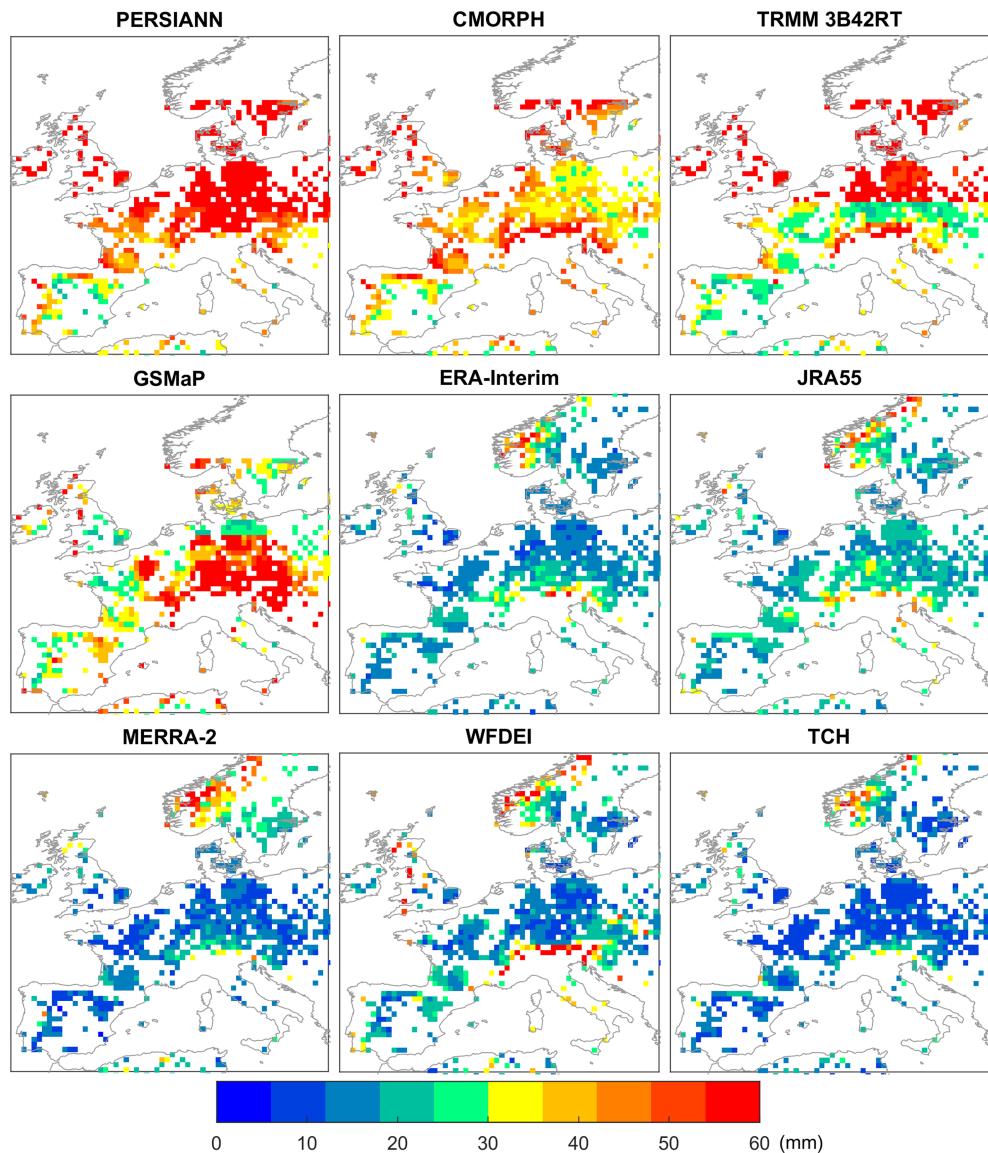


Figure 8. The spatial map of RMSE of the TCH weighted monthly precipitation and individual precipitation data sets (without gauged data) using gauged areas for validation in parts of Europe during 2003–2016.

The TCH merged data are compared with the multisatellite-retrieved GPM IMERG data in the validation areas. It can be seen that a substantial improvement is demonstrated in the TCH merged data over the GPM IMERG data (Figure 10), suggesting an advantage of the TCH merged data compared to state-of-the-art integrated precipitation data. The advantage is significant both at monthly and daily scales.

4.3. Spatiotemporal Patterns of Interannual Precipitation

The TCH weighted data set exhibits good performance compared with other individual products as the difference with references in the weighted data set is overall comparable to the best individual data, or even lower, both in parts of Europe (Figure 11) and in parts of North America (Figure 12).

The magnitudes of annual and seasonal precipitation in the weighted data set are generally consistent with CPC in parts of North America (Figure 13), although some differences exist. The weighted precipitation data avoid the extremely low or extremely high values exhibited in some data sets like PERSIANN and CMORPH and can largely reproduce the seasonal variations in CPC.

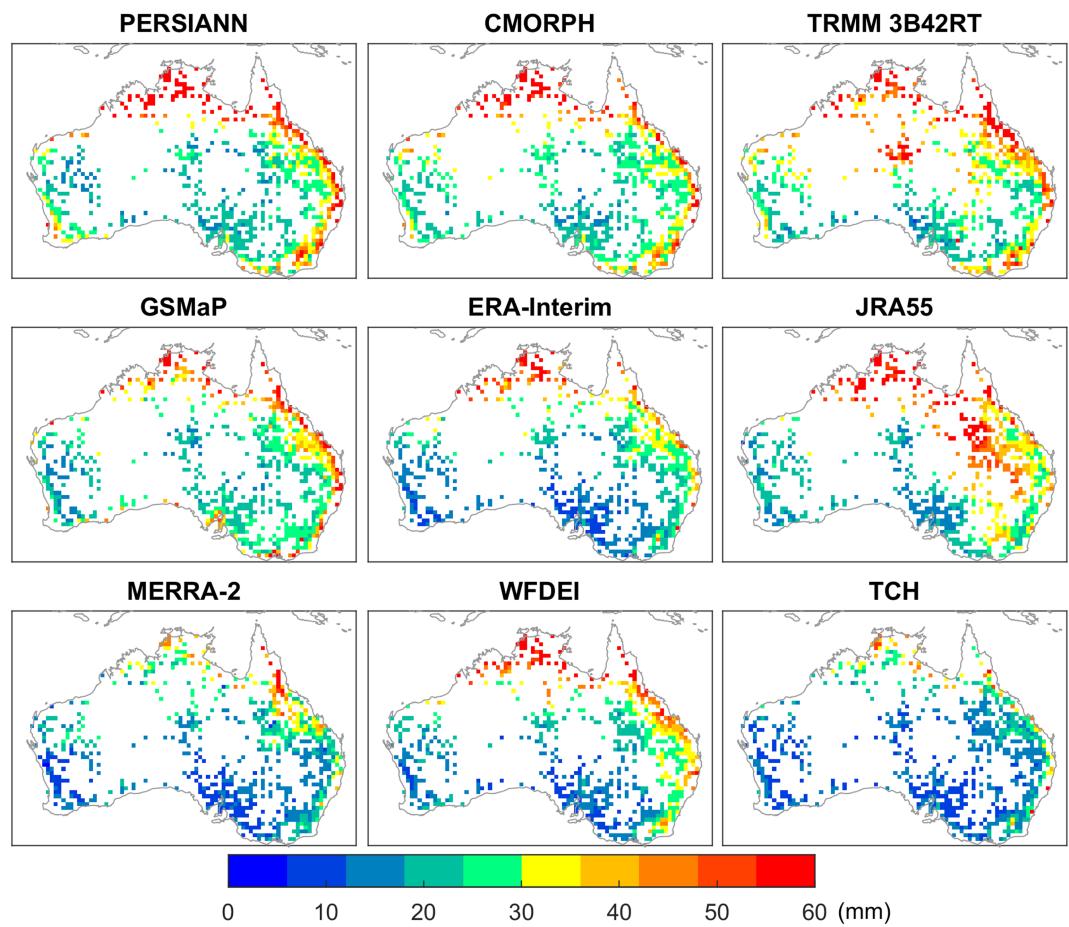


Figure 9. The spatial map of RMSE of the TCH weighted monthly precipitation (without gauged data) and individual precipitation data sets using gauged areas for validation in Australia during 2003–2016.

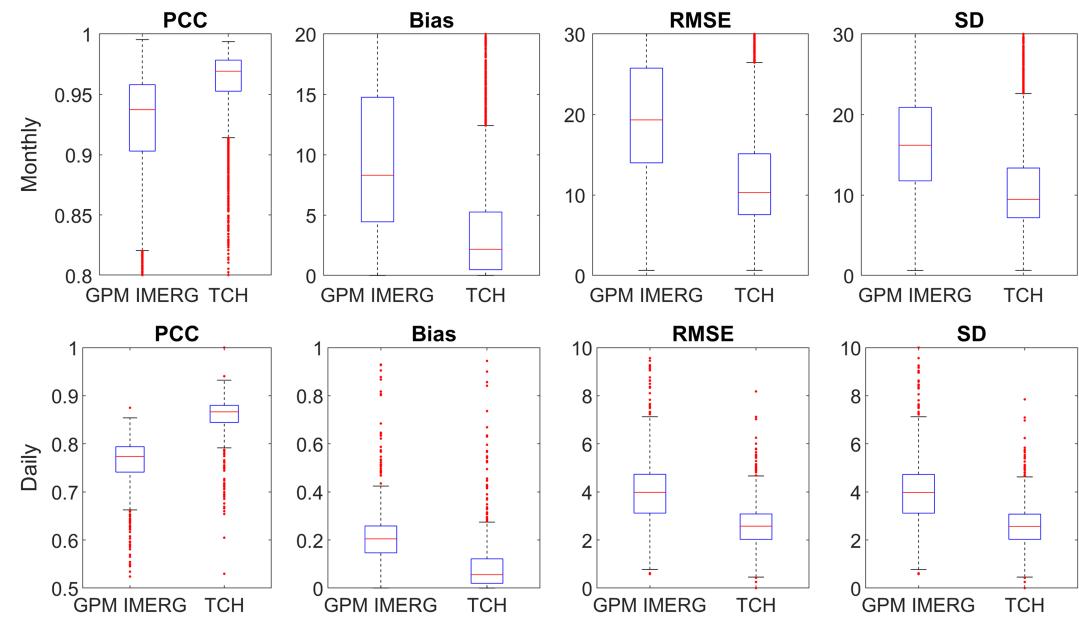


Figure 10. A comparison of the TCH weighted precipitation data set with GPM IMERG precipitation data at monthly and daily scales. Four metrics are calculated: (first column) PCC, (second column) Bias (mm), (third column) RMSE (mm), and (last column) SD (mm).

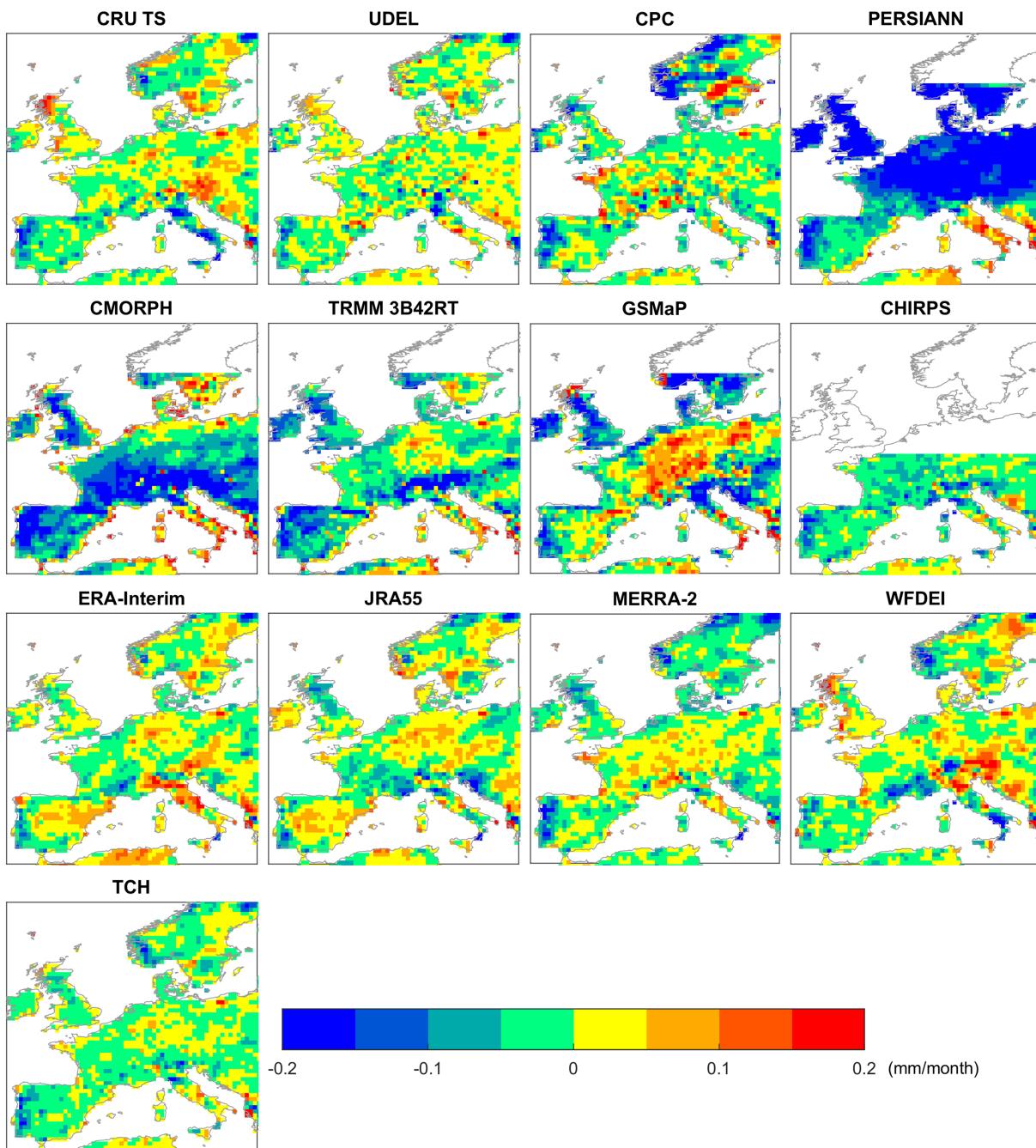


Figure 11. Precipitation trend difference between involved data sets and GPCC in parts of Europe during 2003–2016.

4.4. Daily Data Merging

The 11 daily precipitation products are merged into a new daily precipitation data using the same approach with that of monthly data. As daily precipitation data are commonly used in hydrological simulations, the daily data merge experiment is conducted to demonstrate the validity of the TCH merging approach. A large improvement is seen in the weighted daily precipitation data over the individual precipitation data sets in terms of the overall correlation, RMSE, and SD (Figure 14). The Bias measurement also suggests an improvement over the majority of the individual data. When the gauge-based data are included, a substantial improvement can be seen in the weighted data set over the satellite-based and reanalysis products (Figures 15 and S6).

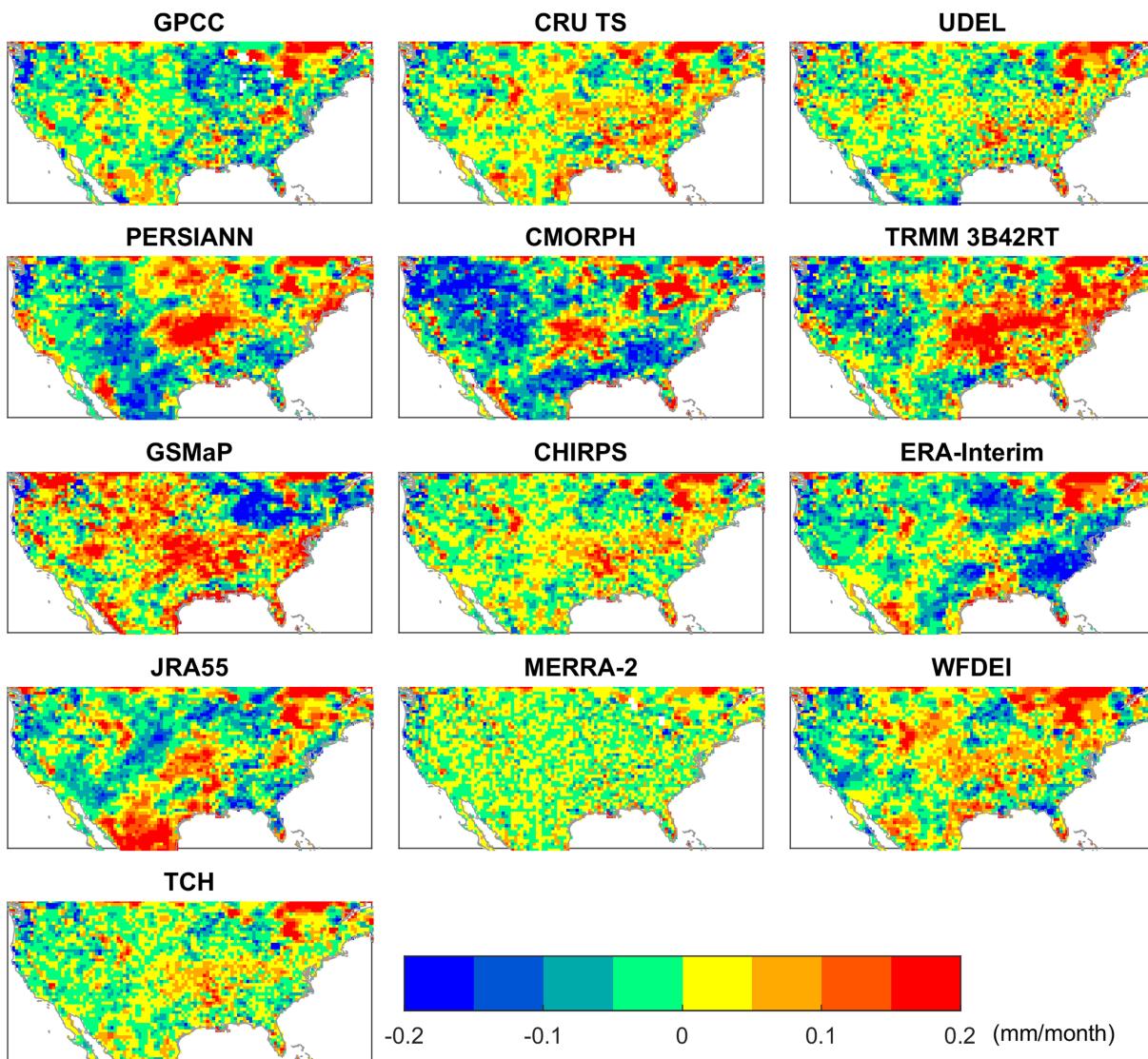


Figure 12. Precipitation trend difference between involved data sets and CPC in parts of North America during 2003–2016.

The weights in the daily data merging (Figure 16) has some similarities and differences with that of monthly data merge. The MERRA-2 data are assigned large weights in many areas except the Africa. In Africa, PERSIANN, GSMAp, and CHIRPS have larger weights than the reanalysis data and gauge-based data, which may be a result of limited in situ data in these areas. The individual precipitation data contributes differently to the weight in different areas, which is also a reflection of its uncertainty (Figure S7).

4.5. Comparison with Other Merging Methods

Another two optimally merging methods, namely, arithmetic mean and one-outlier-removed (OOR) mean, are used to compare with the TCH merging method. The arithmetic mean and OOR mean methods are used in a previous study (Shen et al., 2014) and show effective merging result. Therefore, it is useful to see if the TCH merging method outperforms the previously used methods. The monthly and daily merged data sets without gauge-based data by TCH and the two previously used methods are compared. The monthly and daily results are shown in Figures 17 and 18, respectively. It is seen that the TCH merging results outperform the arithmetic mean and OOR mean methods in terms of Bias, RMSE, and SD measurements at both monthly and daily scales. The advantage of the TCH-based merging method lies in the uncertainty quantification of individual precipitation products, while the arithmetic and OOR mean methods are unweighted average without taking the uncertainty of individual data into consideration.

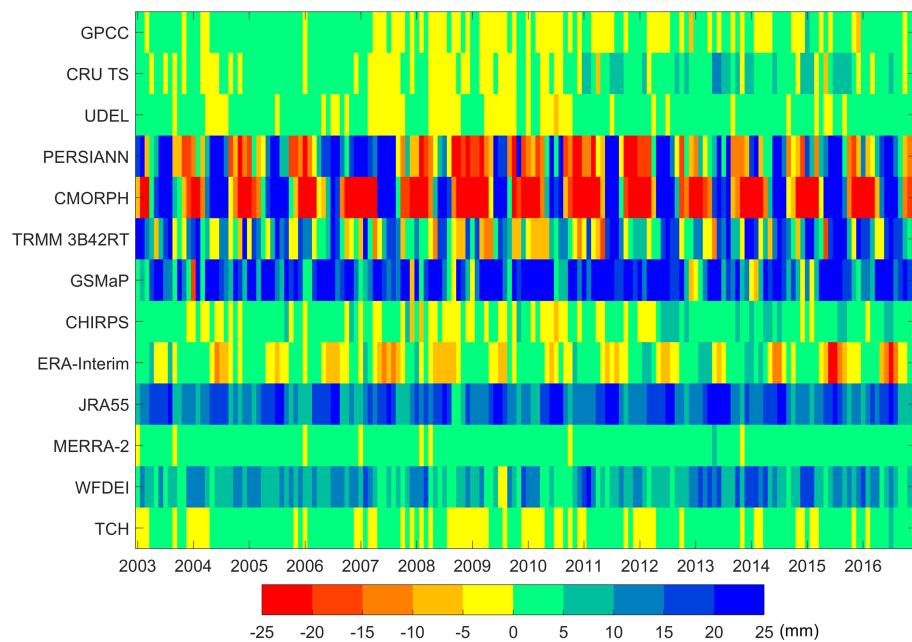


Figure 13. Temporal precipitation pattern of the involved data sets in parts of North America during 2003–2016.

5. Discussion

5.1. TCH Uncertainty Estimation

The generalized TCH method can be used to derive the relative uncertainty of individual precipitation products. Although the TCH estimated uncertainty may underestimate the absolute uncertainty as compared to gauge-based validation (Figure S8), the relative uncertainty among individual precipitation products can be well retained. The underestimation of the uncertainty based on the TCH method is probably due to the error

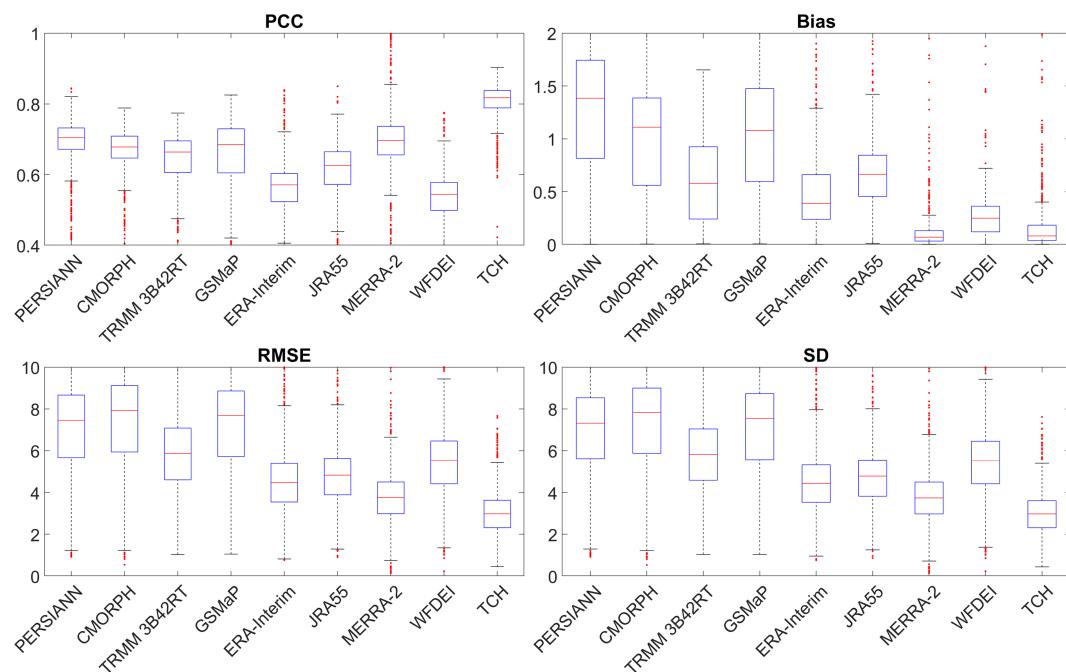


Figure 14. The performance of daily TCH weighted precipitation (without gauged data) using gauged areas for validation. Four metrics are calculated: (top left) PCC, (top right) Bias (mm), (bottom left) RMSE (mm), and (bottom right) SD (mm). The bias statistic is shown in its absolute value. The boxplots are truncated to better demonstrate their differences.

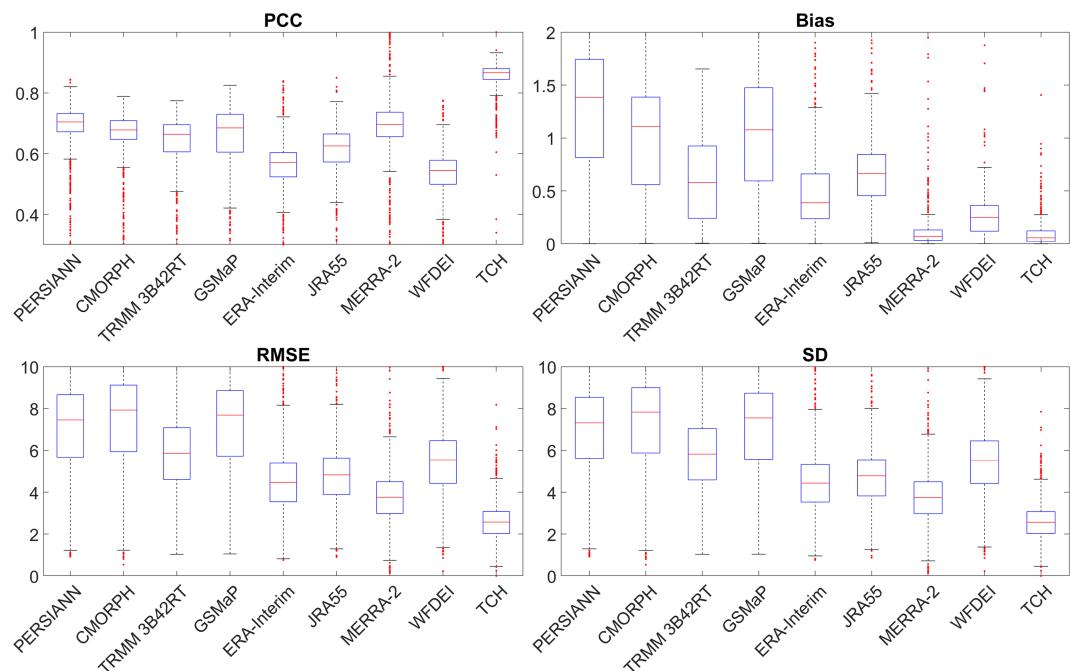


Figure 15. The performance of daily TCH weighted precipitation (with gauged data) using gauged areas for validation. Four metrics are calculated: (top left) PCC, (top right) Bias (mm), (bottom left) RMSE (mm), and (bottom right) SD (mm). The bias statistic is shown in its absolute value. The boxplots are truncated to better demonstrate their differences.

of representativeness (O'Carroll et al., 2008; Swinbank et al., 2012), which indicates the fact that the unknown truth estimated by TCH in a grid is not equal to the observations represented by certain number of gauges. Other factors may also contribute to the underestimation of uncertainty, for example, error correlation. In this study, the 0.5° grids with the coverage of at least one rain gauges with minimal uncertainty are used for validation, but this is far away from the unknown truth for a grid cell. Precipitation is spatially heterogeneous, and a 5-km distance may cause a substantial difference of precipitation intensity. Current high-resolution precipitation products have reached an ~ 5 -km spatial resolution (e.g., CHIRPS); however, the global rain gauges relative to the number of 5-km grids is pretty small. If a 5-km grid is regarded as the truth when a rain gauge is covered, a total of 100 stations are needed for an approximately 0.5° grid, while this is not possible regarding current gauge networks. Therefore, the error of representativeness exists between the gauge-measured precipitation and the gridded precipitation. For example, rain gauges located in central districts of a city measure a strong rainfall event, while a 50-km grid including this station cannot be assumed to have the same rainfall intensity in all of its subareas.

The weighted precipitation can generally reproduce the gauge-measured precipitation with high correlation and low RMSE, based on the inverse of error variance-covariance matrix estimated by TCH. Multiple precipitation data can be selected and applied in TCH if they are generally independent with small error cross-correlation. In our experiment, the positive definiteness can be ensured and a unique solution is obtained for nearly all the locations.

The distribution of precipitation error is investigated by the one-sample Kolmogorov-Smirnov test. A value of one indicate that the null hypothesis that the data come from a normal distribution is rejected and a value of zero indicates the null hypothesis cannot be rejected at 5% significance level. The precipitation error in equation 1 often follows a skewed distribution, which is validated in gauged areas (Figure S9). However, the non-Gaussian precipitation error does not impact the uncertainty estimation as the error in the TCH estimation is not limited to Gaussian error assumption.

5.2. The Performance of Weighted Precipitation

The weighted precipitation has a smaller uncertainty than individual precipitation products validated by gauged areas (Figures 6, 7, 14, and 15). None of the individual precipitation data set is always optimal in

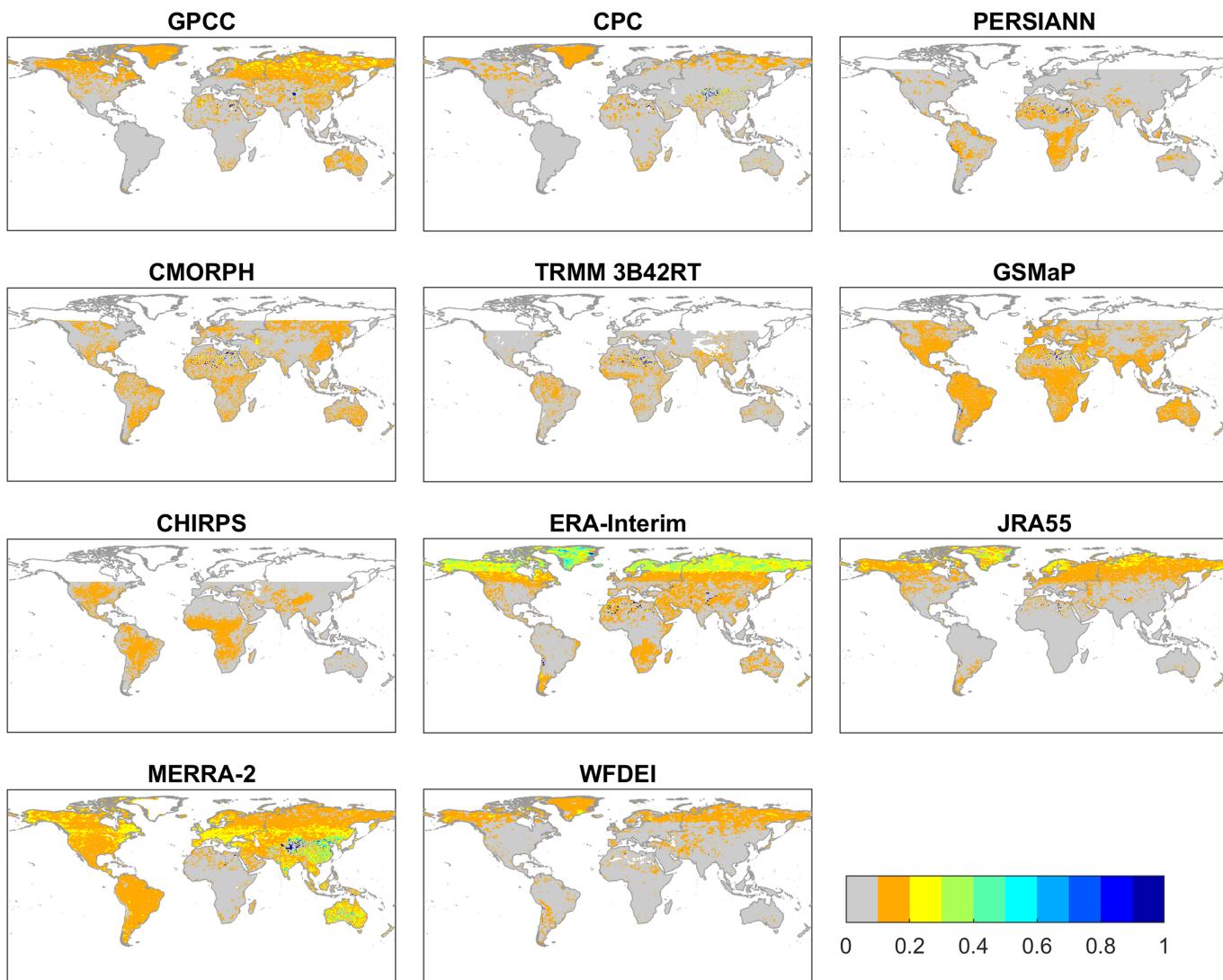


Figure 16. The weights of TCH based on the 11 daily precipitation products during 2003–2016.

global lands. The performance of individual precipitation data is regionally oriented, such as the superiority of CHIRPS data in Africa, MERRA-2 in Australia and Europe, and CRU TS in North Asia. The inverse error variance-covariance weighting can integrate precipitation information from various data sets, combining the strength of multiple data. The inverse variance-covariance weighting is impacted by outliers. Therefore, the precipitation data are filtered to keep the data with high SNR and remove abnormal observations or simulations. The climatology and data distribution of the individual precipitation data sets are rescaled to the gauged data with maximum gauge coverage when gauges are available and rescaled to the data with minimal uncertainty when no gauge exists. This step reduces the impact of climatology bias to some degree and ensures a consistent handling of second-order discrepancies between different precipitation data sets. Theoretically, the weighted average can reach the minimal error variance based on the inverse of variance-covariance matrix as weights. However, in practical applications, the weighted results may not always have minimal error variance because the uncertainty estimation in TCH are impacted by different data sets and the systematic error is not known. Therefore, the weighted precipitation data set has some underlying uncertainties.

The TCH-based merging method used in this study can not only enable the converge of between multiple precipitation products but also lead to a real improvement in terms of absolute accuracy (bias + random error) and uncertainty (refer to random error in this study) reduction. The real improvement can be

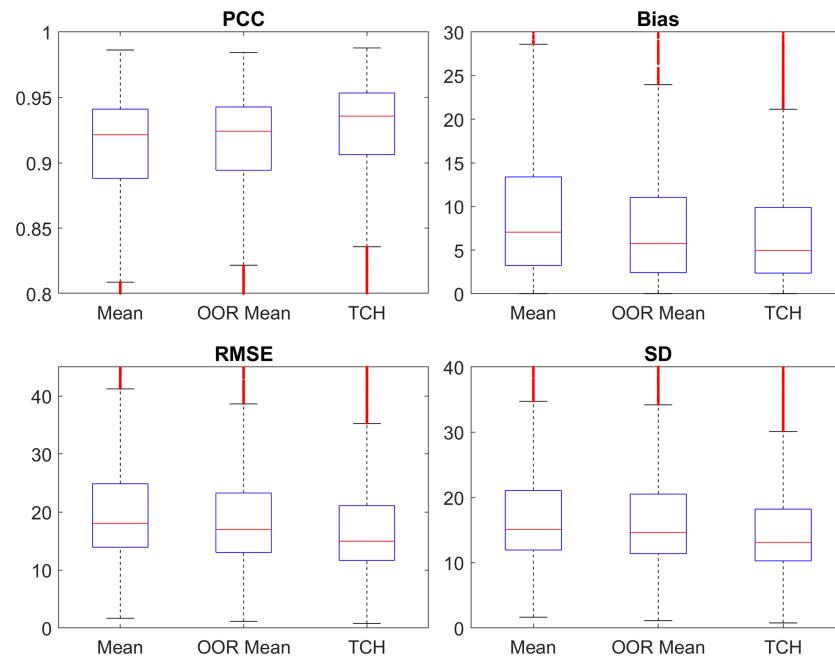


Figure 17. A comparison of the monthly merged precipitation data sets (without gauged data) by TCH and two previously used methods.

clearly seen in section 4 of this paper. Figures 6 and 14 demonstrate the performance of the weighted data without using gauge-based data and validated at gauge-based areas at monthly and daily scales, respectively. The two figures clearly show the advantage of the weighted data set in terms of PCC, RMSE, and SD metrics. The RMSE measurement intrinsically contains bias and random error and thus can be regarded as absolute accuracy. The SD measurement is considered as an evaluation of random error. It can be seen that Figures 6 and 14 both demonstrate the superiority of the weighted data set regarding

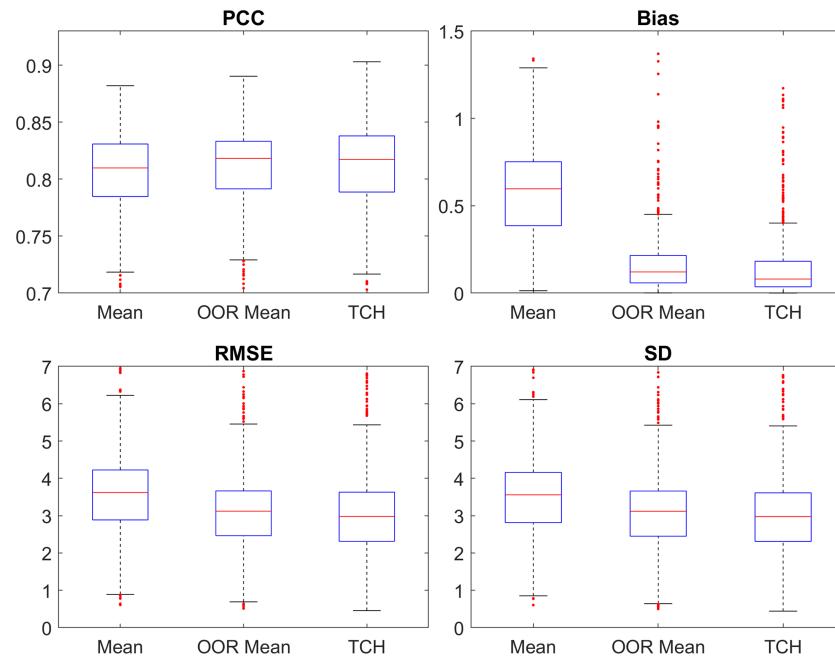


Figure 18. A comparison of the daily merged precipitation data sets (without gauged data) by TCH and two previously used methods.

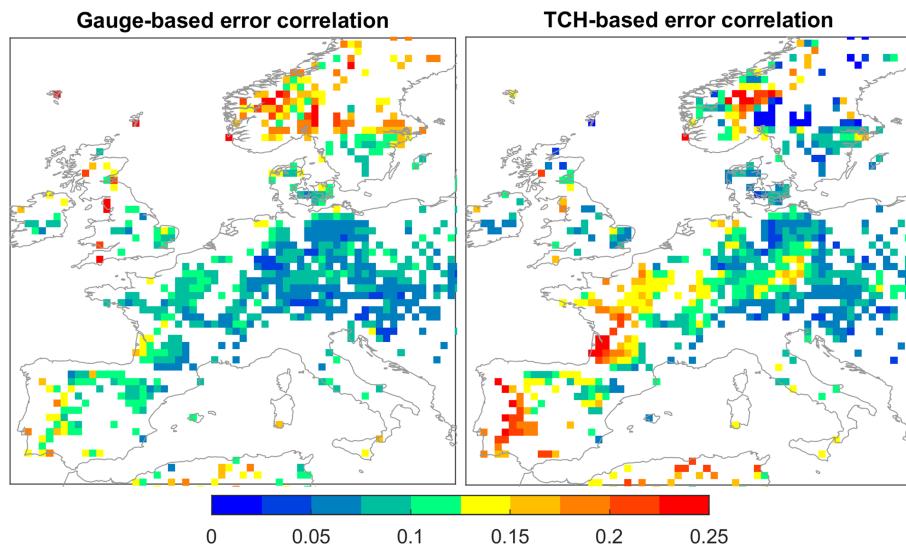


Figure 19. The error cross-correlation validated by gauged data and estimated in the TCH optimization process in parts of Europe. The monthly precipitation data are used.

RMSE and SD over the individual data sets. This conclusion can also be extended to ungauged or poorly gauged areas because no gauged data are used for merging in the results of Figures 6 and 14. The validation in ungauged or poorly gauged areas cannot be applied because no dense gauges are available. What can we do is to validate the merged data without using gauges in gauged areas (Figures 6 and 14). This is reasonable because no gauged data are used in the merging process. This is the commonly used way in merging/validating studies. It is true that the truth is unknown in the ungauged or poorly gauged areas. However, the accuracy is likely to be improved in ungauged or poorly gauged areas in a similar way as that of the gauged areas when using the TCH merging scheme, despite the spatial heterogeneity.

5.3. Error Cross-Correlation

The TCH method was first proposed under the assumption that the involved three or more data sets are independent from each other. However, the generalized TCH method estimates the error variance by an iterative optimization procedure that minimizes the global error cross-correlation. Here, the error cross-correlation ($\sum_{i=1}^n \sum_{j=1, j>i}^n \frac{\rho_{ij}^2}{n*(n-1)/2}$) is calculated using validated data to compare with that obtained from TCH. The error cross-correlation of individual precipitation products validated by gauged data and by TCH are plotted in Figures 19, S10, and S11. It is seen that the error cross-correlation estimated by TCH is comparable to that validated by gauged data, although some differences exist. The TCH approach provides a feasible way to estimate the error cross-correlation under some uncertainty. Different areas have different uncertainties, which is due to numerous factors. For example, the satellite sensors have different retrieval accuracy over different vegetation types. On the other hand, some areas do not have in situ stations and are interpolated. Figure 19 demonstrates the overall averaged error cross-correlation using multiple data sets. For a specific type of precipitation product, the error cross-correlation may be larger, especially for satellite-based data (Figures S12 and S13).

6. Conclusion

Precipitation data from gauge-based, satellite-estimated, and reanalysis data sets have substantial uncertainties. In this study, the generalized TCH method is developed to examine the relative uncertainty of 13 monthly precipitation products and 11 daily products on a global scale. The uncertainty of each precipitation product is analyzed spatially and temporally. The uncertainty of each precipitation data set varies in different places. CHIRPS, CRU TS, and WFDEI exhibit lower uncertainty in Africa than other individual precipitation products. Overall, MERRA-2 exhibits the minimum uncertainty among all the individual precipitation products. However, substantial uncertainty is seen in northern South America, Central Africa, and

Malay Archipelago in the MERRA-2 data. Given the uncertainty in individual precipitation products, a weighting scheme is developed to merge the individual precipitation data based on the inverse of error variance-covariance matrix. The weighting experiments suggest that the weighted precipitation can reduce the second-order errors of individual precipitation data substantially.

The TCH estimated uncertainty is validated using gauged data with minimal uncertainty among all the individual precipitation products. The underestimation of TCH uncertainty is likely a result of error of representativeness. However, the relative size of individual precipitation uncertainty can be well obtained. The generalized TCH method is capable of quantifying the relative uncertainty in gauge-based areas and can be extended to a global scale for ungauged and sparsely gauged areas. Overall, the TCH weighting scheme is able to generate a merged precipitation data set with substantially reduced random error. In gauged areas, the individual precipitation data sets are rescaled to the gauged data with maximum gauge coverage and weighted based on TCH. In ungauged areas, the individual precipitation data sets are rescaled to the data with minimal uncertainty estimated by TCH and weighted. In this way, the second-order error is substantially reduced and the first-order system bias is related to gauged data in gauged areas and is related to the data with minimal random error in ungauged areas due to the rescaling process.

The reduction of system bias is highly dependent on (hydrometeorological) gauged stations. However, the number of stations is limited in some areas due to no gauges or sparse gauges. Using in situ stations is almost the only way to approach to the precipitation truth to reduce the system bias. Therefore, we do not focus on this part in this study and admit the limitation of the TCH method in system bias reduction. It is very challenging to improve precipitation data accuracy at a global scale. We improved the accuracy of precipitation data by reducing the random error of precipitation products. Overall, the second order error of precipitation data is reduced. This method may not improve the precipitation accuracy at every grid of global regions but can improve it as a whole. This moves a step forward toward the challenge of precipitation estimation.

The TCH merged precipitation data sets are compared with the state-of-the-art multisatellite-fused precipitation product and exhibits a substantial improvement in terms of correlation, systematic bias, and random error over GPM IMERG both at monthly and daily scales. The TCH merging results are also compared with two previously used merging methods, namely, arithmetic mean and OOR mean, and exhibit superiority over them. Therefore, the developed TCH merging approach can be regarded as an effective way to merge multiple precipitation data sets for hydrometeorological applications.

Acknowledgments

The code and the weighted data set can be accessed from the HydroShare public repository (<https://www.hydroshare.org/resource/63c8cd8f63ec49ebabdca45826219a60/>), operated by The Consortium of Universities for the Advancement of Hydrologic Science, Inc. This work was supported by the National Key R&D Program (2018YFB2100500), the grants from the National Key Research and Development Program of China (2017YFB0503803), the Fundamental Research Funds for the Central Universities (2042017GF0057), and the National Natural Science Foundation of China program (41801339, 41771422, 41601406, and 41890822).

References

- Abbaszadeh, P., Moradkhani, H., & Daescu, D. N. (2019). The quest for model uncertainty quantification: A hybrid ensemble and variational data assimilation framework. *Water Resources Research*, 55, 2407–2431. <https://doi.org/10.1029/2018WR023629>
- Bastola, S., & Misra, V. (2014). Evaluation of dynamically downscaled reanalysis precipitation data for hydrological application. *Hydrological Processes*, 28(4), 1989–2002.
- Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., et al. (2019). Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences*, 23(1), 207–224. <https://doi.org/10.5194/hess-23-207-2019>
- Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & de Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21(1), 589–615. <https://doi.org/10.5194/hess-21-589-2017>
- Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., & 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>
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (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>
- Ferreira, V. G., Montecino, H. D., Yakubu, C. I., & Heck, B. (2016). Uncertainties of the Gravity Recovery and Climate Experiment time-variable gravity-field solutions based on three-cornered hat method. *Journal of Applied Remote Sensing*, 10(1), 015015.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, 2(1). <https://doi.org/10.1038/sdata.2015.66>
- Galindo, FJ, & Palacio, J. (1999). Estimating the instabilities of N correlated clocks.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14), 5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>
- Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple collocation-based merging of satellite soil moisture retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 6780–6792.
- Gruber, A., Su, C.-H., Zwiewack, S., Crow, W., Dorigo, W., & Wagner, W. (2016). Recent advances in (soil moisture) triple collocation analysis. *International Journal of Applied Earth Observation and Geoinformation*, 45, 200–211.

- Harris, I., Jones, P. D., Osborn, T. J., & 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.
- Hollinger, J. Lo, R., Poe, G., Savage, R., & Pierce, J. (1987). Special sensor microwave/imager user's guide. NRL Tech. Rpt.
- Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., et al. (2014). The Global Precipitation Measurement mission. *Bulletin of the American Meteorological Society*, 95(5), 701–722. <https://doi.org/10.1175/BAMS-D-13-00164.1>
- Huffman, G. J., Adler, R. F., Bolvin, D. T., & Nelkin, E. J. (2010). *The TRMM Multi-Satellite Precipitation Analysis (TMPA). Satellite rainfall applications for surface hydrology* (pp. 3–22). Dordrecht: Springer.
- Huffman, GJ, Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P. and Yoo, S.H. (2015). NASA Global Precipitation Measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm theoretical basis document, version 4, 30.
- Jalota, S., Sood, A., Chahal, G., & Choudhury, B. (2006). Crop water productivity of cotton (*Gossypium hirsutum* L.)–wheat (*Triticum aestivum* L.) system as influenced by deficit irrigation, soil texture and precipitation. *Agricultural Water Management*, 84(1-2), 137–146.
- 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. *Journal of Hydrometeorology*, 5(3), 487–503.
- Kawanishi, T., Sezai, T., Ito, Y., Imaoka, K., Takeshima, T., Ishido, Y., et al. (2003). The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASA's contribution to the EOS for global energy and water cycle studies. *IEEE Transactions on Geoscience and Remote Sensing*, 41(2), 184–194. <https://doi.org/10.1109/TGRS.2002.808331>
- Khajehi, S., Ahmadalipour, A., Moradkhani, H. (2017). An Effective Post-Processing of the North American Multi-Model Ensemble (NMME) Precipitation Forecasts over the Continental US. *Climate Dynamics*. <https://doi.org/10.1007/s00382-017-3934-0>
- Kobayashi, S., Ota, Y., Harada, Y., Ebina, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, 93(1), 5–48. <https://doi.org/10.2151/jmsj.2015-001>
- Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., et al. (2007). Global precipitation map using satellite-borne microwave radiometers by the GSMAp project: Production and validation. *IEEE Transactions on Geoscience and Remote Sensing*, 45(7), 2259–2275. <https://doi.org/10.1109/TGRS.2007.895337>
- Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & Simpson, J. (1998). The Tropical Rainfall Measuring Mission (TRMM) sensor package. *Journal of Atmospheric and Oceanic Technology*, 15(3), 809–817.
- Kummerow, C., Hong, Y., Olson, W. S., Yang, S., Adler, R. F., McCollum, J., et al. (2001). The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. *Journal of Applied Meteorology*, 40(11), 1801–1820. [https://doi.org/10.1175/1520-0450\(2001\)040<1801:TEOTGP>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2)
- Legates, D. R., & Willmott, C. J. (1990). Mean seasonal and spatial variability in gauge-corrected, global precipitation. *International Journal of Climatology*, 10(2), 111–127.
- Liu, ZJ, Yang, WK, Lv, ZC, Zhu, XW, & Ou, G. (2013). Application of three-cornered hats for frequency stability analysis of atomic clock. Paper presented at the Applied Mechanics and Materials.
- Long, D., Longuevergne, L., & Scanlon, B. R. (2014). Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. *Water Resources Research*, 50, 1131–1151. <https://doi.org/10.1002/2013WR014581>
- Massari, C., Crow, W., & Brocca, L. (2017). An assessment of the performance of global rainfall estimates without ground-based observations. *Hydrology and Earth System Sciences*, 21(9), 4347–4361.
- McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., & Stoffelen, A. (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. *Geophysical Research Letters*, 41, 6229–6236. <https://doi.org/10.1002/2014GL061322>
- Nguyen, P., Shearer, E. J., Tran, H., Ombadi, M., Hayatbini, N., Palacios, T., et al. (2019). The CHRS Data Portal, an easily accessible public repository for PERSIANN global satellite precipitation data. *Scientific data*, 6.
- O'Carroll, A. G., Eyre, J. R., & Saunders, R. W. (2008). Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations. *Journal of Atmospheric and Oceanic Technology*, 25(7), 1197–1207.
- Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., & Geenens, G. (2018). Data-driven model uncertainty estimation in hydrologic data assimilation. *Water Resources Research*, 54, 1252–1280. <https://doi.org/10.1002/2018WR022627>
- Premoli, A., & Tavella, P. (1993). A revisited three-cornered hat method for estimating frequency standard instability. *IEEE Transactions on Instrumentation and Measurement*, 42(1), 7–13.
- Reichle, R. H., Crow, W. T., & Keppenne, C. L. (2008). An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resources Research*, 44, W03423. <https://doi.org/10.1029/2007WR006357>
- Rieckh, T., & Anthes, R. (2018). Evaluating two methods of estimating error variances using simulated data sets with known errors. *Atmospheric Measurement Techniques*, 11(7), 4309–4325.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, 24(14), 3624–3648. <https://doi.org/10.1175/JCLI-D-11-00015.1>
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & 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(1-2), 15–40.
- Schneider, U., Ziese, M., Meyer-Christoffer, A., Finger, P., Rustemeier, E., & Becker, A. (2016). The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources. *Proceedings of the International Association of Hydrological Sciences*, 374, 29–34.
- Serreze, M. C., Clark, M. P., & Bromwich, D. H. (2003). Monitoring precipitation over the Arctic terrestrial drainage system: Data requirements, shortcomings, and applications of atmospheric reanalysis. *Journal of Hydrometeorology*, 4(2), 387–407.
- Shen, Y., Xiong, A., Hong, Y., Yu, J., Pan, Y., Chen, Z., & Saharia, M. (2014). Uncertainty analysis of five satellite-based precipitation products and evaluation of three optimally merged multi-algorithm products over the Tibetan Plateau. *International Journal of Remote Sensing*, 35(19), 6843–6858. <https://doi.org/10.1080/01431161.2014.960612>
- Siddique-E-Akbor, A. H. M., Hossain, F., Sikder, S., Shum, C. K., Tseng, S., Yi, Y., et al. (2014). Satellite precipitation data–driven hydrological modeling for water resources management in the Ganges, Brahmaputra, and Meghna Basins. *Earth Interactions*, 18(17), 1–25. <https://doi.org/10.1175/EI-D-14-0017.1>
- Sorooshian, S., Nguyen, P., Sellars, S., Braithwaite, D., AghaKouchak, A., & Hsu, K. (2014). Satellite-based remote sensing estimation of precipitation for early warning systems. *Extreme natural hazards, disaster risks and societal implications*, 1, 99.
- Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research*, 103(C4), 7755–7766.

- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56, 79–107. <https://doi.org/10.1002/2017RG000574>
- Swinbank, R., Shutyaev, V., & Lahoz, W. A. (2012). *Data assimilation for the earth system* (Vol. 26). Dordrecht: Springer Science & Business Media.
- Tavella, P., & Premoli, A. (1994). Estimating the instabilities of N clocks by measuring differences of their readings. *Metrologia*, 30(5), 479.
- Torcaso, F., Ekstrom, C., Burt, E., & Matsakis, D. (1998). Estimating frequency stability and cross-correlations. Paper presented at the Proceedings of the 30th Annual Precise Time and Time Interval Systems and Applications Meeting, Reston, Virginia.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., et al. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612), 2961–3012. <https://doi.org/10.1256/qj.04.176>
- Ushio, T., Okamoto, K. I., Iguchi, T., Takahashi, N., Iwanami, K., Aonashi, K., et al. (2003). The global satellite mapping of precipitation (GSMP) project. *Aqua (AMSR-E)*, 2004.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, 50, 7505–7514. <https://doi.org/10.1002/2014WR015638>
- Willmott, C. J., & Matsuura, K. (1995). Smart interpolation of annually averaged air temperature in the United States. *Journal of Applied Meteorology*, 34(12), 2577–2586.
- Xie, P., Chen, M., & Shi, W. (2010). CPC unified gauge-based analysis of global daily precipitation. Paper presented at the Preprints, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc.
- Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., & Yang, S. (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>
- Xu, L., Chen, N., & Zhang, X. (2019). Global drought trends under 1.5 and 2 °C warming. *International Journal of Climatology*, 39(4), 2375–2385.
- Xu, L., Chen, N., Zhang, X., & Chen, Z. (2018). An evaluation of statistical, NMME and hybrid models for drought prediction in China. *Journal of Hydrology*, 566, 235–249.
- Xu, L., Chen, N., Zhang, X., & Chen, Z. (2019). Spatiotemporal changes in China's terrestrial water storage from GRACE satellites and its possible drivers. *Journal of Geophysical Research: Atmospheres*, 124, 11976–11993. <https://doi.org/10.1029/2019JD031147>
- Yilmaz, M. T., & Crow, W. T. (2014). Evaluation of assumptions in soil moisture triple collocation analysis. *Journal of Hydrometeorology*, 15(3), 1293–1302.
- Ziese, M., Rauthe-Schöch, A., Becker, A., Finger, P., Meyer-Christoffer, A., & Schneider, U. (2018). GPCC full data daily version. 2018 at 1.0°: Daily land-surface precipitation from rain-gauges built on GTS-based and historic data: Global Precipitation Climatology Centre (GPCC, <http://gpcc.dwd.de/>) at Deutscher Wetterdienst.