Easy-to-use spatial Random Forest-based downscaling-calibration method for producing precipitation data with high resolution and high accuracy

Chuanfa Chen[1,2], Baojian Hu[1,2], Yanyan Li[1,2]

[1] College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China

[2] Key Laboratory of Geomatics and Digital Technology of Shandong Province, Shandong University of Science and Technology, Qingdao 266590, China

Correspondence to: Yanyan Li (yylee@whu.edu.cn)

Abstract. Precipitation data with high resolution and high accuracy is significantly important in numerous hydrological applications. To enhance the spatial resolution and accuracy of satellite-based precipitation products, an easy-to-use downscaling-calibration method based on spatial Random Forest (SRF-DC) is proposed in this study, where the spatial correlation of precipitation measurements between neighboring locations is considered. SRF-DC consists of two main stages. First, the satellite-based precipitation is downscaled by SRF with the incorporation of high-resolution variables including latitude, longitude, Normalized Difference Vegetation Index (NDVI), digital elevation model (DEM), terrain slope, aspect, relief, and land surface temperatures. Then, the downscaled precipitation is calibrated by SRF with rain gauge observations and the aforementioned high-resolution variables. The monthly Integrated MultisatellitE Retrievals for Global Precipitation Measurement (IMERG) over Sichuan province, China from 2015 to 2019 was processed using SRF-DC, and its results were compared with those of classical methods including geographically weighted regression (GWR), artificial neural network (ANN), random forest (RF), kriging interpolation only on gauge measurements, bilinear interpolation-based downscaling and then SRF-based calibration (Bi-SRF), and SRF-based downscaling and then geographical difference analysis (GDA)-based calibration (SRF-GDA). Comparative analyses with respect to root mean square error (RMSE), mean absolute error (MAE) and correlation coefficient (CC) demonstrate that: (1) SRF-DC outperforms the classical methods as well as the original IMERG; (2) the monthly-based SRF estimation is slightly more accurate than the annual-based SRF fraction disaggregation method; (3) SRF-based downscaling and calibration performs better than bilinear downscaling (Bi-SRF) and GDA-based calibration (SRF-GDA); (4) kriging is more accurate than GWR and ANN, whereas its precipitation map loses

detailed spatial precipitation patterns; and (5) based on the variable importance rank of RF, the precipitation interpolated by kriging on the rain gauge measurements is the most important variable, indicating the significance of incorporating spatial autocorrelation for precipitation estimation.

1. Introduction

Precipitation is an important variable for promoting our understanding of hydrological cycle and water resource management (Chen et al., 2010). Previous studies have showed that about 70-80% of hydrological modeling errors are caused by precipitation uncertainties (Gebregiorgis and Hossain, 2013). However, precipitation is also one of the most difficult meteorological factors to estimate due to its high spatial and temporal heterogeneity (Beck et al., 2019). Although point-based rain gauge observations are reliable and accurate, it is difficult to reflect the spatial precipitation pattern because of the sparse and uneven distribution of meteorological stations, especially in remote and mountainous areas (Ullah et al., 2020).

During the past decades, diverse satellite-based precipitation datasets have been produced, such as the Climate Hazards Group Infrared Precipitation with Station data (CHIRPS, 0.05°) (Funk et al., 2015), the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR, 0.25°) (Ashouri et al., 2015), the Climate Prediction Center (CPC) morphing technique (CMORPH, 0.25°) (Haile et al., 2013), the Multi-Source Weighted-Ensemble Precipitation (MSWEP, 0.1°) (Beck et al., 2017), the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA, 0.25°) (Huffman et al., 2007) and the Integrated MultisatellitE Retrievals for Global Precipitation Measurement (GPM) mission (IMERG, 0.1°) (Hou et al., 2014). Nevertheless, these products are characterized by considerable systematic biases due to the shortcomings of retrieval algorithms, sensor capability and spatiotemporal collection frequency (Chen et al., 2018; Wu et al., 2018; Yang et al., 2017). Moreover, their resolutions (from 0.05° to 2.5°) are too coarse for hydrological modeling when applied to local and basin regions (Immerzeel et al., 2009).

As a result, downscaling techniques have been widely adopted to derive high resolution precipitation products. This is generally achieved by firstly modeling the relationship between precipitation and land surface variables at a coarse scale, and then putting the high resolution variables into the constructed model to downscale the precipitation data (Immerzeel et al., 2009; Chen et al., 2010). Immerzeel et al. (2009) employed an exponential regression (ER) to describe the relationship between TRMM and Normalized Difference Vegetation Index (NDVI). Jia et al. (2011) used a multiple linear regression model (MLR) to establish the relationship between TRMM, digital elevation model (DEM) and NDVI. Duan and Bastiaanssen (2013) proposed a downscaling model based on the second-order polynomial relationship between TRMM and NDVI. Considering the heterogeneous relationships between precipitation and land surface variables across the study area, geographically weighted regression (GWR) was widely used (Chen et al., 2014; Chen et al., 2015; Xu et al., 2015; Li et al., 2019; Chen et al., 2020c; Lu et al., 2020; Zhao et al., 2018). In the recent decade, some data-driven machine learning (ML) methods were employed to downscale satellite-based precipitation products, such as random forest (RF) (Shi et al., 2015; Zhang et al., 2021), support vector machine (SVM) (Jing et al., 2016; Chen et al., 2010) and artificial neural network (ANN) (Elnashar et al., 2020), and showed more accurate results than the statistical methods. However, the downscaled precipitation products inherently contain large systematic biases.

To alleviate the inherent biases, many calibration methods have been proposed to merge gauge observations and satellite-based precipitation, such as nonparametric kernel smoothing method (Li and Shao, 2010), geographical difference analysis (GDA) (Cheema and Bastiaanssen, 2012), geographical ratio analysis (GRA) (Duan and Bastiaanssen, 2013), conditional merging (CM) (Berndt et al., 2014), quantile mapping (Chen et al., 2013; Zhang and Tang, 2015), optimal interpolation (Xie and Xiong, 2011; Lu et al., 2020; Wu et al., 2018), GWR (Chen et al., 2018; Lu et al., 2019; Chao et al., 2018) and geostatistical interpolation (Park et al., 2017). Nevertheless, these methods are based on some strict assumptions, which might be not satisfied in reality (Zhang et al., 2021; Wu et al., 2020). To this end, ML-based calibration methods have been widely used, such as Quantile Regression Forest (QRF) (Bhuiyan et al., 2018), ANN (Yang and Luo, 2014; Pham et al., 2020), deep neural network (Tao et al., 2016), RF (Baez-Villanueva et al., 2020), convolutional neural network (CNN) (Wu et al., 2020), SVM and extreme learning machine (Zhang et al., 2021).

Compared to the statistical methods, the merits of the ML-based methods are as follows (Zhang et al., 2021; Hengl et al., 2018): (i) they require no strict statistical assumption; (ii) they can capture the complex and nonlinear relationship between precipitation and its influence factors; (iii) they generally outperform the statistical methods. However, ML-based methods were simply taken as statistical tools without considering the spatial autocorrelation of precipitation measurements between adjacent locations. Moreover, they were adopted in either downscaling or calibration of precipitation. Specifically, some (Karbalaye Ghorbanpour et al., 2021; Yan et al., 2021; Jing et al., 2016) attempted to use the ML methods for downscaling and then use the classical method (e.g. GDA) for calibration, while some (Zhang et al., 2021) employed the classical interpolation methods (e.g. bilinear interpolation) for downscaling and then used the ML methods for calibration. However, we regard that the use of ML methods in both downscaling and calibration could improve the accuracy of precipitation. To the best of our knowledge, no previous studies have used the ML technique in both downscaling and calibration (Karbalaye Ghorbanpour et al., 2021; Yan et al., 2021).

Based on aforementioned discussion, the objectives of this study are twofold: (i) to develop an easy-to-use spatial RF (SRF) by incorporating spatial autocorrelation for precipitation estimation, and (ii) to propose a downscaling-calibration method based on SRF (SRF-DC) for producing high resolution and high accuracy precipitation products. RF is taken as the basic model in this study owing to its high interpolation accuracy and low computational cost (Mohsenzadeh Karimi et al., 2020; Belgiu et al., 2016).

SRF-DC consists of two main steps. First, the precipitation data is downscaled by SRF with the incorporation of high resolution environmental variables, including DEM, NDVI, land surface temperatures (LSTs), terrain parameters, latitude and longitude, as recommended in previous studies (Jing et al., 2016; Li et al., 2019). Second, SRF and the environmental variables are further used to merge the downscaled precipitation data and gauge observations to boost the accuracy of the precipitation data. The merit of SRF-DC lies in the use of SRF for both downscaling and calibration of precipitation products, with the incorporation of high-resolution environmental variables.

2 Study area and dataset

2.1. Study area

Sichuan province between 97°21'-108°31'E and 26°03'-34°19'N (Fig. 1) is situated between the Qinghai-Tibet Plateau and the Plain of the Middle-and-lower Reaches of Yangtze River, with an area of

486,000 km^2. Its topography is very complex, including mountains, hills, plain basins and plateaus, and the elevations range from approximately 180 m in the east to 7100 m in the west. Such a variety of complex topography results in different climates across the study region. Specifically, the east basin has subtropical monsoon climate. The weather is generally warm, humid and foggy with much cloud, fog and rain but less sunshine. While in the west plateau, the weather is relatively cool or cold. The climate is featured by a long cold winter, a very short summer and rich sunshine but less rainfall. Annual precipitation shows significant spatial heterogeneity, varying from about 400 mm in the west to 1800 mm in the east. Moreover, more than 80% precipitation occurs between July and September. The high spatial and temporal variability of precipitation makes the study site ideal for evaluating satellite-based precipitation estimates (Zhang et al., 2021; Karbalaye Ghorbanpour et al., 2021).

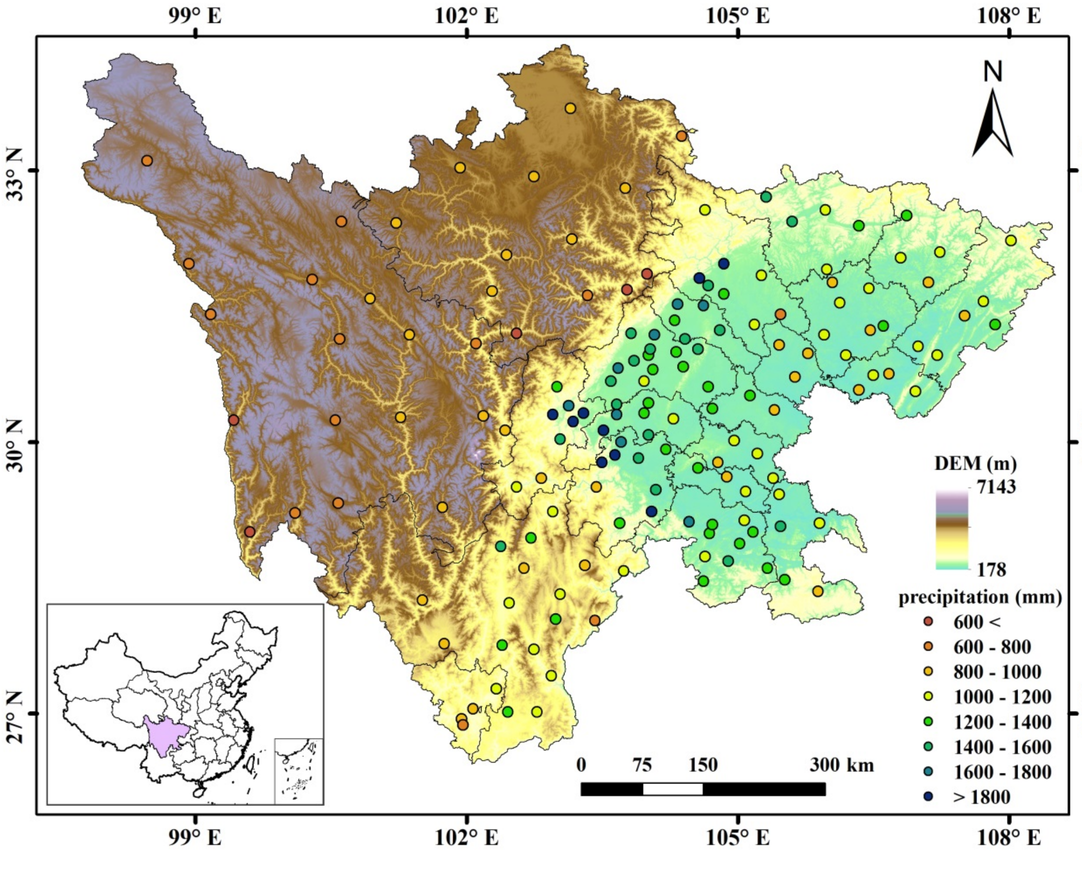


Fig. 1 Topography, rain gauges and geographic location of Sichuan province in China

2.2. Dataset

2.2.1. Rain gauge observations

The study region has 156 rain gauge stations, which shows an uneven distribution with high density in the east and low density in the west (Fig. 1). The average cover area of one rain gauge observation is about 3115 km^2. Daily precipitation of all the stations for the period 2015–2019 was collected from the China Meteorological Data Service Center (CMDSC, http://data.cma.cn/). The data quality was guaranteed based on some strict quality controls, such as manual inspection, outlier check and spatiotemporal consistency verification (Zhao and Yatagai, 2014). After that, the monthly precipitation was produced by aggregating the daily precipitation of rain gauges for each month.

2.2.2. Integrated MultisatellitE Retrievals for Global Precipitation Measurement (IMERG)

As the successor of TRMM, the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA) initiated the next-generation global precipitation observation mission (Hou et al., 2014). The IMERG products were generated by assimilating all microwave and infrared (IR) estimates, together with gauge observations (Huffman et al., 2019). It has the spatial resolution of 0.1° × 0.1° with the coverage from 60°S-60°N. IMERG provides three different products including Early, Late, and Final Runs, which were estimated about 4 hours, 14 hours, and 3.5 months after observation time, respectively. Due to the incorporation of the Global Precipitation Climatology Centre (GPCC) rain gauge data, IMERG Final Run is more accurate than the others (Lu et al., 2019). Thus, the monthly IMERG V06B Final Run product was adopted in the study. It was downloaded from https://gpm.nasa.gov/data/.

The average monthly precipitation of all rain gauges and that of IMERG at the corresponding grid cells from 2015-2019 over Sichuan province are shown in Fig. 2. Obviously, IMERG has an overestimation in most months and the wettest month is July 2018.

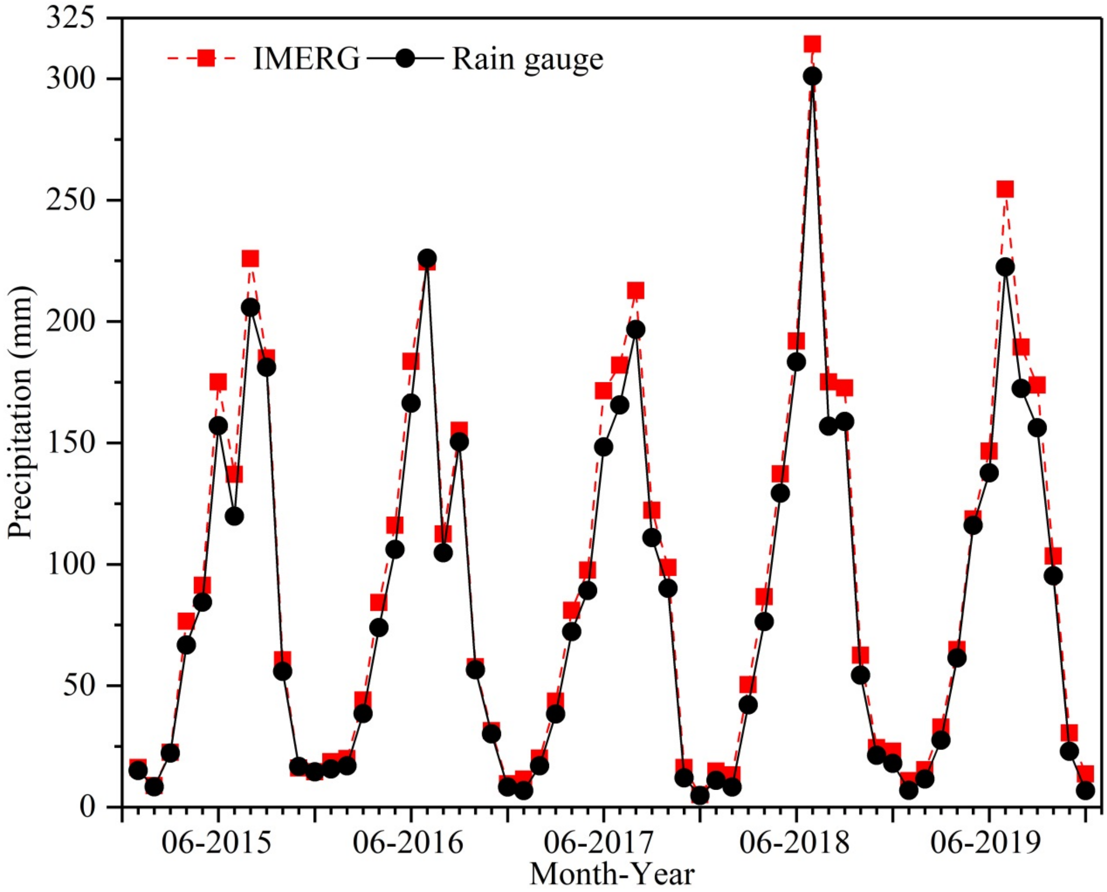


Fig. 2 Average monthly precipitation of all rain gauges and that of IMERG at the corresponding grid cells from 2015-2019 over Sichuan province

2.2.3. Environmental variables

Vegetation types have a significant impact on fluxes of sensible and latent heat into the atmosphere, apparently influencing the humidity of the lower atmosphere and further affecting moist convection (Spracklen et al., 2012). Therefore, as an indicator of vegetation activity, NDVI has been widely adopted to estimate precipitation (Wu et al., 2019; Immerzeel et al., 2009). In this study, the Moderate Resolution Imaging Spectroradiometer (MODIS) monthly NDVI with the resolution of 1 km (MOD13A3) from 2015 to 2019 (https://search.earthdata.nasa.gov/) was used.

Precipitation can influence LTS both in the daytime and at night; rain leads to cool temperatures, and droughts often couple with heat waves (Trenberth and Shea, 2005; Jing et al., 2016). Thus, the daytime LST (LST_D), nighttime LST (LST_N), and the difference between daytime and nighttime LSTs (LST_{D-N}) at the monthly scale were used in this study. Here, MODIS 8-day LST with the resolution of 1 km (MOD11A2) from 2015 to 2019 was downloaded from https://ladsweb.modaps.eosdis.nasa.gov/ and then temporally averaged into the monthly LST products.

Topography could affect the regional atmospheric circulation and the spatial pattern of precipitation through its thermal and dynamic forcing mechanisms (Jing et al., 2016; Jia et al., 2011). With the increase of elevations, the relative humidity of the air masses increases through expansion and cooling of the rising air masses, which brings precipitation (Jing et al., 2016). Thus, the precipitation-DEM relationship has been widely employed to downscale satellite precipitation dataset. Here, the Shuttle Radar Topography Mission (SRTM) DEM (Shortridge and Messina, 2011) was used. The SRTM DEM with the spatial resolution of 90 m was downloaded from http://srtm.csi.cgiar.org/ and then resampled to 1 km by the pixel averaging method. Since precipitation tends to be influenced by terrain variability and terrain orientation, DEM derivatives including slope, aspect and terrain relief (Chen et al., 2020a) were also used in the study. These derivatives were extracted from the SRTM DEM using ArcGIS 10.3.

The detailed information of all the datasets used in the study is shown in Table 1.

Table 1 Datasets used in the study

Data Type	Product	Spatial resolution	Temporal resolution	Source
Meteorological data	IMERG	10 km	Monthly	https://gpm.nasa.gov/data.
	Rain gauge observations	-	Daily	http://data.cma.cn/
Land surface data	SRTM DEM	30 m	-	http://srtm.csi.cgiar.org/
	slope, aspect, terrain relief	30 m	-	Derived from SRTM DEM
	NDVI	1 km	Monthly	https://search.earthdata.nasa.gov/
	LST	1 km	8-days	https://ladsweb.modaps.eosdis.nasa.gov

3. Methodology

The flowchart of SRF-DC is illustrated in Fig. 3, which includes three stages: data processing, IMERG downscaling and downscaled IMERG calibration. It is noted that each IMERG pixel represents the areal average precipitation within it, whereas rain gauge measurements are point-based. Therefore, downscaling before calibration can decrease scale mismatch between pixel-based areal precipitation and gauge-based point measurements.

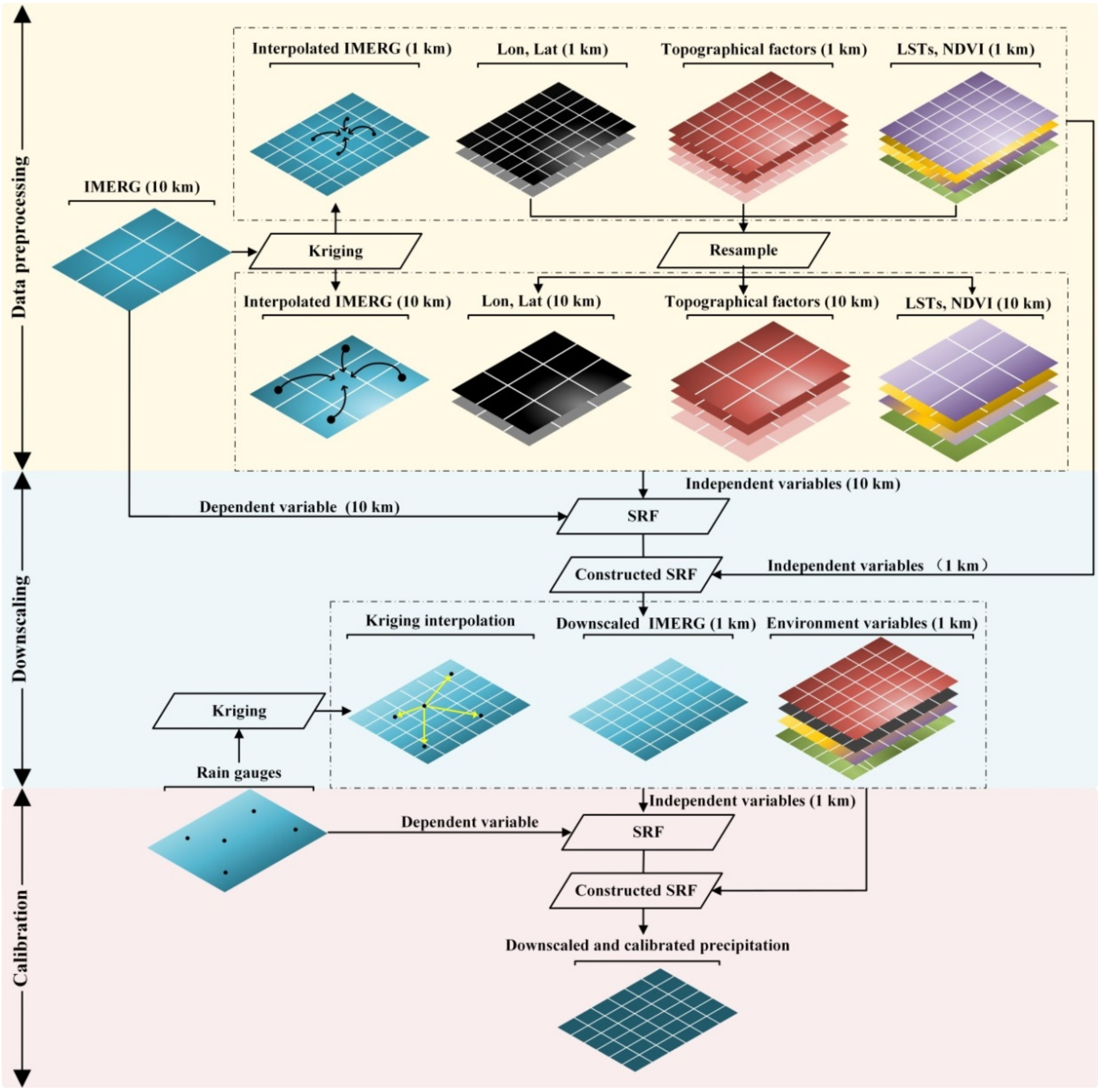


Fig. 3 Flowchart of SRF-DC in this study

3.1. Random Forest (RF)

RF is an ensemble of several tree predictors such that each tree relies on a random and independent selection of some samples and features but with the same distribution (Breiman, 2001). The general framework of RF is shown in Fig. 4. Specifically, each decision tree is constructed by randomly collecting some training data with replacement, while the others are used to assess the tree performance (sample bagging). When constructing each tree, only a random subset of features is selected at each decision node (feature bagging). In the end, the majority vote for classification or the average prediction of all trees for regression is used to obtain the final output. Overall, RF includes three parameters to set: number of trees, depth of the tree, and number of features.

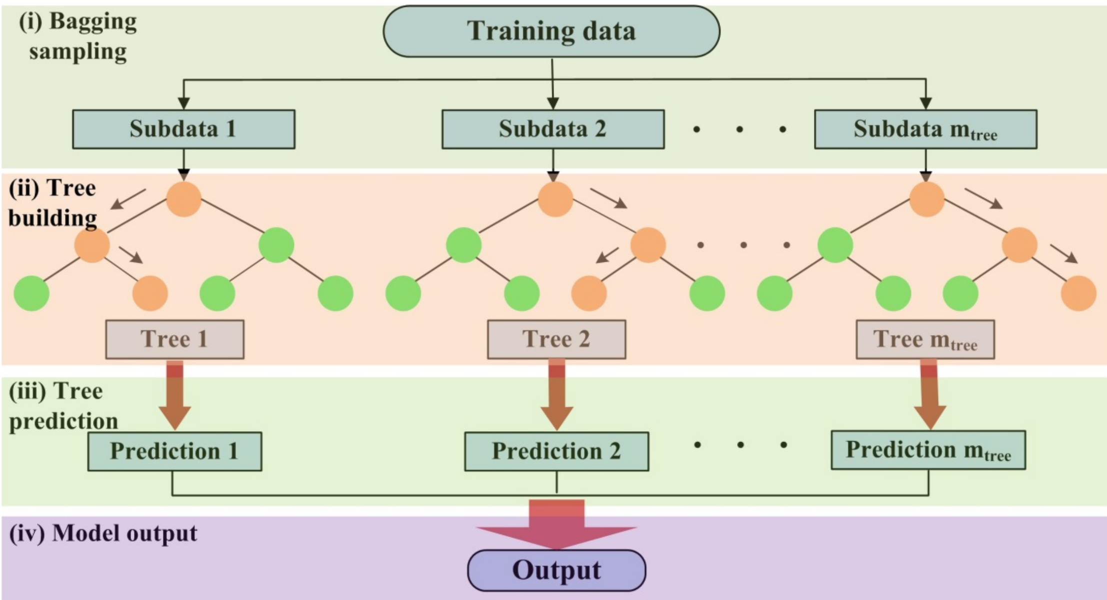


Fig. 4 General framework of RF

Meanwhile, RF can evaluate the relative importance of each predictor by means of the out-of-bag (OOB) observations, i.e. the samples without being used for model construction. Specifically, to measure the importance of the *i*th predictor, its values are permuted while the values of the other predictors remain unchanged. Then, the OOB error based on the permuted samples is computed. Next, the importance score of the *i*th predictor is computed by averaging the difference between the OOB errors before and after the permutation. With the estimated scores, the importance of each variable is ranked.

In this study, the RF regression model was performed with the freely available codes, downloaded from the website (https://code.google.com/archive/p/randomforest-matlab/downloads).

3.2. Spatial Random Forest (SRF)

In essence, the classical RF is a non-spatial statistical technique for spatial prediction, as it neglects sampling locations and general sampling pattern (Hengl et al., 2018). This can potentially cause sub-optimal estimations, especially when the spatial autocorrelation between dependent variables is high. To this end, a spatial RF (SRF) is proposed in this study. The general formulation of SRF is as follows:

$$p\left(s_0\right) = f\left(\boldsymbol{X}_s, \boldsymbol{X}_{ns}\right) + e \tag{1}$$

where $p(s_0)$ is the estimated precipitation at location s_0, e is the fitting residual, $f(\bullet)$ is the function constructed by SRF, and $\boldsymbol{X}_s$ and $\boldsymbol{X}_{ns}$ are the spatial and non-spatial covariates, respectively.

In addition to spatial coordinates, one spatial covariate (X_s) is computed to account for the spatial autocorrelation of precipitation measurements between neighboring locations:

$$X_s\left(s_0\right)=\sum_{i=1}^{n} w_i z\left(s_i\right) \tag{2}$$

where s_i is the ith neighbor of s_0, $z(s_i)$ is the precipitation data of s_i, w_i is its weight, and n is the number of neighbors.

In previous studies (Zhang et al., 2021; Li et al., 2017), the inverse distance weights (IDW) were widely used. However, the IDW method only resorts to the spatial distance between the estimated location and its neighbor locations, and does not consider the spatial autocorrelation between the neighboring locations. To overcome this limitation, ordinary kriging (OK)-based variogram is adopted to estimate the interpolation weights in this study by solving the following linear system:

$$\begin{pmatrix} \gamma\left(\boldsymbol{x}_1-\boldsymbol{x}_1\right) & \cdots & \gamma\left(\boldsymbol{x}_1-\boldsymbol{x}_n\right) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma\left(\boldsymbol{x}_n-\boldsymbol{x}_1\right) & \cdots & \gamma\left(\boldsymbol{x}_n-\boldsymbol{x}_n\right) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma\left(\boldsymbol{x}_1-\boldsymbol{x}_0\right) \\ \vdots \\ \gamma\left(\boldsymbol{x}_n-\boldsymbol{x}_0\right) \\ 1 \end{pmatrix} \tag{3}$$

where μ is Lagrange parameter and $\gamma\left(\cdot\right)$ is the semivariogram.

It can be concluded that the variogram-based weights consider the spatial autocorrelation not only between the known locations, but also between the known locations and the interpolated location (Berndt and Haberlandt, 2018). In practice, the experimental semivariogram can be estimated from sample data as follows (Goovaerts, 2000):

$$\gamma\left(h\right)=\frac{1}{2n}\sum_{i=1}^{n}\left(z\left(s_i\right)-z\left(s_i+h\right)\right)^2 \tag{4}$$

where n is the number of data pairs with the attribute z separated by distance h.

To obtain the semivariogram at any h, a theoretical semivariogram model should be fitted to the experimental values. There are four commonly used theoretical semivariogram models: the spherical, Gaussian, exponential, and power models. The best one with the highest fitting R^2 was used in the study.

3.3. Working procedure of the proposed method

The detailed steps of SRF-DC are as follows (Fig. 3):

(1) Each pixel value of the 10 km IMERG was re-estimated by OK interpolation with its *k* nearest neighbors (e.g. *k*=8) to obtain the interpolated IMERG (termed as $I_s^{10\text{km}}$), the 10 km IMERG was interpolated by OK to obtain the interpolated 1 km IMERG ($I_s^{1\text{km}}$), and the gauge observations were interpolated by OK to produce the 1 km precipitation map ($P_s^{1\text{km}}$). This step aims to provide spatial variables for SRF, i.e. X_s in Eq. (1). Since the semivariogram model cannot be accurately estimated from the sparse gauge measurements, the satellite-based precipitation was used to derive the model, as suggested by Chen et al. (2020c). To estimate $I_s^{10\text{km}}$ and $I_s^{1\text{km}}$, the raster-based 10 km IMERG was first transformed into the point-based form with spatial coordinates and precipitation values, and then the scattered points were interpolated by OK to produce raster-based maps.

(2) The negative NDVI values were excluded from the original data, which mainly belong to snow and water bodies in the study site. The removed values were interpolated by OK with their neighbors to avoid information loss.

(3) The 1 km environmental variables $\boldsymbol{X}_{ns}^{1\text{km}}$ (i.e. NDVI, LST_D, LST_N, $LST_{D\text{-}N}$, DEM, slope, aspect, terrain relief, latitude and longitude) were resampled to the 10 km resolution $\boldsymbol{X}_{ns}^{10\text{km}}$ by the pixel averaging method because the average value reflects the overall trend within each 10 km pixel and reduces the influence of outliers in the 1 km pixels.

(4) The relationship between $\boldsymbol{X}_{ns}^{10\text{km}}$, $I_s^{10\text{km}}$ and the original 10 km IMERG ($D^{10\text{km}}$) was constructed by SRF:

$$D^{10\text{km}}\left(s_0\right) = f_{\text{downscale}}\left(I_s^{10\text{km}}\left(s_0\right), \boldsymbol{X}_{ns}^{10\text{km}}\left(s_0\right)\right) + e^{10\text{km}}\left(s_0\right) \tag{5}$$

where *e* is the fitting residual.

(5) The 10 km IMERG ($D^{10\text{km}}$) was downscaled to 1 km ($D^{1\text{km}}$) by applying the constructed model in step (4) to $\boldsymbol{X}_{ns}^{1\text{km}}$ and $I_s^{1\text{km}}$:

$$D^{1\text{km}} = f_{\text{downscale}}\left(I_s^{1\text{km}}, \boldsymbol{X}_{ns}^{1\text{km}}\right) \tag{6}$$

(6) The relationship between the 1 km predictors and the gauge observations (*G*) was constructed by

SRF:

$$G\left(s_0\right)=f_{\text{calibrate}}\left(P_s^{1\text{km}}\left(s_0\right),D^{1\text{km}}\left(s_0\right),\boldsymbol{X}_{ns}^{1\text{km}}\left(s_0\right)\right)+e^{1\text{km}}\left(s_0\right) \tag{7}$$

(7) The 1 km precipitation data ($C^{1\text{km}}$) was produced based on the constructed relationship in step (6):

$$C^{1\text{km}}=f_{\text{calibrate}}\left(P_s^{1\text{km}},D^{1\text{km}},\boldsymbol{X}_{ns}^{1\text{km}}\right) \tag{8}$$

In this study, residual correction was ignored during downscaling and calibration, as many previous studies (Karbalaye Ghorbanpour et al., 2021; Lu et al., 2019) demonstrated that residual correction on the ML-based technique could decrease prediction accuracy.

3.4. Comparative methods

In the study, the performance of SRF-DC was comparatively assessed under three manners. Firstly, we compared the results of SRF-DC with those of the classical methods including GWR, RF and BPNN. Secondly, SRF-DC was compared with two frameworks: (i) the IMERG was first downscaled by the bilinear interpolation and then calibrated by SRF (termed as Bi-SRF), and (ii) the IMERG was first downscaled by SRF and then calibrated by GDA (termed as SRF-GDA). Thirdly, SRF-DC at the monthly scale was compared with the annual-based SRF fraction disaggregation method (termed as SRFdis). Specifically, the IMMERG was first downscaled and calibrated by SRF on an annual scale and then the estimated annual precipitation was disaggregated into monthly precipitation using monthly fractions, as proposed by Duan and Bastiaanssen (2013). Finally, SRF-DC was compared with OK interpolation only on gauge measurements (termed as kriging). Overall, SRF-DC was compared with seven classical methods in this study including GWR, RF, BPNN, Bi-SRF, SRF-GDA, SRFdis and kriging.

To quantitatively analyze the performance of all the methods, all rain gauge observations were randomly divided into *l* folds (e.g. *l*=10), where the *l*-1 folds (i.e. training/validating data) was used to construct the model, while the remaining fold (i.e. testing data) to assess the performance of the model (Xu and Goodacre, 2018). During model construction, the *l*-1 folds were randomly divided into training and validating datasets with the proportions of 80% and 20%, respectively, where the former was used to train the model and the latter to validate the model. Then, the performance of the model with the optimized parameters was assessed using the testing data. The aforementioned process was repeated *l* times until all folds were taken as the testing data.

3.5. Accuracy measures

We comparatively analyzed the performance of all methods with four accuracy measures including root mean square error (RMSE), mean error (ME), mean absolute error (MAE) and correlation coefficient (CC) (Jing et al., 2016; Sharifi et al., 2019). They are respectively expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(E_i - O_i\right)^2} \tag{9}$$

$$ME = \frac{\sum_{i=1}^{n}\left(E_i - O_i\right)}{n} \tag{10}$$

$$MAE = \frac{\sum_{i=1}^{n}\left|E_i - O_i\right|}{n} \tag{11}$$

$$CC = \frac{\sum_{i=1}^{n}\left(E_i - \bar{E}\right)\left(O_i - \bar{O}\right)}{\sqrt{\sum_{i=1}^{n}\left(E_i - \bar{E}\right)^2} \times \sqrt{\sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2}} \tag{12}$$

where n is the number of testing points, and E_i and O_i are the estimated and observed precipitation at location i, respectively.

4. Results and analysis

Fig. 5 illustrates the scatterplots between the predicted and observed precipitation on a monthly scale from 2015 to 2019. Results show that the original IMMERG is heavily biased with the ME value of 8.01 mm. In contrast, except for kriging, all the other models greatly reduce the bias with the ME values approximate to zero. In other words, the models with the incorporation of high resolution variables become unbiased. With respect to RMSE, MAE and CC, BPNN produces worse results than the original IMERG. The performance of GWR is also unsatisfactory. This is mainly attributed to the complex relationship between precipitation and predictors, which cannot be properly described by the two models. RF and kriging perform better than IMERG. The four SRF-based methods including SRF-DC, Bi-SRF, SRF-GDA and SRFdis outperform the other methods. This indicates the importance of spatial autocorrelation for precipitation estimation. Moreover, among the four versions of SRF, SRF-GDA has the lowest accuracy, indicating that SRF is more important for calibration than downscaling. SRF-DC with the RMSE, MAE and CC values of 32.20 mm, 18.77 mm and 0.937 produces the best result. Thus, it can be concluded that (i) SRF-based downscaling and calibration is more effective than bilinear downscaling (Bi-SRF) and GDA-based calibration (SRF-GDA) and (ii) there is no obvious time latency for vegetation response to precipitation in the study site, as SRF-DC on the monthly scale is slightly more accurate than SRFdis on the annual scale.

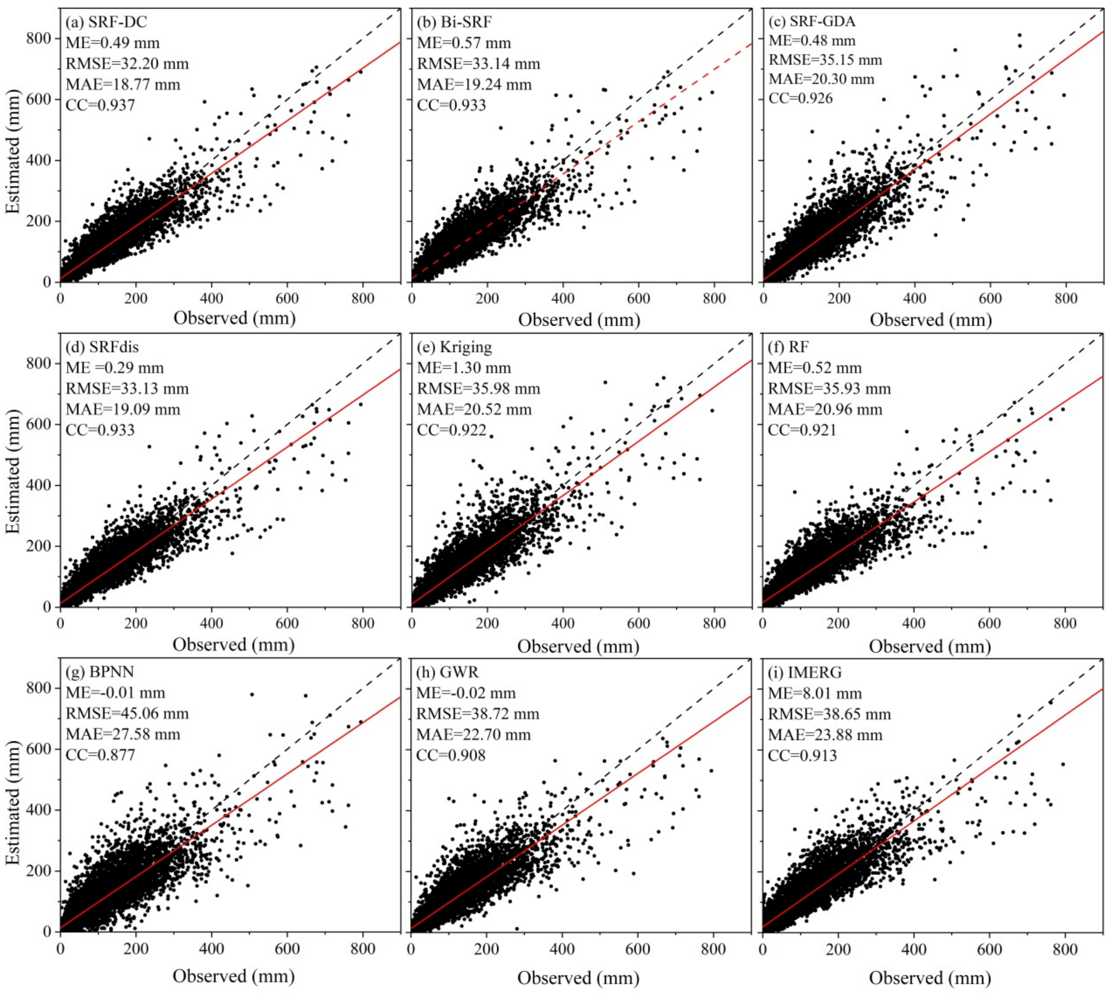


Fig. 5 Scatterplots between the estimated and observed precipitation on a monthly scale from 2015 to 2019. Fitting line with the red color models the relationship between the observed and estimated precipitation.

However, as shown in Fig. 5, all the methods significantly underestimate precipitation when the values are greater than 400 mm. To quantitatively analyze the performance of all methods on the high precipitation, their accuracy measures are shown in Table 2. Results show that all methods have poor results for these observations. A possible reason is that high precipitation is often caused by complicated environmental factors, which cannot be sufficiently explained by the constructed predictors-precipitation relationships. In terms of ME, SRF-GDA ranks the first, which is followed by kriging and SRF-DC. However, their ME values are less than -70 mm. With respect to RMSE and MAE, kriging performs the best, which is closely followed by SRF-DC, while with respect to CC, SRF-DC with the value of 0.64 outperforms the others.

Table 2 Accuracy measures of all methods for estimating high precipitation (i.e. values greater than 400 mm)

Method	ME (mm)	RMSE (mm)	MAE (mm)	CC
SRF-DC	-105.54	149.80	124.82	0.64
Bi-SRF	-110.96	156.81	130.67	0.60
SRF-GDA	-74.21	150.10	126.02	0.55
SRFdis	-117.31	160.11	137.29	0.61
Kriging	-86.25	146.94	119.53	0.58
RF	-141.53	177.71	150.83	0.61
BPNN	-118.88	171.23	142.00	0.57
GWR	-139.02	178.85	145.19	0.57
IMERG	-136.22	173.24	143.69	0.55

Fig. 6 shows the boxplots of the four accuracy measures. Obviously, BPNN obtains the lowest accuracy. It is followed by GWR and IMERG. RF and kriging show better results than BPNN, GWR and IMERG. The four methods based on SRF seem more accurate than the classical methods. Moreover, SRF-DC slightly outperforms the other SRF-based methods, which highlights the benefit of including spatial autocorrelation for downscaling and calibration of satellite-based precipitation.

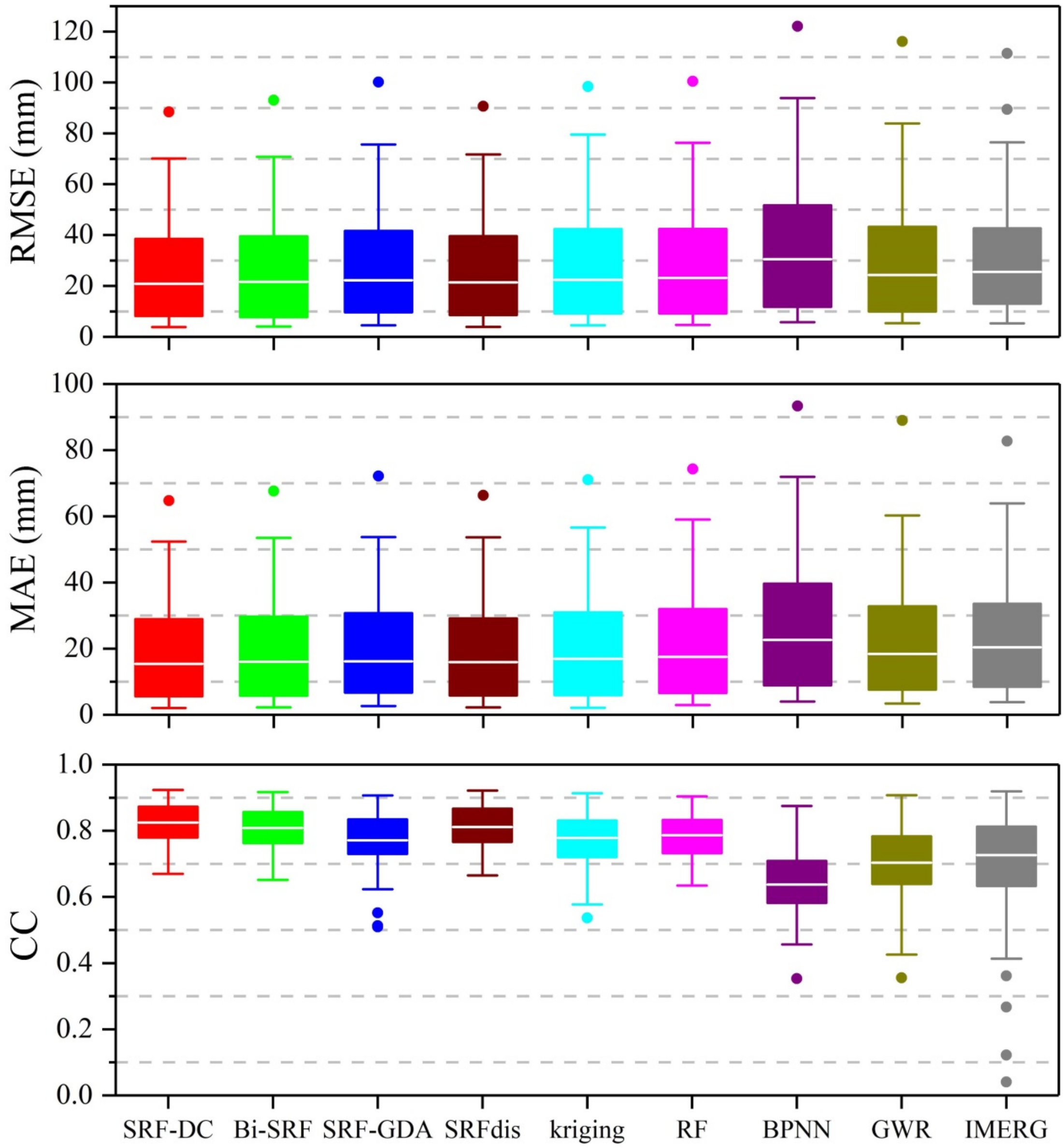


Fig. 6 Boxplots of RMSE, MAE and CC values of all the methods on a monthly scale during 2015-2019

Fig. 7 shows the RMSE spatial distributions of SRF-DC, SRFdis, RF, BPNN, kriging and GWR on all gauge stations. Overall, the RMSEs tend to be larger in the middle area, since it has higher precipitation than the other areas (Fig. 1). BPNN (Fig. 7d) yields the poorest result, where many stations have the RMSE values greater than 60 mm. It is followed by GWR (Fig. 7f). RF (Fig. 7c) and kriging (Fig. 7e) are better than GWR and BPNN at most stations. SRF-DC (Fig. 7a) and SRFdis (Fig. 7b) are more accurate than the classical methods, especially at the stations in the middle area.

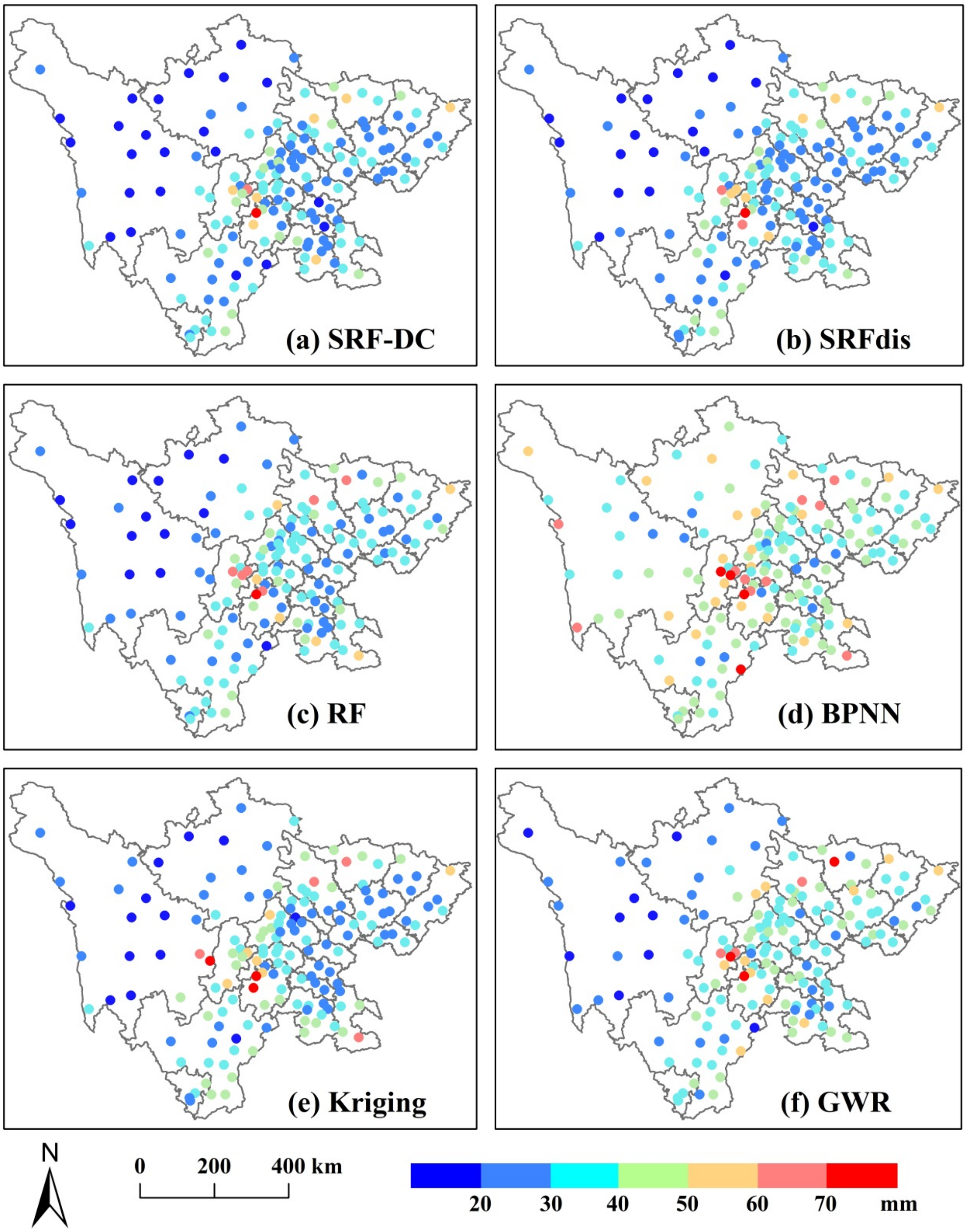


Fig. 7 RMSE distributions of SRF-DC and some representative methods for all gauge stations on a monthly scale during 2015-2019

Since the wettest month is July 2018 (Fig. 2), it is taken as an example to show the precipitation maps of SRF-DC and some classical methods. Moreover, the semivariogram of kriging derived from the original IMMERG and its prediction error map are shown, since they play an important role in the performance of kriging and SRF-based methods. Results (Fig. 8) indicate that precipitation produced by all the methods have spatial distribution patterns similar to IMERG, with much high precipitation in the middle and low precipitation in the east. The ML-based methods have more spatial details of precipitation than IMERG due to the inclusion of high-resolution predictors for precipitation estimation. The kriging map is so smooth that many details and variations of precipitation pattern are lost. This is expected as it only uses ground measurements for the interpolation. RF shows obvious unnatural discontinuities at the bottom. GWR suffers from systematic anomalies, with the values clearly greater than their neighbors. In comparison, SRF-DC produces a good precipitation map.

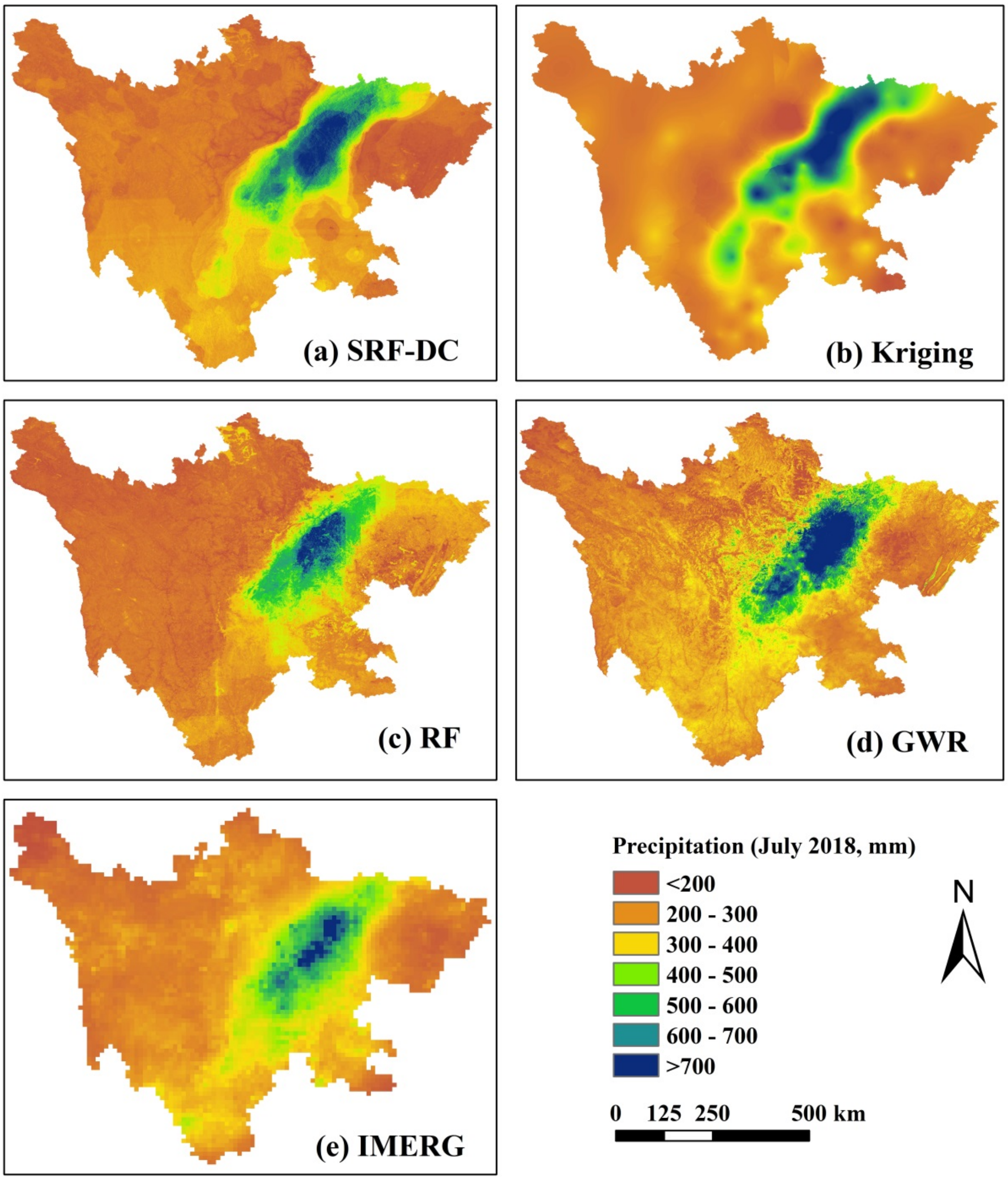
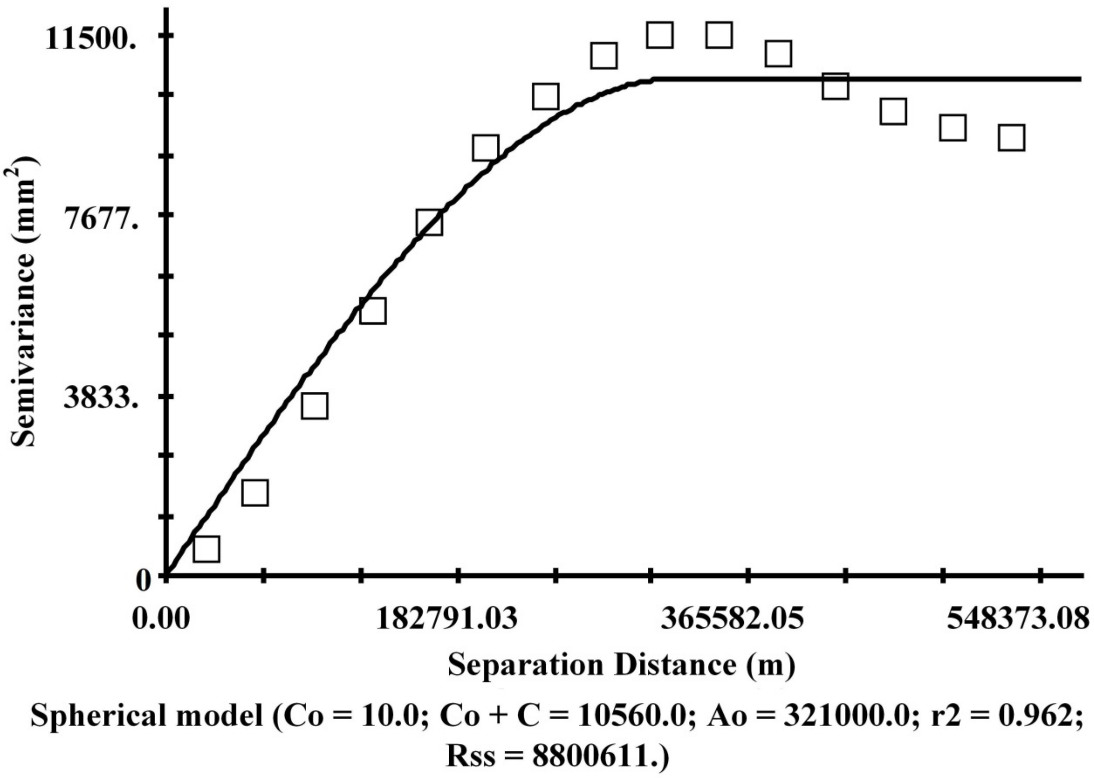
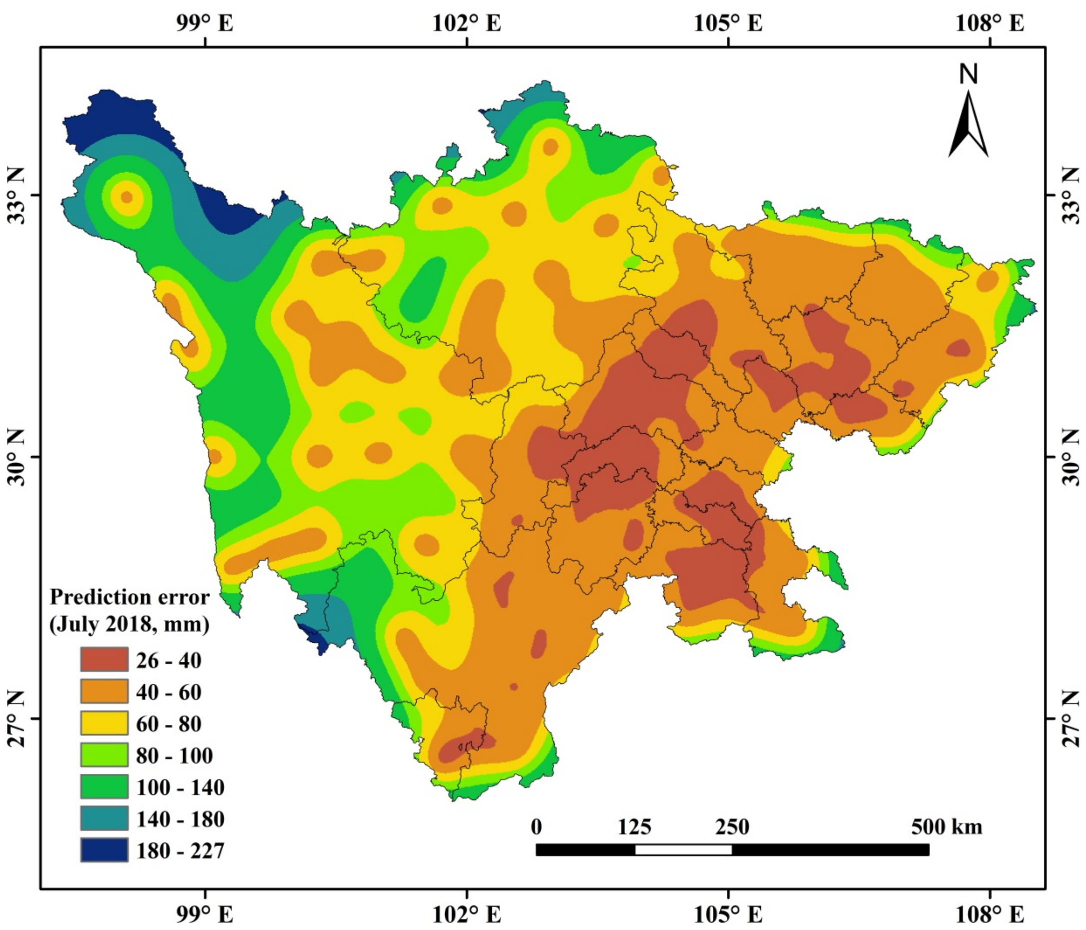


Fig. 8 Downscaled and calibrated precipitation maps of SRF-DC and some representative methods on the wettest month (July 2018)

The semivariogram and prediction error map of OK are shown in Fig. 9. Obviously, OK has a spherical model with the nugget variance (C_0) of 10.0 m^2, sill (C_0+C) of 10,560 m^2, residual sum of squares (Rss) of 8,800,611 m^2, range (A_0) of 321,000 m, and fitting R^2 of 0.962, respectively (Fig. 9a). The prediction error map (Fig. 9b) illustrates that the errors in the west are larger than in the east, and in the boundary are larger than in the inner. It can be inferred that large errors are mainly located in the areas with the sparse distribution of rain gauges. Moreover, the error magnitudes are not related to RMSE distribution (Fig. 7) and precipitation pattern (Fig. 8).



(a) Semivariogram



(b) Prediction error

Fig. 9 Semivariogram and prediction error map of kriging on the wettest month (July 2018)

5. Discussion

For downscaling and calibration of satellite-based precipitation, the three most important factors for constructing predictors-precipitation relationships are model, predictor and temporal scale (Chen et al., 2020b). Thus, they should be carefully selected to produce accurate precipitation data.

5.1. Model

In previous studies, the most commonly adopted model was GWR (Xu et al., 2015; Chen et al., 2015; Zhao et al., 2018), since it considers the spatial variation between the predictors and precipitation. However, due to the sparse distribution of rain gauge stations (Lu et al., 2019), GWR produced worse results than the original IMERG in the study region. RF and kriging outperformed GWR. Nevertheless, the two methods have some shortcomings. For example, the precipitation map of kriging was so smooth that many details were lost, and RF did not consider the spatial autocorrelation of precipitation measurements. In comparison, SRF-based methods with the consideration of spatial autocorrelation information demonstrated higher accuracy than the classical methods. Moreover, SRF-DC yielded slightly better results than Bi-SRF, SRF-GDA and SRFdis.

5.2. Environmental predictors

NDVI, latitude, longitude and DEM-based parameters were commonly adopted as predictors to estimate precipitation (Shi et al., 2015). However, satellite-based precipitation across regions with no relationship with NDVI could not be estimated, such as in barren or snow areas (Xu et al., 2015). Jing et al. (2016) indicated that the downscaled models including LST features (LSTs) performed better than those without LSTs. Thus, in addition to NDVI and DEM-related parameters, daytime LST (LST_D), nighttime LST (LST_N), and difference between day and night LSTs (LST_{D-N}) were used in this study.

Based on RF, the relative importance of each predictor (i.e. predictor importance estimate) is shown in Fig. 10. Obviously, precipitation from kriging interpolation has the most importance. This is because the interpolated value is directly related to precipitation. Kriging estimation is followed by the downscaled precipitation. Longitude is the third most important variable, which is followed by latitude. This result is consistent with that of Karbalaye Ghorbanpour et al. (2021). They indicated that compared to NDVI, LST and DEM, longitude ranks the first with respect to importance score.

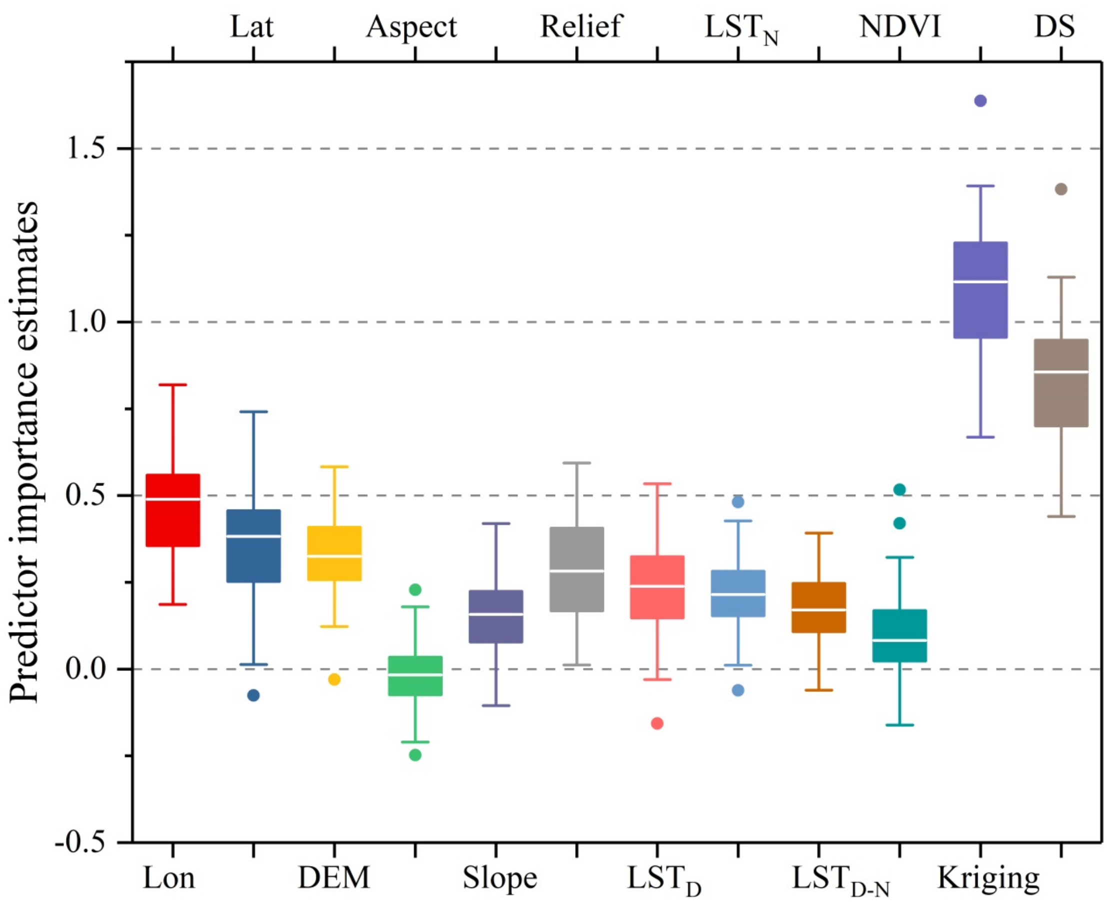


Fig. 10 Predictor importance estimates (Lat: latitude; Lon: longitude; DS: downscaled precipitation)

The three LSTs also have a great impact on the precipitation estimation, where LST_D seems slightly more important than LST_N and LST_{D-N}. NDVI has a slight effect on the precipitation, which ranks last but one. This might be due to the fact that NDVI is influenced by both precipitation and temperature in the study site, and the low temperature above certain elevations hinders the vegetation growth. It is less likely that the response of vegetation to precipitation has the delay, since SRF-DC on the monthly scale is more accurate than SRFdis on the annual scale.

Among the 12 predictors, aspect has the least importance. This conclusion was also obtained by Ma et al. (2017) for downscaling TMPA 3B43 V7 data over the Tibet Plateau. DEM, terrain relief and slope seem more important than aspect, since precipitation is closely related to topography (Jing et al., 2016). The results are consistent with previous studies (Immerzeel et al., 2009; Jing et al., 2016).

5.3. Temporal scale

Temporal scale has a great effect on the selection of predictors for precipitation estimation. There is a debate on whether NDVI should be taken as a predictor for downscaling and calibration of monthly precipitation. Some (Duan and Bastiaanssen, 2013; Immerzeel et al., 2009) argued that NDVI could not be used for monthly precipitation estimation since the response of NDVI to precipitation usually delayed for two or three months. However, some (Brunsell, 2006; Xu et al., 2015; Lu et al., 2019; Chen et al., 2020c) stated that the precipitation-NDVI relationship was hardly time-delayed, since vegetation could influence precipitation by adjusting temperature and air moisture during the growing seasons. Thus, it was possible to estimate precipitation with NDVI at the monthly scale. In this study, it was found that SRF-DC on the monthly scale was slightly more accurate than that on the annual scale (i.e. SRFdis). This indicates that the response of vegetation to precipitation has no obvious time delay, and NDVI can be used for monthly precipitation estimates.

5.4. Easy-to-use feature

Since the classical RF did not consider the spatial information in the modeling process, Hengl et al. (2018) proposed an improved RF for spatial estimation, where the buffer distances between the estimated location and measured locations were taken as the predictors. Motivated by this idea, Baez-Villanueva et al. (2020) presented a RF-based method (RF-MEP) for merging satellite precipitation products and rain gauge measurements, where the spatial distances from all rain gauges to the grid cells in the study site were used as the variables. However, as stated by Baez-Villanueva et al. (2020), RF-MEP has a huge computational cost, since the number of extra input features equals to that of gauge measurements. Moreover, RF-MEP ignores the spatial autocorrelation of precipitation between neighboring locations. In comparison, SRF only requires one extra feature that is estimated by kriging interpolation on the precipitation measurements. Thus, compared to the buffer distance layers-based RF, SRF is highly effective. Moreover, with the variogram-based kriging interpolation, the spatial autocorrelation of precipitation not only between the gauge locations, but also between the estimated location and gauge locations is taken into account. Thus, SRF has the merits of accuracy, effectivity and ease of use.

5.5. Limitations and further researches

Although SRF-DC shows promising results than the classical methods, it still suffers from some limitations, which should be solved in our further researches. Firstly, SRF-DC is more complex than Bi-SRF and SRF-GDA, since SRF is used in both downscaling and calibration. Applying SRF to downscale IMMERG might not be prerequisite since SRF-DC is only slightly better than Bi-SRF. However, SRF should be used to calibrate IMMERG due to the much higher accuracy of SRF-DC than SRF-GDA. Secondly, SRF-DC has low accuracy on high precipitation (e.g. >400 mm) since extreme precipitation is often caused by unpredictable factors. Thus, other available variables such as soil moisture (Fan et al., 2019; Brocca et al., 2019), and meteorological conditions such as cloud properties (Sharifi et al., 2019) could be adopted to further improve IMERG quality. Thirdly, the correction of satellite-based precipitation on higher-temporal scales (e.g. daily or hourly) is challenging (Wu et al., 2020; Chen et al., 2020b; Lima et al., 2021; Sun and Lan, 2021; Yan et al., 2021), since the relationships between environmental variables and precipitation on these scales are far less evident and difficult to capture. Although SRF-DC is general, its performance on these scales should be further assessed. Finally, numerous satellite-based precipitation products have been available, and each one has its shortcomings and advantages for the capture of spatial precipitation patterns (Chen et al., 2020c; Baez-Villanueva et al., 2020). Thus, the fusion of multiple precipitation products based on SRF-DC is an alternative to improve the quality of precipitation data.

6. Conclusions

To enhance the resolution (from 0.1° to 1 km) and accuracy of the monthly IMERG V06B Final Run product, a spatial RF (SRF)-based downscaling and calibration method (SRF-DC) was proposed in this study. The performance of SRF-DC was compared with those of seven methods including GWR, RF, BPNN, Bi-SRF, SRF-GDA, SRFdis and kriging on monthly IMERG from 2015 to 2019 over Sichuan province, China. The main findings and conclusions can be summarized as follows:

(1) The SRF-based methods including SRF-DC, Bi-SRF, SRF-GDA and SRFdis were more accurate than the classical methods. Moreover, SRF-DC performed slightly better than Bi-SRF and SRF-GDA.

(2) The comparison between the monthly-based and annual-based estimation demonstrated that there was no statistically significant difference between them, indicating that NDVI could be used for monthly precipitation estimation in the study site.

(3) Kriging outperformed the original IMERG, BPNN and GWR in terms of RMSE, MAE and CC. However, its interpolation map suffered from the serious loss of spatial precipitation patterns.

(4) Based on the variable importance assessment of RF, the precipitation interpolated by kriging on the gauge measurements was the most important variable, while terrain aspect was the least one. This indicated that considering spatial correlation was beneficial for precipitation estimation.

Overall, SRF-DC is general, robust, accurate and easy-to-use, as it shows promising results in the study area with heterogeneous terrain morphology and precipitation. Thus, it can be easily applied to other regions, where precipitation data with high resolution and high accuracy is urgently required.

Data availability

The gauge data are from the China Meteorological Data Service Center (http://data.cma.cn, last access: January 2021). The GPM data are from https://gpm.nasa.gov/data (last access: January 2021). The GPM data are from http://srtm.csi.cgiar.org/ (last access: January 2021). The MOD13A3 data are from http://www.gscloud.cn/ (last access: January 2021). The MOD11A2 data are from https://ladsweb.modaps.eosdis.nasa.gov (last access: January 2021).

Declaration of Competing Interest

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

Author contributions

CF and YY conceived the idea, and acquired the project and financial support. BJ conducted the detailed analysis. CF contributed to the writing and revisions.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant No. 41804001), Shandong Provincial Natural Science Foundation, China (Grant No. ZR2020YQ26, ZR2019MD007, ZR2019BD006), A Project of Shandong Province Higher Educational Youth Innovation Science and Technology Program (Grant No. 2019KJH007), Shandong Provincial Key Research and Development Program (Major Scientific and Technological Innovation Project) (Grant No. 2019JZZY010429) and by the Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents (Grant No. 2019RCJJ003).

References

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies, Bulletin of the American Meteorological Society, 96, 69-83, 10.1175/bams-d-13-00068.1, 2015.

Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Beck, H. E., McNamara, I., Ribbe, L., Nauditt, A., Birkel, C., Verbist, K., Giraldo-Osorio, J. D., and Xuan Thinh, N.: RF-MEP: A novel Random Forest method for merging gridded precipitation products and ground-based measurements, Remote Sensing of Environment, 239, 111606, https://doi.org/10.1016/j.rse.2019.111606, 2020.

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

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R., and Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment, Bulletin of the American Meteorological Society, 100, 473-500, 10.1175/bams-d-17-0138.1, 2019.

Belgiu, M., Drăguţ, L. J. I. J. o. P., and Sensing, R.: Random forest in remote sensing: A review of applications and future directions, 114, 24-31, 2016.

Berndt, C. and Haberlandt, U.: Spatial interpolation of climate variables in Northern Germany-Influence of temporal resolution and network density, Journal of Hydrology-Regional Studies, 15, 184-202, 10.1016/j.ejrh.2018.02.002, 2018.

Berndt, C., Rabiei, E., and Haberlandt, U.: Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios, Journal of Hydrology, 508, 88-101, https://doi.org/10.1016/j.jhydrol.2013.10.028, 2014.

Bhuiyan, M. A. E., Nikolopoulos, E. I., Anagnostou, E. N., Quintana-Seguí, P., and Barella-Ortiz, A.: A nonparametric statistical technique for combining global precipitation datasets: development and hydrological evaluation over the Iberian Peninsula, Hydrology and Earth System Sciences, 22, 1371-1389, 10.5194/hess-22-1371-2018, 2018.

Breiman, L.: Random Forests, Machine Learning, 45, 5-32, 2001.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., and Wagner, W.: SM2RAIN–ASCAT (2007–2018): global daily satellite rainfall data from ASCAT soil moisture observations, Earth System Science Data, 11, 1583-1601, 10.5194/essd-11-1583-2019, 2019.

Brunsell, N. A.: Characterization of land-surface precipitation feedback regimes with remote sensing, Remote Sensing of Environment, 100, 200-211, https://doi.org/10.1016/j.rse.2005.10.025, 2006.

Chao, L., Zhang, K., Li, Z., Zhu, Y., Wang, J., and Yu, Z.: Geographically weighted regression based methods for merging satellite and gauge precipitation, Journal of Hydrology, 558, 275-289, https://doi.org/10.1016/j.jhydrol.2018.01.042, 2018.

Cheema, M. J. M. and Bastiaanssen, W. G. M.: Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin, International Journal of Remote Sensing, 33, 2603-2627, 10.1080/01431161.2011.617397, 2012.

Chen, C., Yang, S., and Li, Y.: Accuracy Assessment and Correction of SRTM DEM using ICESat/GLAS Data under Data Coregistration, Remote Sensing, 12, 3435, 10.3390/rs12203435, 2020a.

Chen, C., Zhao, S., Duan, Z., and Qin, Z.: An Improved Spatial Downscaling Procedure for TRMM 3B43 Precipitation Product Using Geographically Weighted Regression, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, 4592-4604, 10.1109/JSTARS.2015.2441734, 2015.

Chen, F., Gao, Y., Wang, Y., and Li, X.: A downscaling-merging method for high-resolution daily precipitation estimation, Journal of Hydrology, 581, 124414, https://doi.org/10.1016/j.jhydrol.2019.124414, 2020b.

Chen, F., Liu, Y., Liu, Q., and Li, X.: Spatial downscaling of TRMM 3B43 precipitation considering spatial heterogeneity, International Journal of Remote Sensing, 35, 3074-3093, 10.1080/01431161.2014.902550, 2014.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, Water Resources Research, 49, 4187-4205, https://doi.org/10.1002/wrcr.20331, 2013.

Chen, S.-T., Yu, P.-S., and Tang, Y.-H.: Statistical downscaling of daily precipitation using support vector machines and multivariate analysis, Journal of Hydrology, 385, 13-22, https://doi.org/10.1016/j.jhydrol.2010.01.021, 2010.

Chen, S., Xiong, L., Ma, Q., Kim, J.-S., Chen, J., and Xu, C.-Y.: Improving daily spatial precipitation estimates by merging gauge observation with multiple satellite-based precipitation products based on the geographically weighted ridge regression method, Journal of Hydrology, 589, 125156, https://doi.org/10.1016/j.jhydrol.2020.125156, 2020c.

Chen, Y., Huang, J., Sheng, S., Mansaray, L. R., Liu, Z., Wu, H., and Wang, X.: A new downscaling-integration framework for high-resolution monthly precipitation estimates: Combining rain gauge observations, satellite-derived precipitation data and geographical ancillary data, Remote Sensing of Environment, 214, 154-172, https://doi.org/10.1016/j.rse.2018.05.021, 2018.

Duan, Z. and Bastiaanssen, W. G. M.: First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure, Remote Sensing of Environment, 131, 1-13, https://doi.org/10.1016/j.rse.2012.12.002, 2013.

Elnashar, A., Zeng, H., Wu, B., Zhang, N., Tian, F., Zhang, M., Zhu, W., Yan, N., Chen, Z., Sun, Z., Wu, X., and Li, Y.: Downscaling TRMM Monthly Precipitation Using Google Earth Engine and Google Cloud Computing, Remote Sensing, 12, 10.3390/rs12233860, 2020.

Fan, D., Wu, H., Dong, G., Jiang, X., and Xue, H.: A Temporal Disaggregation Approach for TRMM Monthly Precipitation Products Using AMSR2 Soil Moisture Data, Remote Sensing, 11, 10.3390/rs11242962, 2019.

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

Gebregiorgis, A. S. and Hossain, F.: Understanding the Dependence of Satellite Rainfall Uncertainty on Topography and Climate for Hydrologic Model Simulation, IEEE Transactions on Geoscience and Remote Sensing, 51, 704-718, 10.1109/TGRS.2012.2196282, 2013.

Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol. , 228, 113-129, 2000.

Haile, A. T., Habib, E., and Rientjes, T.: Evaluation of the climate prediction center (CPC) morphing technique (CMORPH) rainfall product on hourly time scales over the source of the Blue Nile River, Hydrological Processes, 27, 1829-1839, https://doi.org/10.1002/hyp.9330, 2013.

Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B., and Gräler, B. J. P.: Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables, PeerJ, 6, e5518, 2018.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T.: The Global Precipitation Measurement Mission, Bulletin of the American Meteorological Society, 95, 701-722, 10.1175/bams-d-13-00164.1, 2014.

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., and Joyce, R.: Algorithm theoretical basis document (ATBD) NASA global precipitation measurement (GPM) integrated multi-satellitE Retrievals for GPM (IMERG). Nasa (December), 29., 2019.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, Journal of Hydrometeorology, 8, 38-55, 10.1175/jhm560.1, 2007.

Immerzeel, W. W., Rutten, M. M., and Droogers, P.: Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula, Remote Sensing of Environment, 113, 362-370, https://doi.org/10.1016/j.rse.2008.10.004, 2009.

Jia, S., Zhu, W., Lű, A., and Yan, T.: A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China, Remote Sensing of Environment, 115, 3069-3079, 10.1016/J.RSE.2011.06.009, 2011.

Jing, W., Yang, Y., Yue, X., and Zhao, X.: A Spatial Downscaling Algorithm for Satellite-Based Precipitation over the Tibetan Plateau Based on NDVI, DEM, and Land Surface Temperature, Remote Sensing, 8, 10.3390/rs8080655, 2016.

Karbalaye Ghorbanpour, A., Hessels, T., Moghim, S., and Afshar, A.: Comparison and assessment of spatial downscaling methods for enhancing the accuracy of satellite-based precipitation over Lake Urmia Basin, Journal of Hydrology, 596, 126055, https://doi.org/10.1016/j.jhydrol.2021.126055, 2021.

Li, M. and Shao, Q.: An improved statistical approach to merge satellite rainfall estimates and raingauge data, Journal of Hydrology, 385, 51-64, https://doi.org/10.1016/j.jhydrol.2010.01.023, 2010.

Li, T., Shen, H., Yuan, Q., Zhang, X., and Zhang, L.: Estimating ground‐level PM2.5 by fusing satellite and station observations: a geo-intelligent deep learning approach, Geophysical Research Letters, 44, 11,985-911,993, 2017.

Li, Y., Zhang, Y., He, D., Luo, X., and Ji, X.: Spatial Downscaling of the Tropical Rainfall Measuring

Mission Precipitation Using Geographically Weighted Regression Kriging over the Lancang River Basin, China, Chinese Geographical Science, 29, 446-462, 10.1007/s11769-019-1033-3, 2019.

Lima, C. H. R., Kwon, H.-H., and Kim, Y.-T.: A Bayesian Kriging Model Applied for Spatial Downscaling of Daily Rainfall from GCMs, Journal of Hydrology, 126095, https://doi.org/10.1016/j.jhydrol.2021.126095, 2021.

Lu, X., Tang, G., Wang, X., Liu, Y., Wei, M., and Zhang, Y.: The Development of a Two-Step Merging and Downscaling Method for Satellite Precipitation Products, Remote Sensing, 12, 10.3390/rs12030398, 2020.

Lu, X., Tang, G., Wang, X., Liu, Y., Jia, L., Xie, G., Li, S., and Zhang, Y.: Correcting GPM IMERG precipitation data over the Tianshan Mountains in China, Journal of Hydrology, 575, 1239-1252, 10.1016/J.JHYDROL.2019.06.019, 2019.

Ma, Z., Shi, Z., Zhou, Y., Xu, J., Yu, W., and Yang, Y.: A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed, Remote Sensing of Environment, 200, 378-395, 2017.

Mohsenzadeh Karimi, S., Kisi, O., Porrajabali, M., Rouhani-Nia, F., and Shiri, J.: Evaluation of the support vector machine, random forest and geo-statistical methodologies for predicting long-term air temperature, ISH Journal of Hydraulic Engineering, 26, 376-386, 10.1080/09715010.2018.1495583, 2020.

Park, N.-W., Kyriakidis, P. C., and Hong, S.: Geostatistical Integration of Coarse Resolution Satellite Precipitation Products and Rain Gauge Data to Map Precipitation at Fine Spatial Resolutions, Remote Sensing, 9, 255, 2017.

Pham, B. T., Le, L. M., Le, T.-T., Bui, K.-T. T., Le, V. M., Ly, H.-B., and Prakash, I.: Development of advanced artificial intelligence models for daily rainfall prediction, Atmospheric Research, 237, 104845, https://doi.org/10.1016/j.atmosres.2020.104845, 2020.

Sharifi, E., Saghafian, B., and Steinacker, R.: Downscaling Satellite Precipitation Estimates With Multiple Linear Regression, Artificial Neural Networks, and Spline Interpolation Techniques, Journal of Geophysical Research Atmospheres, 124, 789-805, https://doi.org/10.1029/2018JD028795, 2019.

Shi, Y., Song, L., Xia, Z., Lin, Y., Myneni, R. B., Choi, S., Wang, L., Ni, X., Lao, C., and Yang, F.: Mapping Annual Precipitation across Mainland China in the Period 2001–2010 from TRMM3B43 Product Using Spatial Downscaling Approach, Remote Sensing, 7, 5849-5878, 2015.

Shortridge, A. and Messina, J.: Spatial structure and landscape associations of SRTM error, Remote Sensing of Environment, 115, 1576-1587, https://doi.org/10.1016/j.rse.2011.02.017, 2011.

Spracklen, D. V., Arnold, S. R., and Taylor, C. M.: Observations of increased tropical rainfall preceded by air passage over forests, Nature, 489, 282-285, 10.1038/nature11390, 2012.

Sun, L. and Lan, Y.: Statistical downscaling of daily temperature and precipitation over China using deep learning neural models: Localization and comparison with other methods, International Journal of Climatology, 41, 1128-1147, https://doi.org/10.1002/joc.6769, 2021.

Tao, Y., Gao, X., Hsu, K., Sorooshian, S., and Ihler, A.: A Deep Neural Network Modeling Framework to Reduce Bias in Satellite Precipitation Products, Journal of Hydrometeorology, 17, 931-945, 10.1175/jhm-d-15-0075.1, 2016.

Trenberth, K. E. and Shea, D. J.: Relationships between precipitation and surface temperature, 32, https://doi.org/10.1029/2005GL022760, 2005.

Ullah, S., Zuo, Z., Zhang, F., Zheng, J., Huang, S., Lin, Y., Iqbal, I., Sun, Y., Yang, M., and Yan, L.:

GPM-Based Multitemporal Weighted Precipitation Analysis Using GPM_IMERGDF Product and ASTER DEM in EDBF Algorithm, Remote Sensing, 12, 10.3390/rs12193162, 2020.

Wu, H., Yang, Q., Liu, J., and Wang, G.: A spatiotemporal deep fusion model for merging satellite and gauge precipitation in China, Journal of Hydrology, 584, 124664, https://doi.org/10.1016/j.jhydrol.2020.124664, 2020.

Wu, T., Feng, F., Lin, Q., and Bai, H.: Advanced Method to Capture the Time-Lag Effects between Annual NDVI and Precipitation Variation Using RNN in the Arid and Semi-Arid Grasslands, Water, 11, 1789, 2019.

Wu, Z., Zhang, Y., Sun, Z., Lin, Q., and He, H.: Improvement of a combination of TMPA (or IMERG) and ground-based precipitation and application to a typical region of the East China Plain, Science of The Total Environment, 640-641, 1165-1175, https://doi.org/10.1016/j.scitotenv.2018.05.272, 2018.

Xie, P. and Xiong, A.-Y.: A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD016118, 2011.

Xu, S., Wu, C., Wang, L., Gonsamo, A., Shen, Y., and Niu, Z.: A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics, Remote Sensing of Environment, 162, 119-140, https://doi.org/10.1016/j.rse.2015.02.024, 2015.

Xu, Y. and Goodacre, R.: On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning, J Anal Test, 2, 249-262, 10.1007/s41664-018-0068-2, 2018.

Yan, X., Chen, H., Tian, B., Sheng, S., Wang, J., and Kim, J.-S.: A Downscaling–Merging Scheme for Improving Daily Spatial Precipitation Estimates Based on Random Forest and Cokriging, Remote Sensing, 13, 2040, 2021.

Yang, Y. and Luo, Y.: Using the Back Propagation Neural Network Approach to Bias Correct TMPA Data in the Arid Region of Northwest China, Journal of Hydrometeorology, 15, 459-473, 10.1175/jhm-d-13-041.1, 2014.

Yang, Z., Hsu, K., Sorooshian, S., Xu, X., Braithwaite, D., Zhang, Y., and Verbist, K. M. J.: Merging high‐resolution satellite‐based precipitation fields and point‐scale rain gauge measurements—A case study in Chile, Journal of Geophysical Research: Atmospheres, 122, 5267-5284, 10.1002/2016JD026177, 2017.

Zhang, L., Li, X., Zheng, D., Zhang, K., Ma, Q., Zhao, Y., and Ge, Y.: Merging multiple satellite-based precipitation products and gauge observations using a novel double machine learning approach, Journal of Hydrology, 594, 10.1016/J.JHYDROL.2021.125969, 2021.

Zhang, X. and Tang, Q.: Combining satellite precipitation and long-term ground observations for hydrological monitoring in China, Journal of Geophysical Research: Atmospheres, 120, 6426-6443, https://doi.org/10.1002/2015JD023400, 2015.

Zhao, N., Yue, T., Chen, C., Zhao, M., and Fan, Z.: An improved statistical downscaling scheme of Tropical Rainfall Measuring Mission precipitation in the Heihe River basin, China, International Journal of Climatology, 38, 3309-3322, https://doi.org/10.1002/joc.5502, 2018.

Zhao, T. and Yatagai, A.: Evaluation of TRMM 3B42 product using a new gauge-based analysis of daily precipitation over China, International Journal of Climatology, 34, 2749-2762, https://doi.org/10.1002/joc.3872, 2014.

717