Heliyon 10 (2024) e36368

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

# Assessing a machine learning-based downscaling framework for obtaining 1km daily precipitation from GPM data

Tao Sun [a], Nana Yan [b], Weiwei Zhu [b,**], Qifeng Zhuang [a,*]

[a] College of Geomatics Science and Technology, Nanjing Tech University, Nanjing, 211816, China
[b] Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences. Beijing 100101, China

ARTICLE INFO

*Keywords:*
Spatiotemporal downscaling
Feature factor
Machine learning
Downscaling framework

ABSTRACT

Hydro-meteorological monitoring through satellites in arid and semi-arid regions is constrained by the coarse spatial resolution of precipitation data, which impedes detailed analyses. The objective of this study is to evaluate various machine learning techniques for developing a downscaling framework that generates high spatio-temporal resolution precipitation products. Focusing on the Hai River Basin, we evaluated three machine learning approaches—Extreme Gradient Boosting (XGBoost), Random Forest (RF), and Back Propagation (BP) neural networks. These methods integrate environmental variables including land surface temperature (LST), Normalized Difference Vegetation Index (NDVI), Digital Elevation Model (DEM), Precipitable Water Vapor (PWV), and albedo, to downscale the 0.1° spatial resolution Global Precipitation Measurement (GPM) product to a 1 km resolution. We further refined the results with residual correction and calibration using terrestrial rain gauge data. Subsequently, utilizing the 1 km annual precipitation, we employed the moving average window method to derive monthly and daily precipitation. The results demonstrated that the XGBoost method, calibrated with Geographical Difference Analysis (GDA) and Kriging spatial interpolation, proved to be the most accurate, achieving a Mean Absolute Error (MAE) of 58.40 mm for the annual product, representing a 14 % improvement over the original data. The monthly and daily products achieved MAE values of 11.61 mm and 1.79 mm, respectively, thus enhancing spatial resolution while maintaining accuracy comparable to the original product. In the Hai River Basin, key factors including longitude, latitude, DEM, LST_night, and PWV demonstrated greater importance and stability than other factors, thereby enhancing the model's precipitation prediction capabilities. This study provides a comprehensive assessment of the annual, monthly, and daily high-temporal and high-spatial resolution downscaling processes of precipitation, serving as an important reference for hydrology and related fields.

## 1. Introduction

Precipitation is a critical component of the global water cycle, playing a vital role in ecological, meteorological, and hydrological studies [1,2]. Accurate precipitation measurement is essential in various scientific and applied fields such as agriculture, flood

* Corresponding author.
** Corresponding author.
*E-mail addresses:* zhuww@aircas.ac.cn (W. Zhu), zhuangqf@njtech.edu.cn (Q. Zhuang).

https://doi.org/10.1016/j.heliyon.2024.e36368
Received 17 January 2024; Received in revised form 14 August 2024; Accepted 14 August 2024
Available online 22 August 2024

forecasting, hydrological process simulation, and water resource monitoring [3,4]. Traditional precipitation measurements, such as those based on rain gauges, fail to capture the spatiotemporal variations of precipitation, and their sparse distribution in some areas affects interpolation results [5]. The increasing frequency of extreme precipitation events in recent years have heightened attention to the consequences of extreme precipitation [6], making it particularly important to obtain more accurate precipitation data in areas with frequent extreme weather and sparse rain gauge coverage.

Compared to traditional methods, satellite remote sensing technology for precipitation measurement offers advantages such as high spatial and temporal resolution, extensive coverage, and independence from topographical constraints. It has become a vital approach for acquiring precipitation data. Satellite Precipitation Estimation (SPE) products, which are a significant source of global precipitation data, have garnered increasing attention. Since 1980, various SPE products have been released, including the TRMM Multi-Satellite Precipitation Analysis (TMPA) [7], the Integrated Multi-satellite Retrievals for GPM (IMERG) [8], and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR) [9]. However, these existing satellite precipitation products, with spatial resolutions ranging from 0.1° to 0.25°, are too coarse for watershed-scale hydrological modeling. This necessitates spatial downscaling to obtain high-quality precipitation products.

Spatial downscaling techniques are primarily categorized into two types: dynamic downscaling and statistical downscaling. Dynamic downscaling is based on the mathematical representation of physical processes and considers the interactions among various complex atmospheric, oceanic, and terrestrial processes. However, this method is computationally intensive and demands extensive data, which limits its practical application [10]. In contrast, statistical downscaling establishes empirical statistical relationships between target and auxiliary variables, offering higher efficiency [11]. This approach is more widely applied in spatial downscaling.

A variety of regression models have been employed to utilize hydrometeorological variables as predictors, providing downscaled outcomes [12]. For instance, the correlation between precipitation and the Normalized Difference Vegetation Index (NDVI) at varying spatial resolutions was initially explored using an Exponential Regression (ER) model [13]. Jia et al. established the relationships among precipitation, vegetation, and topography using a Multiple Linear Regression (MLR) model [14]. Xu et al. developed a Geographically Weighted Regression (GWR) model for high-resolution precipitation estimation based on NDVI and the Digital Elevation Model (DEM) [15], while Li et al. compared two models, Geographic Weighted Regression (GWR) and its modified version, Multiscale Geographic Weighted Regression (MGWR), based on Relative Humidity (RHU), Wind Speed (WIN), and Sunshine Duration (SSD) [16]. Given the complex nonlinear relationships between precipitation and surface features, machine learning techniques are more suitable for downscaling precipitation at large spatial scales than traditional statistical regression algorithms [17]. Ma et al. applied the Random Forest model using NDVI, LST, and topographic data to obtain annual downscaled results (1 km) of IMERG and TMPA over the Tibetan Plateau [18]. Zhang et al. used Random Forest to link satellite precipitation estimates from seven precipitation zones in the Lancang-Mekong River Basin with predictors like latitude, longitude, DEM, aspect, slope, and NDVI at a 1 km resolution [19], Shen et al. conducted spatial downscaling in China using the Gradient Boosting Decision Tree (GBDT) method combined with NDVI and LST [20], Wang et al. downscaled satellite precipitation products to 1 km using a neural network model combined with the Enhanced Vegetation Index (EVI), LST, Evapotranspiration (ET), DEM, slope, and geographical coordinates [21].

Most downscaling studies have traditionally been conducted on an annual scale [22]. However, when analyzing precipitation on daily and monthly scales, it is crucial to consider the lagged response of NDVI to precipitation [23], which can extend up to 2–3 months [24]. This suggests that direct application of NDVI downscaling methods on monthly and daily scales presents greater challenges. Xu et al. developed a daily scale GWR model based on the relationship between precipitation and cloud properties: Cloud Effective Radius (CER), Cloud Optical Thickness (COT), and Cloud Water Path (CWP). They obtained daily precipitation data by calibrating the downscaled results using a Geographical Difference Analysis (GDA) method [25]. Zeng et al. have developed a novel remote sensing precipitation downscaling technique, named Local Precipitation Downscaling based on Vegetation Index Adaptive Lag phase (LPVIAL) using a newly developed Vegetation Index based on Adaptive Lag Phase (VIAL) [26]. By employing the NDVI, land surface coverage type (MCD12Q1), and rain gauge data as auxiliary parameters, this technique was applied to IMERG Final Run daily precipitation products for the Yangtze River Delta region from 2010 to 2017. The technique has successfully downscaled these products from a resolution of 0.1°–1 km, and subsequently upscaled the time resolution to 16-day intervals, which is a decrease in temporal resolution compared to the original data. Zandi et al. used model stacking and local weighted linear regression to generate high-resolution monthly precipitation data, and compared the applicability of these two methods in complex regions [27]. Kumar et al. employed three Deep learning methods—super-resolution convolutional neural network ( SRCNN ) , stacked SRCNN, and DeepSD—to downscale summer monsoon data from the India Meteorological Department (IMD) and TRMM datasets, achieving a fourfold reduction in resolution. They found that the DeepSD method yielded the best results [28]. In subsequent studies, they continued to use SR-GAN along with the other three methods (DeepSD, ConvLSTM, and UNET) to downscale IMD gridded products. Validation analysis showed that the SR-GAN method outperformed the other three methods by leveraging the concept of perceptual loss, overcoming the problem of smooth reconstruction, and thus capturing finer details of the data under consideration [29].

In the field of precipitation downscaling methods, identifying a suitable framework is crucial. Notable studies include Zhu et al., who evaluated various machine learning methods for annual-scale results, emphasizing the importance of residual correction and rain gauge calibration [30]. Karbalaye Ghorbanpour et al. devised a downscaling framework for both annual and monthly scales, although they omitted discussions on residual correction [31]. Moreover, Nasseri et al. established a monthly downscaling framework that relies on rain gauge calibration for multi-source precipitation products, significantly improving the accuracy of original rainfall data [32]. Furthermore, Zandi et al. developed a monthly downscaling framework that merges precipitation products with surface features, successfully producing high-precision monthly precipitation maps in topographically complex regions of Iran using the Area-to-Point Kriging (ATPK) method and Local Weighted Linear Regression (LWLR) [33]. However, previous downscaling frameworks have not simultaneously addressed the completeness of calibration, the spatial non-stationarity relationships between precipitation and

environmental factors, and the temporal resolution aspects.

Despite significant advancements in precipitation downscaling methods, there is still a lack of a comprehensive evaluation framework. Such a framework would enable researchers to effectively utilize a combination of satellite precipitation data and ground-based rain gauge observations, achieving high spatial resolution downscaling of precipitation products annually, monthly, and daily. To fill this gap, we have utilized Integrated Multi-satellite Retrievals for GPM (IMERG) data spanning from 2003 to 2018 for downscaling purposes. The goal is to assess and validate the adaptability of the downscaling framework to produce high-resolution precipitation data, both temporally and spatially. By incorporating variables such as Precipitable Water Vapor (PWV), albedo, and others as predictive factors, we applied machine learning models including XGBoost, Random Forest, and Back Propagation Neural Networks. Initially, spatial downscaling is performed on an annual scale, with subsequent corrections of the results. For monthly and daily data, a moving window technique is employed to segment the annual data into more granular datasets. We conducted a comprehensive evaluation of this downscaling framework, aimed at acquiring accurate and reliable multi-scale, high-resolution precipitation products, which provide robust support for climate research and water resource management in the region.

## 2. Materials and method

### 2.1. Study area

The Haihe River Basin, spanning between 112° and 120°E longitude and 35° and 43°N latitude, covers a vast area of 318,200 square kilometers and constituting 3.3 % of China's total land area. The basin's diverse topography includes hills, undulating mountains, alluvial plains, and fluvial terraces, which contribute to its complex climatic patterns [34]. Located within the temperate semi-humid and semi-arid continental monsoon climate zone, the region endures dry winters and springs with minimal rainfall, hot and humid summers, and periods of concentrated heavy rains. The average annual precipitation increases from southwest to northeast, typically ranging from 600 to 700 mm. Influenced by the Taihang Mountains, the windward slopes of the Wutai and Yanshan mountains receive higher rainfall, with peaks reaching 750–800 mm. In contrast, the northwest part of the basin receives the least precipitation, with amounts as low as 350 mm. This significant seasonal and interannual variability in rainfall results in frequent alternating periods of drought and flooding within the basin [35]. Fig. 1 illustrates the boundary and elevation map of the Haihe River Basin.

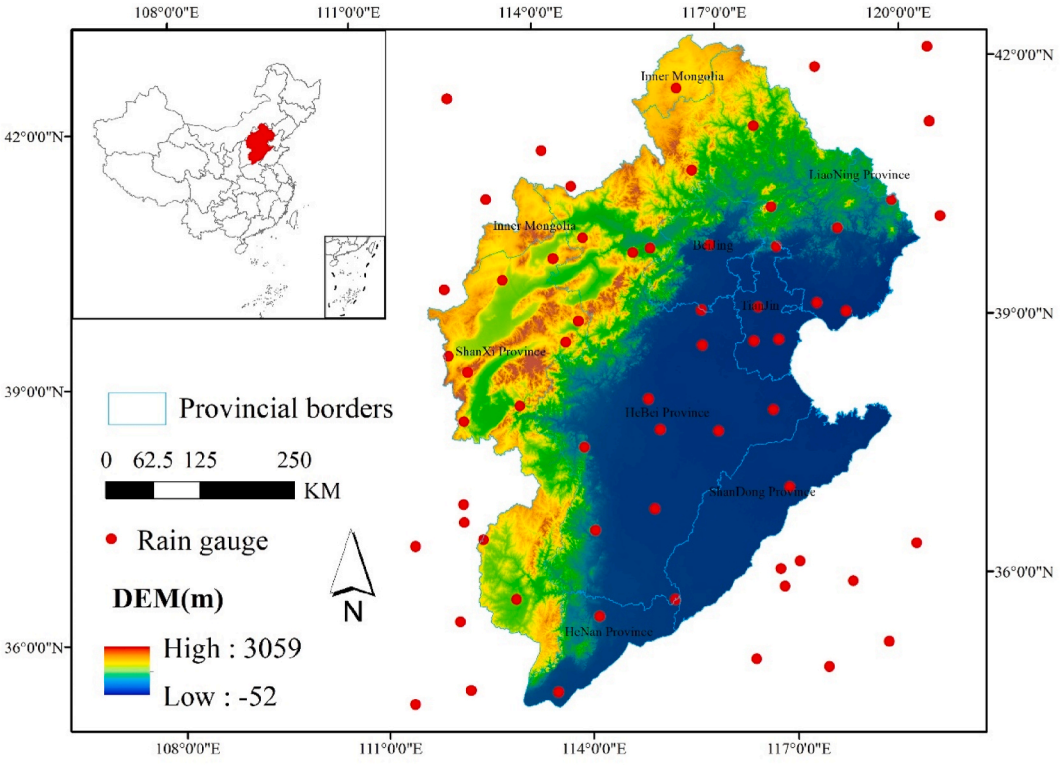


**Fig. 1.** The location of the Haihe River Basin and the distribution of rain gauges.

### 2.2. Dataset

In this study, we utilized the China Surface Climate Data Daily Dataset (SURF_CLI_CHN_MUL_DAY_V3.0), which is collected from meteorological stations across mainland China. The distribution of these stations within the Haihe River Basin is notably uneven. To ensure the validity of our validation results, we included data from 39 stations within the basin and 24 peripheral stations. Furthermore, to compute the annual precipitation at each rain gauge station, we aggregated daily precipitation measurements into annual totals. The locations of these meteorological stations are depicted in Fig. 1.

The Global Precipitation Measurement (GPM) mission, an international satellite initiative, includes the Integrated Multi-satellite Retrievals for GPM (IMERG), which serves as GPM's Level-3 multi-satellite precipitation algorithm. This algorithm integrates microwave and infrared data with precipitation measurements from the GPM constellation to produce a consistent, global, gridded precipitation product. The GPM IMERG V06B product provides precipitation data with a spatial resolution of 0.1°, available in half-hourly, daily, and monthly formats. In our study, we utilized daily and monthly data, aggregating them to compute annual precipitation values.

The Moderate-resolution Imaging Spectroradiometer (MODIS) provides a wide range of products, which were extensively utilized in this study to derive most of the environmental variables. Specifically, the annual mean NDVI values were calculated from the monthly averages of the MOD13A3 product. Likewise, the annual mean LST was calculated by averaging the eight-day interval data from MOD11A2, providing the annual averages for both Land Surface Temperature during Daytime (LST_Day) and Land Surface Temperature during Nighttime (LST_Night). Albedo and Precipitable Water Vapor (PWV) values were obtained from MOD09GA and MOD05, respectively. Annual mean albedo values were derived using the Google Earth Engine (GEE). In a manner similar to other data treatments, MOD05 data was compiled monthly to calculate annual averages. Additional environmental variables like the DEM were acquired from the Shuttle Radar Topography Mission (SRTM) dataset, with a spatial resolution of 90m in TIF format. This DEM data was resampled to a 1 km resolution using ArcGIS, facilitating the computation of slope, aspect, and geographic coordinates. Detailed information about the datasets used in this study is summarized in Table 1.

### 2.3. Downscaling framework

XGBoost, proposed by Chen, is a machine learning algorithm designed to sequentially combine weak learners into a robust final model [36]. Each base learner in the sequence learns from its predecessors and reduces their errors, resulting in a final learner that exhibits minimal bias and variance during training [37].

Random Forest, a tree-based ensemble learning technique introduced by Breiman, enhances model accuracy and stability by combining multiple decision trees. This method integrates the Bagging concept with the random selection of features to create a robust ensemble [38]. Recognized for its resistance to noise, Random Forest effectively prevents overfitting through parameter adjustments [39] and has been widely applied in various research studies [20,30,40–42].

The Backpropagation (BP) Neural Network is a multi-layer, feedforward network that incorporates activation functions to enhance performance. This network optimizes the weights and thresholds of neurons in each layer by backpropagating errors calculated during training. Thanks to their ability to autonomously adjust the weights of input variables through learning, BP neural networks are known for their exceptional fitting capabilities. BP neural networks have been utilized in downscaling endeavors in previous research [43].

The annual downscaling process employs three machine learning algorithms and is implemented in a Python environment. Specifically, we use scikit-learn for XGBoost and Random Forest, and TensorFlow's tf.keras API for the Backpropagation Neural Network. The process is detailed in Fig. 2 and encompasses the following steps.

1. Spatial resolution NDVI, LST, PWV, and Albedo data, along with 90m resolution DEM, were resampled to a 10 km spatial resolution using a geographic coordinate system. During this process, any negative or invalid values were excluded from the environmental variables.
2. The environmental variables initially at 1 km resolution were resampled to 10 km using the nearest-neighbor method. At the 10 km resolution, the annual NDVI, LST, and PWV were computed by averaging their respective monthly NDVI values, 8-day LST readings, and monthly PWV measurements taken within that year. Additionally, the annual mean albedo was calculated using Google Earth Engine (GEE). The aim here is to convert all auxiliary variables to an annual scale.

**Table 1**
Datasets used in this study.

| Data Type | Products | Dataset | Resolution |
|---|---|---|---|
| Meteorological data | Precipitation | GPM_IMERG | 0.1°, daily |
| | Precipitation | Station Observation | -,daily |
| Environmental variables | Normalized Difference Vegetation Index(NDVI) | MOD13A3 | 1 km, 16 days |
| | Precipitable Water Vapor(PWV) | MOD05 | 1 km, daily |
| | Land Surface Temperature(LST) | MOD11A2 | 1 km, 8 days |
| | Albedo | MOD09GA | 500m, daily |
| | DEM | SRTM | 90 m, - |

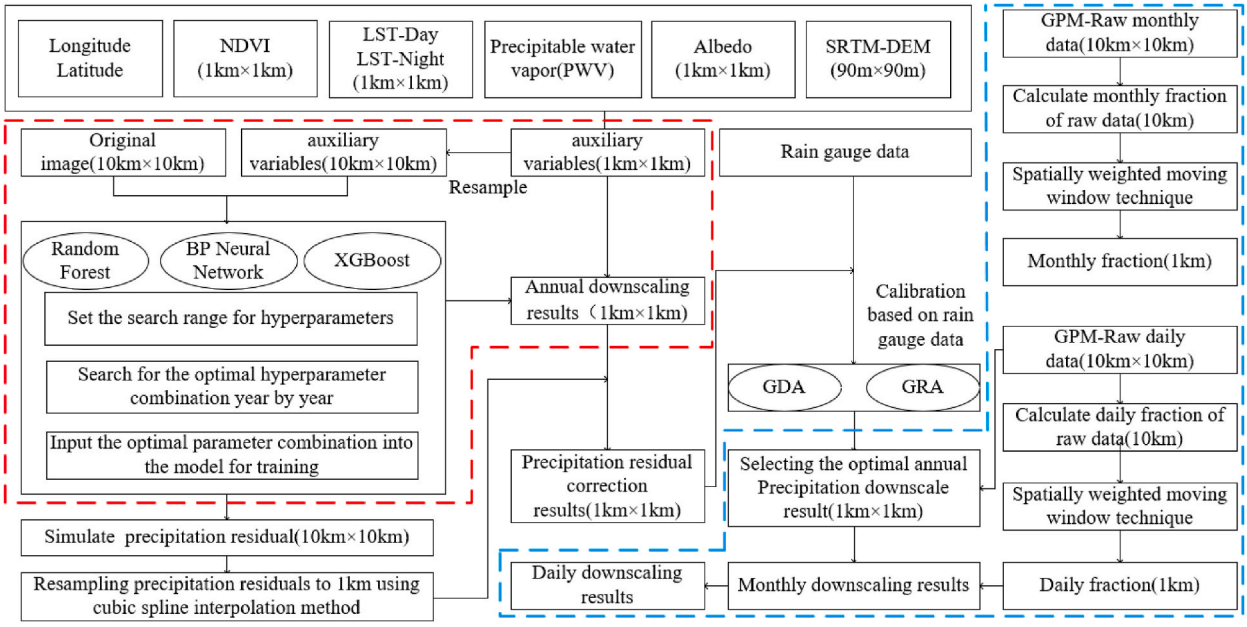


**Fig. 2.** Flow chart for downscaling and calibration using three machine learning methods: XGBoost, Random Forest, and BP neural network.(The content within the red dashed box represents the annual downscaling process, while the content within the blue dashed box represents the monthly and daily downscaling process.).

3. The downscaling model was trained with a series of annual environmental variables at a 10 km resolution, including IMERG precipitation, NDVI, LST, PWV, Albedo, and DEM, as well as latitude and longitude information, this training was conducted through the XGBoost, RF, and BP Neural Network techniques.
4. Hyperparameter optimization is a crucial aspect of XGBoost and RF, involving adjustments of parameter combinations to prevent model overfitting. For both machine learning models, hyperparameter optimization employs three common methods: Grid Search Algorithm, Bayesian Algorithm, and Genetic Algorithm. The optimal combination is determined annually through 10-fold cross-validation. In the Random Forest model, optimized parameters include n_estimators (number of decision trees), max_depth (maximum depth of the tree), min_samples_split (minimum number of samples required to split a node), min_samples_leaf (minimum number of samples required at a leaf node), and max_features (number of features). Parameters optimized in the XGBoost model include max_depth, n_estimators, learning_rate(the rate at which the model is learning, with a lower rate requiring more boosting rounds but often improving the model's generalization), gamma (penalty term for complexity), reg_alpha (L1 regularization that minimizes the absolute values of coefficients, potentially setting some to zero to enhance feature selection), reg_lambda (L2 regularization that minimizes the squared values of coefficients, reducing their magnitudes to prevent overfitting while retaining all features), min_child_weight (minimum sum of instance weight needed in a child node to further partition it, used to control over-fitting), colsample_bytree (proportion of features sampled for each tree), and subsample(the fraction of the training data used for growing each tree, used to increase randomness and prevent overfitting). The annual hyperparameter optimization results are summarized in Tables S1 and S2.
5. The activation function chosen for the BP neural network is the efficient RELU function. Despite its advantages, the BP network has limitations, such as its inability to perform global searches and its susceptibility to falling into local minima. To overcome these limitations, this study integrates the Genetic Algorithm (GA). GA leverages genetic operations such as selection, mutation, and crossover, facilitating the selection of high-fitness individuals in each generation. This approach effectively mitigates the local minima issue inherent in BP neural networks and enhances their nonlinear expression capabilities. Consequently, the GA-BP neural network is adopted as the method for downscaling model construction.
6. The downscaling models developed from the three algorithmic approaches are applied to 1 km resolution environmental variables, producing high spatial resolution IMERG outputs.

### *2.4. Annual data decomposition scheme*

Based on the annual results obtained from three machine learning models [discussed further in Section 3.1], we conducted a verification analysis and selected datasets with more accurate annual results. These selected datasets were subsequently decomposed into daily and monthly data using a spatially weighted moving window approach [44].Taking the monthly data calculation as an example: First, we calculated the monthly fractions of the original data by comparing each month's figures with the annual data (as shown in Equation (1)). These monthly fraction data were then resampled to a 1 km resolution using the nearest-neighbor approach.

Following this step, a moving window technique was applied across all the monthly fraction data, calculating the number of adjacent pixels and their respective areas. The monthly fractions at 1 km resolution were determined using a specific formula (as shown in Equation (2)). Finally, these results were multiplied with the downscaled 1 km resolution annual data to yield the 1 km resolution monthly data. The formula for calculating the original data's monthly fractions is as follows:

$$F(10km) = \frac{IMERG_{original,i}}{\sum_{i=1}^{12} IMERG_{original,i}} \quad (1)$$

In equation (1), $F(10km)$ represents the original monthly fraction, $IMERG_{original,i}$ denotes the original precipitation data for the i-th month, and $\sum_{i=1}^{12} IMERG_{original,i}$ signifies the summation of monthly data for the year in which the i-th month falls.

The calculation formula for the monthly results at a 1 km resolution is as follows:

$$F(1km) = \frac{1}{\sum_{1}^{n} S_i} \sum_{1}^{n} S_i \times F(10km)_i \quad (2)$$

In equation (2), $F(1km)$ represents the monthly fraction at 1 km, n denotes the number of pixels covered by the moving window, $S_i$ indicates the area of each pixel under the coverage of the moving window, and $F(10km)_i$ signifies the pixel value corresponding to the covered window location.

The approach to decomposing daily data closely follows the technique used for monthly data analysis. This process starts with the calculation of a daily fraction for each day and then iteratively traverses the entire image. Using a specific formula, a 1 km score is determined and subsequently multiplied with the equivalent monthly 1 km data to derive the final dataset. we implemented this method using the rasterio and numpy libraries in the Python environment. For a more detailed understanding of the method, we recommend referring to the original literature.

### 2.5. Residual correction and rain gauge calibration

Initially, residual analysis is performed using three machine learning models—XGBoost, Random Forest (RF), and Backpropagation (BP) Neural Network—to simulate data at 10 km resolution and calculate estimated precipitation. The discrepancies between these estimates and the original coarse resolution data are then refined to 1 km resolution using cubic spline interpolation. This refined data is adjusted by subtracting it from the initial downscaled results to derive the final refined dataset.

Geographical Difference Analysis (GDA), developed by Cheema et al., employs residual analysis to minimize discrepancies between downscaled results and rain gauge measurements [45]. In this study, we strategically selected 31 sites both within and beyond the study area as interpolation points, and 32 sites as validation points post-interpolation. The rain gauge data at each location is calculated annually and interpolated from the original data using kriging interpolation and inverse distance weighting interpolation at a 1 km resolution. These interpolations were conducted in batch using the arcpy library in Python. The finally data is obtained by adding the interpolated result to the downscaled data. Detailed calculation formulas and their explanations can be found in Appendix A (Equations S1-S3).

Geographical Ratio Analysis (GRA) is another calibration method that is widely employed in the field [11,44], Similar in its foundational approach to Geographical Difference Analysis (GDA), GRA's primary distinction lies in its use of ratio analysis for interpolation purposes. This method involves multiplying the downscaled results by the interpolated data to produce the final output. Detailed computation formulas and their explanations can be found in Appendix B (Equations S4-S6).

### 2.6. Validation

Hyperparameter optimization employs 10-fold cross-validation to evaluate the downscaling model's performance. We segment the original IMERG and corresponding environmental data into ten subsets, using nine for model training and one to assess predictive performance. This cross-validation cycle is repeated ten times, ensuring comprehensive validation of each subset. Model accuracy is gauged using three key metrics: the coefficient of determination ($R^2$), root mean square error (RMSE), and Kling-Gupta efficiency (KGE). $R^2$ quantifies the model's explanatory power regarding observed data variability. RMSE measures the quality of the model fit, and KGE, by integrating correlation, bias, and variation coefficients, offers a holistic assessment of the downscaled precipitation data's performance. Detailed methodologies for these metrics can be found in Appendix C (Equations S7-S10).

In order to further validate the accuracy of downscaling products, we also verified the downscaling results using rain gauge data. In addition to employing $R^2$, RMSE, and KGE as metrics, we also used the Mean Absolute Error (MAE) to evaluate our models. The computation method for MAE can be found in Appendix D (Equation S11).

In the daily scale assessment, we have added several metrics to the existing evaluation indices to assess the ability of the daily products to capture precipitation events. These metrics include the Probability of detection (POD), the false alarm ratio (FAR), and the critical success index (CSI). The computation methods for these metrics can be found in Appendix E (Equations S12-S14).

## 3. Results and discussion

### 3.1. Annual scale analysis

Fig. 3 features a scatter plot that demonstrates the model's ability to fit the data, displaying predictions from the original IMERG dataset for the years 2003–2018 at a 10 km spatial resolution, using the Random Forest, XGBoost, and BP Neural Network algorithms. Notably, the XGBoost and BP Neural Network models exhibit excellent fitting capabilities, with their scatter plots closely adhering to the 1:1 line. In contrast, the Random Forest model displays somewhat weaker performance, particularly tending to underestimate precipitation in the range of 800 mm–1200 mm. Table 2 summarizes the optimal models derived from a 10-fold cross-validation conducted over the years 2003–2018, evaluating their precision on test sets. Despite undergoing hyperparameter optimization, the annual accuracy results clearly demonstrate that the XGBoost and BP Neural Network models outperform the Random Forest in terms of fit, as evidenced by their significantly lower error rates. Although the XGBoost model and BP Neural Network show comparable accuracies, the XGBoost demonstrates a higher degree of fit, with KGE values ranging between 0.908 and 0.996, outstripping the BP Neural Network's precision. Regarding root mean square error, the XGBoost model records a notably lower error, with just 10.61 mm recorded in 2014, underscoring its superior effectiveness in simulating annual precipitation of IMERG based on environmental variables at a 10 km scale.

After evaluating various models' ability to simulate the 10 km IMERG precipitation data, we further assessed the practicality of the downscaling models by corroborating their accuracy with rain gauge verification. This analysis included validation against 63 sites distributed both within and outside the Haihe River Basin. Fig. 4 presents scatter plots of the original data (4A) and the downscaled results from three methods—XGBoost (4B), Random Forest (4C), and BP Neural Network (4D)—each validated against gauge data from 2003 to 2018, generally showing that precipitation predominantly ranged between 400 mm and 600 mm. Notably, the GPM data displayed pronounced overestimation, consistent with findings reported by several scholars [18,20,46]. Despite machine learning models' capability to simulate precipitation, original results remained superior, achieving an RMSE of 97.47 mm, MAE of 73.08 mm, an $R^2$ of 0.75, and a KGE of 0.8. These metrics suggest that while machine learning techniques can simulate precipitation effectively, they do not necessarily enhance the quality of results. Fig. 5 presents the rain gauge data as validation, plotting the Taylor diagram of the three models and the original data, from which it's evident that the original data aligns closely with the validation values. Comparing the three machine learning methods, the XGBoost algorithm yields the best outcomes, similar to the original data, followed by the BP Neural Network model, with the RF model lagging behind the others. This pattern mirrors the precision of the training model results and also indicates that the model training process did not result in overfitting.

Fig. 6 presents annual Taylor statistical diagrams, showing that in certain years, such as 2005, 2007, and 2014, the simulation results of the three models closely approached those of the original data. This indicates that in these years, the models performed well, matching the original data. In other years, the Random Forest model demonstrated clear differences from the other models and the original data, with its simulation results being less accurate than those of the other two models. However, the XGBoost and BP Neural Network models in most years closely matched the original data and were more akin to rain gauge data compared to the RF model, indicating their superior downscaling performance. In 2017 and 2018, XGBoost slightly outperformed the BP Neural Network model.

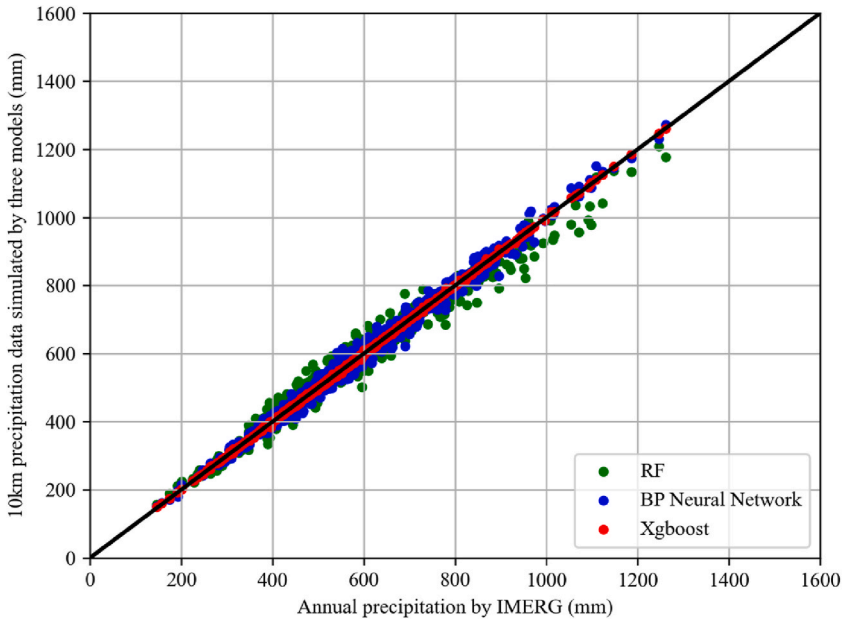


**Fig. 3.** Scatter plot of simulation results from different models.

**Table 2**
Training accuracy of random forest, XGBoost and BP neural network models.

| Method | Random Forest | | | XGBoost | | | GA-BP | | |
|---|---|---|---|---|---|---|---|---|---|
| Year | $R^2$ | KGE | RMSE (mm) | $R^2$ | KGE | RMSE ( mm ) | $R^2$ | KGE | RMSE ( mm ) |
| 2003 | 0.995 | 0.885 | 13.22 | 0.997 | 0.970 | 13.86 | 0.995 | 0.965 | 16.38 |
| 2004 | 0.992 | 0.875 | 15.97 | 0.992 | 0.988 | 15.12 | 0.993 | 0.994 | 15.5 |
| 2005 | 0.994 | 0.990 | 20.89 | 0.993 | 0.976 | 17.86 | 0.991 | 0.974 | 20.77 |
| 2006 | 0.985 | 0.852 | 21.2 | 0.984 | 0.956 | 17.63 | 0.983 | 0.935 | 18.96 |
| 2007 | 0.991 | 0.979 | 22.67 | 0.993 | 0.952 | 21.74 | 0.99 | 0.952 | 21.46 |
| 2008 | 0.989 | 0.989 | 15.95 | 0.993 | 0.996 | 14.06 | 0.993 | 0.988 | 13.67 |
| 2009 | 0.996 | 0.926 | 13.5 | 0.997 | 0.977 | 12.36 | 0.993 | 0.975 | 12.9 |
| 2010 | 0.985 | 0.965 | 16.76 | 0.989 | 0.992 | 14.19 | 0.989 | 0.994 | 14.26 |
| 2011 | 0.992 | 0.918 | 17.57 | 0.994 | 0.986 | 15.31 | 0.99 | 0.990 | 16.32 |
| 2012 | 0.991 | 0.841 | 13.84 | 0.992 | 0.985 | 13.11 | 0.994 | 0.977 | 12.94 |
| 2013 | 0.985 | 0.986 | 14.23 | 0.991 | 0.969 | 11.94 | 0.99 | 0.971 | 12.37 |
| 2014 | 0.979 | 0.932 | 15.68 | 0.989 | 0.908 | 10.61 | 0.984 | 0.832 | 15.34 |
| 2015 | 0.975 | 0.987 | 18.96 | 0.98 | 0.977 | 17.06 | 0.981 | 0.961 | 18.51 |
| 2016 | 0.976 | 0.948 | 20.14 | 0.989 | 0.990 | 15.31 | 0.992 | 0.994 | 16.02 |
| 2017 | 0.977 | 0.822 | 20.15 | 0.986 | 0.954 | 15.36 | 0.985 | 0.941 | 21.76 |
| 2018 | 0.987 | 0.878 | 21.38 | 0.984 | 0.992 | 20.52 | 0.981 | 0.973 | 21.2 |

To visually highlight the differences in the downscaling results of the different models across various years, an annual line graph of the root mean square error was plotted, as shown in Fig. 7. This graph reveals that the original data consistently produced the best RMSE, with the lowest RMSE recorded in 2015, indicating the highest accuracy for that year. The precision of the XGBoost model was slightly less than that of the original dataset, closely followed by the BP Neural Network, while the Random Forest model consistently showed the weakest performance. The Random Forest model exhibited significant annual fluctuations in RMSE, highlighting its instability compared to the XGBoost model, which is thus deemed more suitable for downscaling tasks. This observation aligns with previous research, as noted in Ref. [30]. Regarding the trends, the validation results of each year varied, with the downscaled outcomes being influenced by the original data. The magnitude of error fluctuated with the trends of the original data. It was observed that the highest RMSE for all three models occurred in 2018, indicating the poorest performance, while the lowest errors were in 2015, marking the best performance. This fluctuation in the accuracy of downscaled results not only depends on the precision of the regression models but also on the quality of the original satellite precipitation data, a point that has been highlighted in several studies [14,47,48]. The validation metrics, as detailed in Table 3, reveal that the original data exhibits the highest degree of $R^2$ and the lowest MAE. The results of the XGBoost and BP Neural Network models follow closely, with the Random Forest model demonstrating comparatively poorer performance. The accuracy of the downscaled results is notably influenced by the quality of the original data. In 2017, the $R^2$ of all three downscaled results was relatively lower, ranging between 0.46 and 0.48. The lowest MAE and KGE were observed in 2018, with MAE ranging between 87.13 mm and 92.83 mm, and KGE ranging between 0.45 and 0.51. This indicates that the quality of the original data for the years 2017 and 2018 was inferior compared to that of other years.

Fig. 8 illustrates the spatial distribution patterns from 2015 to 2018. Subplots (A-D) represent Random Forest, XGBoost, BP Neural Network, and the original data, respectively, showing that each model effectively enhanced resolution and aligned the spatial patterns with those of the original data. Notably, the southeastern regions exhibit higher precipitation, while the western regions receive less, in keeping with the original data patterns. The high-resolution results obtained through training and learning from the original data have provided a wealth of spatial information and successfully captured the essential spatial characteristics, showcasing the strengths of machine learning algorithms in such applications. Focusing on the results from the three models, the Random Forest model exhibits a slight underestimation in downscaled precipitation, particularly evident in the southern parts of the Haihe River Basin in 2018. In contrast, the XGBoost model's precipitation results demonstrate a smoother transition in the precipitation zones, and the downscaled precipitation distribution of XGBoost more closely mirrors the original IMERG data. The BP Neural Network model, on the other hand, offers a more detailed description, particularly noticeable in the precipitation distribution of the western region in 2015 and the eastern coastal areas in 2017. Compared to the XGBoost and Random Forest models, it provides more detailed insights into precipitation patterns. However, when compared against the original data, the XGBoost model's predictions more closely resemble the original dataset. In sum, when considering the temporal and spatial distribution of precipitation, the XGBoost model emerges as more suitable for precipitation downscaling in regions with uneven precipitation distribution compared to the other two models.

In a comprehensive analysis comparing the fit and downscaling results of three machine learning methods, it is evident that each method capably handles outliers and noise, demonstrating robustness and preventing overfitting. These techniques not only enhance resolution but also maintain the precision and spatial features of the original images, marking them as superior for downscaling tasks. Among these, the XGBoost model stands out for its exceptional performance. The abundance of tunable hyperparameters within XGBoost allows for refined optimization, leading to downscaling models whose simulations and predictions are remarkably congruent with the original data, surpassing the results achieved by the Random Forest and BP Neural Network models. While the Random Forest model is effective, it shows limitations in capturing spatial distribution patterns, particularly in smoothness across certain regions and its tendency to underestimate, which compromises its ability to fully capture spatial heterogeneity [42,49]. Conversely, although the BP Neural Network excels at detailing spatial features, it slightly lags behind XGBoost in accurately simulating the spatial distribution of the original data. Therefore, in evaluating their suitability for annual scale downscaling tasks, the models rank as follows: XGBoost

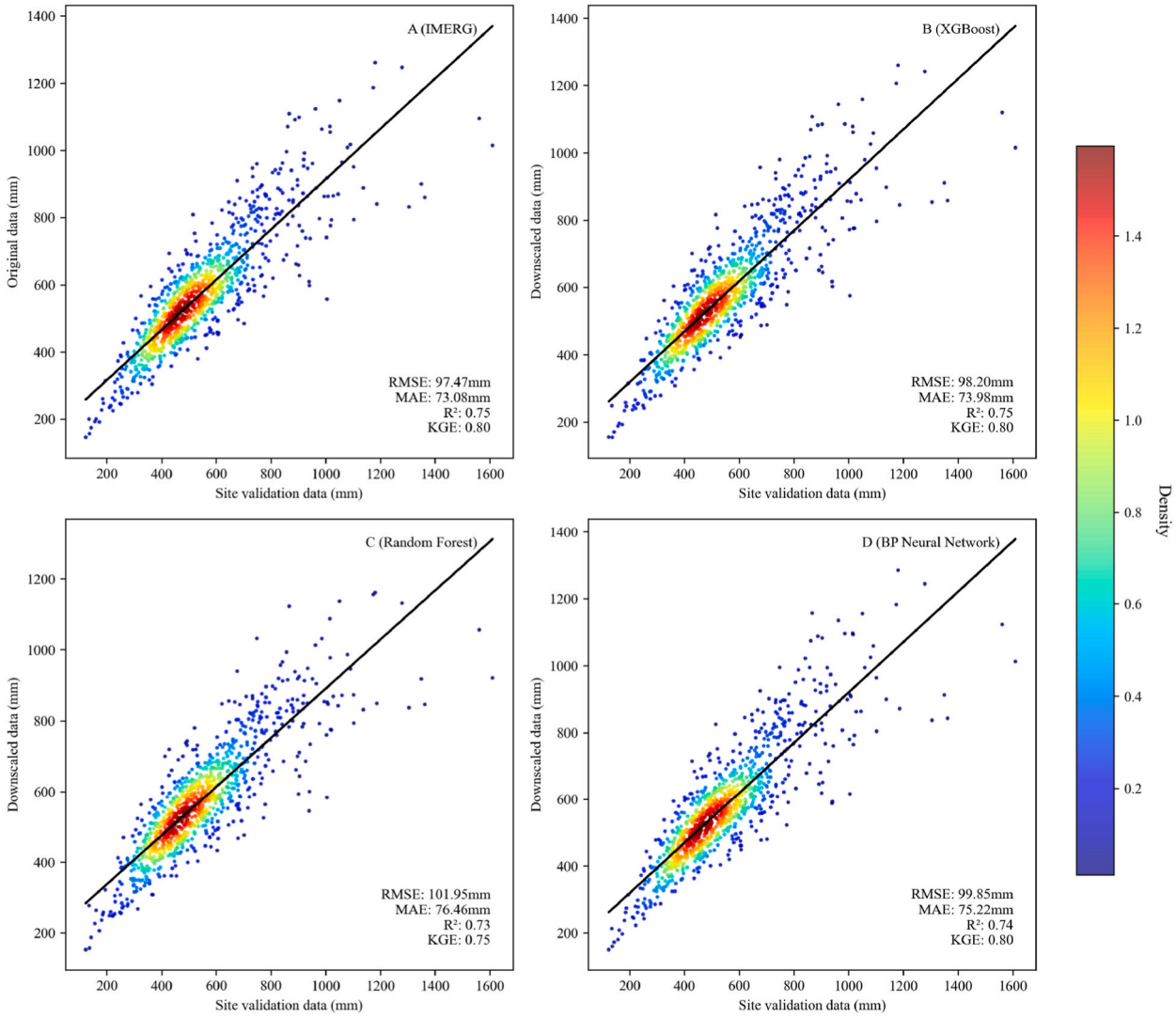


**Fig. 4.** Validation between rain gauges against (A) the original IMERG data, and downscaled results by (B) XGBoost, (C) Random Forest, and (D) BP Neural Network for the year 2003–2018.

> BP Neural Network > Random Forest. Due to its superior generalization abilities and more adept capturing of precipitation, XGBoost is deemed the most suitable model for downscaling efforts.

### *3.2. Residual correction and rain gauge correction results*

We compiled the results of residual correction into a density scatter plot, displayed in Fig. 9. The figure consists of four subplots (A-D), representing the validation results for the original data, XGBoost, Random Forest, and BP Neural Network, respectively. The overall results indicate a significant improvement in accuracy following residual correction.The downscaled accuracies of the three models are summarized with residual corrected accuracies as shown in Table 4. For the XGBoost model, while the fit remained constant, there was a decrease in the MAE from 73.98 mm to 73.27 mm and a reduction in the RMSE from 98.20 mm to 97.55 mm. The Bp Neural Network model maintained its fit at 0.74, with the MAE decreasing from 75.22 mm to 72.37 mm, and the RMSE from 99.85 mm to 97.31 mm. Most notably, the Random Forest model exhibited the greatest improvement in accuracy. Its fit improved from 0.73 to 0.74, the MAE was reduced from 76.46 mm to 72.19 mm, and the RMSE decreased from 101.95 mm to 98.04 mm. Additionally, the KGE indices for all models also improved, with both the BP Neural Network and XGBoost models increasing to 0.8, while the Random Forest's KGE rose to 0.76. These collective results demonstrate noticeable enhancements in precision across all models following the application of residual correction.

Fig. 10 illustrates the spatial distribution of precipitation for the year 2018, comparing three distinct data sets: the original data, the original downscaled results, and the results post-residual correction using three different models-Random Forest(A-B), BP Neural

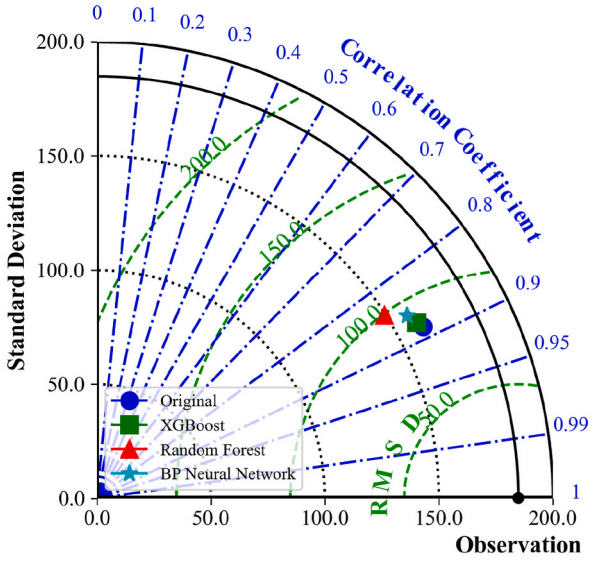


**Fig. 5.** Taylor diagram of rain gauge data validation for original and downscaled outputs.

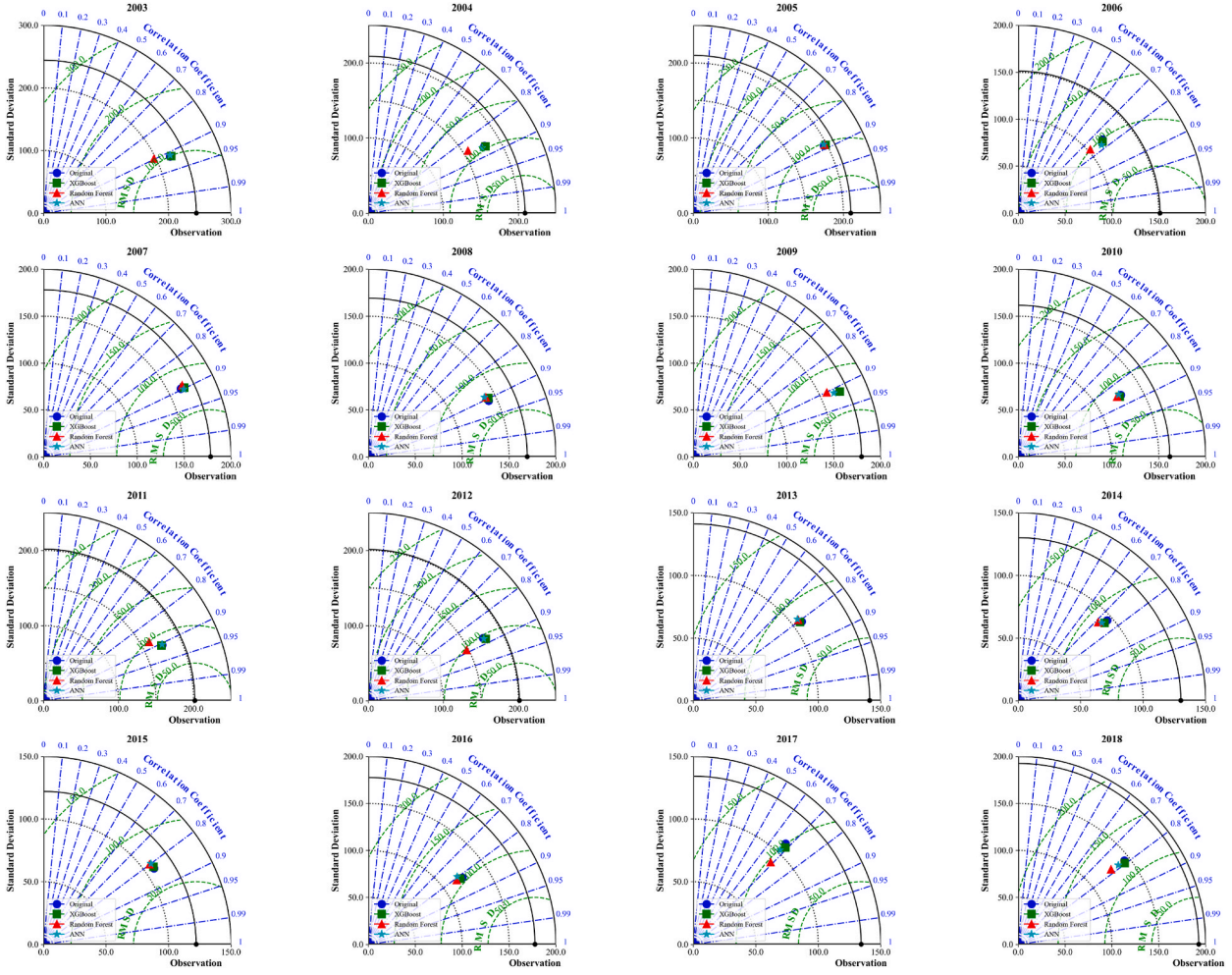


**Fig. 6.** Annual Taylor diagram of rain gauge data validation for original and downscaled outputs.

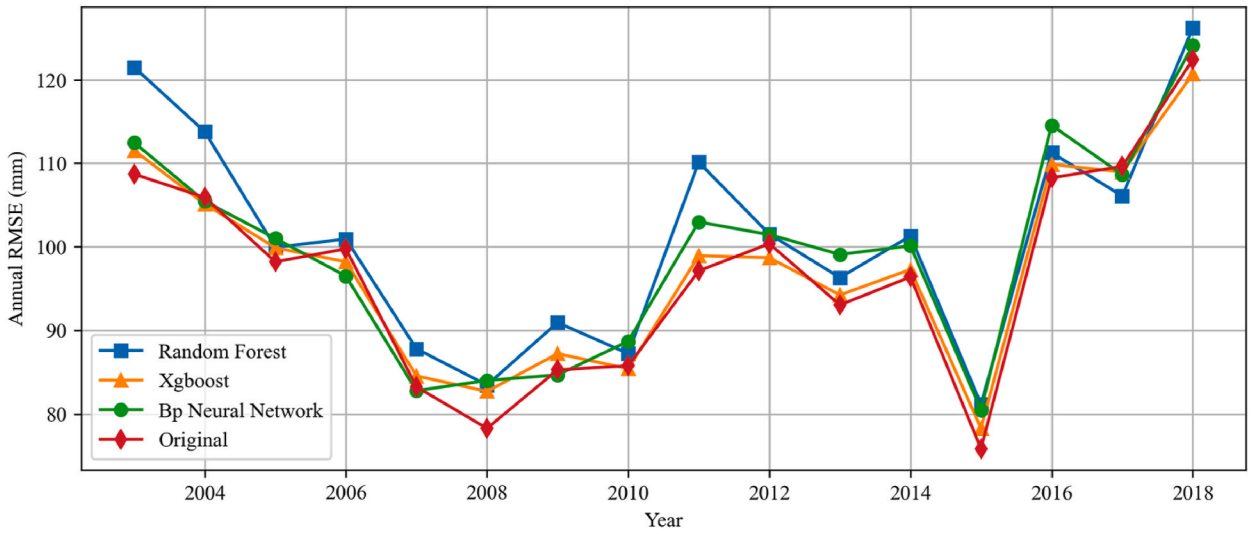


**Fig. 7.** Yearly error comparison (Original data and Downscaled results of three models).

Network (C-D), and XGBoost (E-F). Following residual correction, the RF model displayed a spatial distribution of precipitation in the central region similar to the original data but continued to underestimate precipitation in the southern region. This suggests that the RF model, relative to the other two, has a limited capacity for capturing spatial heterogeneity. After residual correction, the BP Neural Network model showed a spatial distribution closely mirroring the original, surpassing the original downscaled results in simulating rainfall in both central and southern areas. The XGBoost model's spatial distribution was already akin to the original; therefore, the post-correction spatial distribution showed no significant changes. While improving in accuracy, the XGBoost model maintained the original image's spatial distribution.

After applying residual correction, we integrated the corrected data using two methodologies, Geographical Difference Analysis (GDA) and Geographical Ratio Analysis (GRA), each employing both Kriging and Inverse Distance Weighting (IDW) interpolation methods. This combination was used to calibrate the results of the XGBoost model post-residual correction. Calibration was conducted at 31 selected sites, with the results illustrated in Fig. 11.

In terms of validation accuracy, integrating the GDA method with Kriging interpolation produced more favorable correction outcomes. Compared to the results following residual correction with the XGBoost model (Fig. 11B), the GDA-Kriging approach (Fig. 11E), despite a slight reduction in model fit, achieved significant improvements in precision, reducing the RMSE from 86.65 mm to 76.89 mm and the MAE from 68.52 mm to 58.4 mm. Additionally, the KGE also improved from 0.88 to 0.91. While other methods also showed improvements, they did not reach the level of precision provided by the GDA-Kriging approach. Therefore, the application of GDA-Kriging is particularly well-suited for calibration within this downscaling framework.

Fig. 12 displays the spatial distribution maps of the rainfall calibration results alongside the residual correction and original data: (A) GDA-IDW, (B) GDA-Kriging, (C) GRA-Kriging, and (D) GRA-IDW. Spatially, the results of the GDA and GRA exhibit similar distributions, with variations emerging primarily from the choice of interpolation method. IDW interpolation shows greater spatial variability in precipitation patterns, providing more detailed spatial information. Compared to the residual correction outcomes, the IDW-calibrated results show reduced precipitation in the west and south. Conversely, the Kriging interpolation maintains the spatial distribution of the original data, closely mirroring the layout of the initial imagery.

In this section, we evaluate the application and effectiveness of residual correction alongside rain gauge correction. The utility of residual correction in downscaling models has sparked considerable debate among researchers. Some argue that such corrections are not suitable for integration into downscaling processes [30,48,50], while others believe they significantly improve downscaling results [11,51]. This study employs three computational models — XGBoost, RF, and BP Neural Network — in conjunction with cubic spline interpolation to integrate residual correction into the downscaling data analysis. This approach yields a noticeable improvement when compared to the original downscaling results. Spatially, the correction alters the distribution in certain areas, aligning more closely with the original data. Overall, the application of cubic spline interpolation in residual correction markedly refines the results of downscaling, validating its inclusion in the downscaling framework. Among the models evaluated, XGBoost demonstrates superior accuracy in residual correction, outperforming the others. Consequently, the focus shifts to XGBoost, the most efficient model, for further refinement in rain gauge correction.

For rain gauge correction, we analyze four calibration methodologies applied to the XGBoost model post-residual correction: GDA-IDW, GDA-Kriging, GRA-IDW, and GRA-Kriging. In assessing the results for accuracy and spatial distribution, it is evident that although the spatial distribution results of the GDA and GRA methods are similar, the GDA approaches demonstrate greater precision than the GRA methods. Regarding interpolation techniques, IDW highlights greater spatial variability in precipitation patterns, enhancing spatial details consistent with prior studies [30,49,52–54]. Conversely, Kriging interpolation, while not enhancing the spatial heterogeneity of precipitation, effectively preserves the spatio-temporal distribution of the original image and achieves greater

**Table 3**
Validation of annual IMERG data and downscaled results by the machine learning algorithms against the gauge observations at annual scale (2003–2018).

| Year | Random Forest | | | | XGBoost | | | | BP Neural Network | | | | Original | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $R^2$ | KGE | MAE (mm) | RMSE (mm) | $R^2$ | KGE | MAE (mm) | RMSE ( mm ) | $R^2$ | KGE | MAE (mm) | RMSE ( mm ) | $R^2$ | KGE | MAE (mm) | RMSE ( mm ) |
| 2003 | 0.8 | 0.68 | 88.07 | 121.51 | 0.83 | 0.66 | 83.09 | 111.55 | 0.83 | 0.56 | 84.27 | 112.49 | 0.83 | 0.68 | 79.92 | 108.7 |
| 2004 | 0.72 | 0.60 | 67.72 | 113.79 | 0.76 | 0.59 | 65.48 | 105.13 | 0.75 | 0.49 | 64.63 | 105.45 | 0.75 | 0.60 | 66.51 | 105.92 |
| 2005 | 0.79 | 0.70 | 73.89 | 99.93 | 0.79 | 0.68 | 74.69 | 99.88 | 0.78 | 0.70 | 73.52 | 101.01 | 0.79 | 0.70 | 71.04 | 98.2 |
| 2006 | 0.56 | 0.67 | 74.15 | 100.93 | 0.58 | 0.64 | 70.13 | 98.2 | 0.6 | 0.62 | 70.36 | 96.49 | 0.57 | 0.68 | 69.54 | 99.75 |
| 2007 | 0.79 | 0.83 | 67.8 | 87.8 | 0.81 | 0.80 | 64.58 | 84.56 | 0.81 | 0.81 | 63.85 | 82.8 | 0.8 | 0.83 | 63.76 | 83.25 |
| 2008 | 0.8 | 0.76 | 66.09 | 83.47 | 0.81 | 0.76 | 66.05 | 82.71 | 0.79 | 0.75 | 65.19 | 84.02 | 0.82 | 0.77 | 62.4 | 78.29 |
| 2009 | 0.81 | 0.81 | 73.92 | 90.94 | 0.84 | 0.80 | 72.79 | 87.22 | 0.83 | 0.71 | 69.1 | 84.68 | 0.83 | 0.81 | 70.24 | 85.29 |
| 2010 | 0.73 | 0.63 | 62.81 | 87.23 | 0.74 | 0.63 | 60.33 | 85.42 | 0.72 | 0.60 | 64.65 | 88.7 | 0.73 | 0.63 | 61.88 | 85.77 |
| 2011 | 0.76 | 0.68 | 89.68 | 110.13 | 0.82 | 0.67 | 81.4 | 98.93 | 0.82 | 0.61 | 84.57 | 102.97 | 0.82 | 0.67 | 79.06 | 97.12 |
| 2012 | 0.79 | 0.85 | 82.26 | 101.52 | 0.78 | 0.82 | 77.95 | 98.69 | 0.78 | 0.88 | 81.42 | 101.46 | 0.77 | 0.85 | 80.4 | 100.37 |
| 2013 | 0.63 | 0.65 | 77.42 | 96.31 | 0.65 | 0.67 | 76.21 | 94.24 | 0.62 | 0.66 | 81.02 | 99.08 | 0.65 | 0.65 | 75.97 | 93.04 |
| 2014 | 0.51 | 0.47 | 78.27 | 101.28 | 0.55 | 0.51 | 74.96 | 97.28 | 0.54 | 0.42 | 76.74 | 100.11 | 0.55 | 0.43 | 73.71 | 96.39 |
| 2015 | 0.64 | 0.80 | 67.14 | 81.16 | 0.67 | 0.79 | 66.01 | 78.28 | 0.64 | 0.78 | 66.58 | 80.48 | 0.68 | 0.81 | 63.36 | 75.84 |
| 2016 | 0.65 | 0.52 | 81.43 | 111.27 | 0.66 | 0.56 | 79.23 | 109.87 | 0.63 | 0.53 | 83.65 | 114.55 | 0.67 | 0.59 | 78.05 | 108.26 |
| 2017 | 0.47 | 0.66 | 80.74 | 106.09 | 0.48 | 0.63 | 84.26 | 108.96 | 0.46 | 0.55 | 85.89 | 108.65 | 0.46 | 0.66 | 86 | 109.64 |
| 2018 | 0.61 | 0.51 | 92.83 | 126.22 | 0.63 | 0.50 | 87.13 | 120.71 | 0.62 | 0.45 | 88.77 | 124.14 | 0.62 | 0.52 | 88.24 | 122.44 |

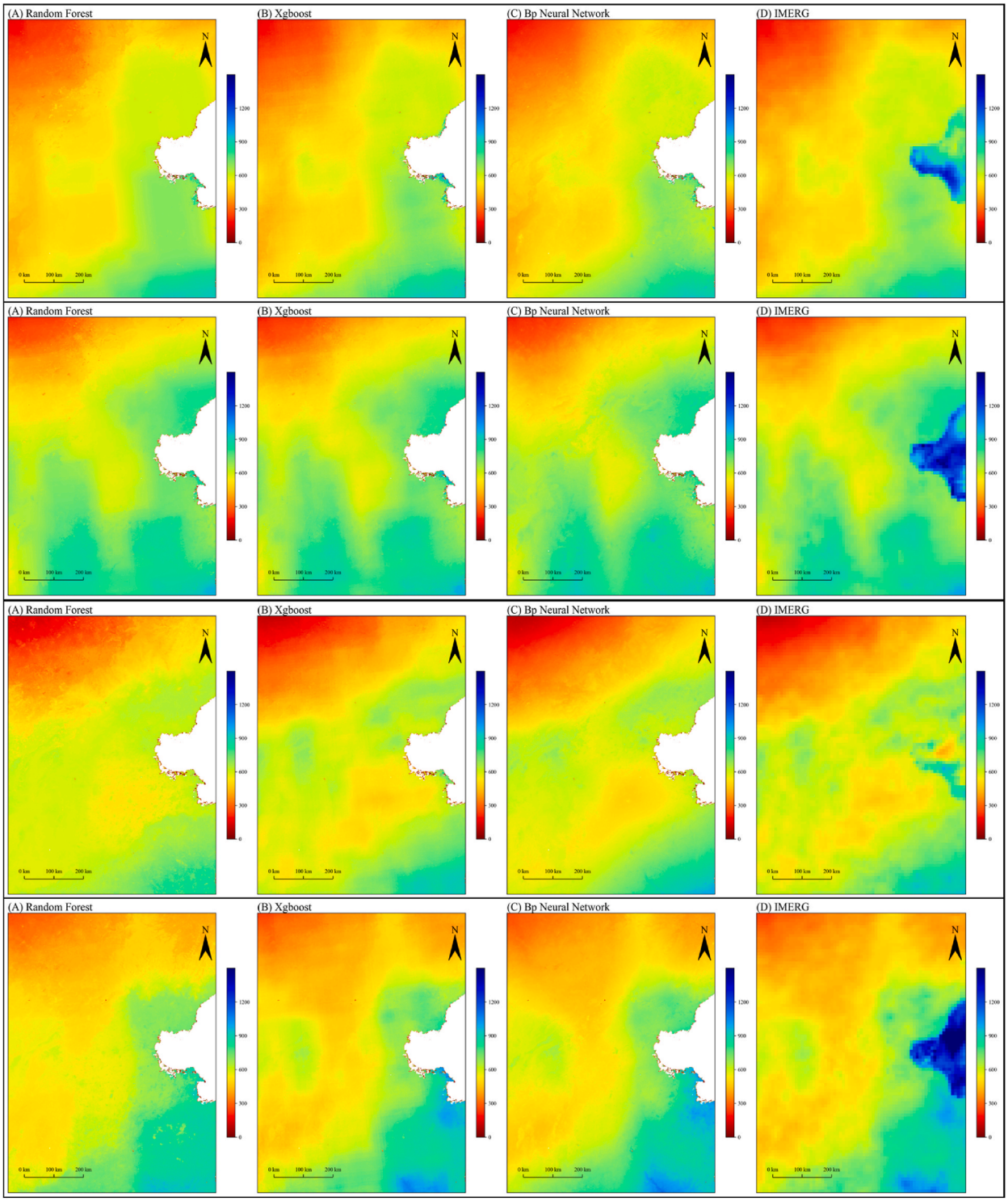


**Fig. 8.** Comparison of precipitation spatial distribution between original images (D) and downscaled results by (A) Random Forest, (B) XGBoost and (C) BP Neural Network from 2015 to 2018.

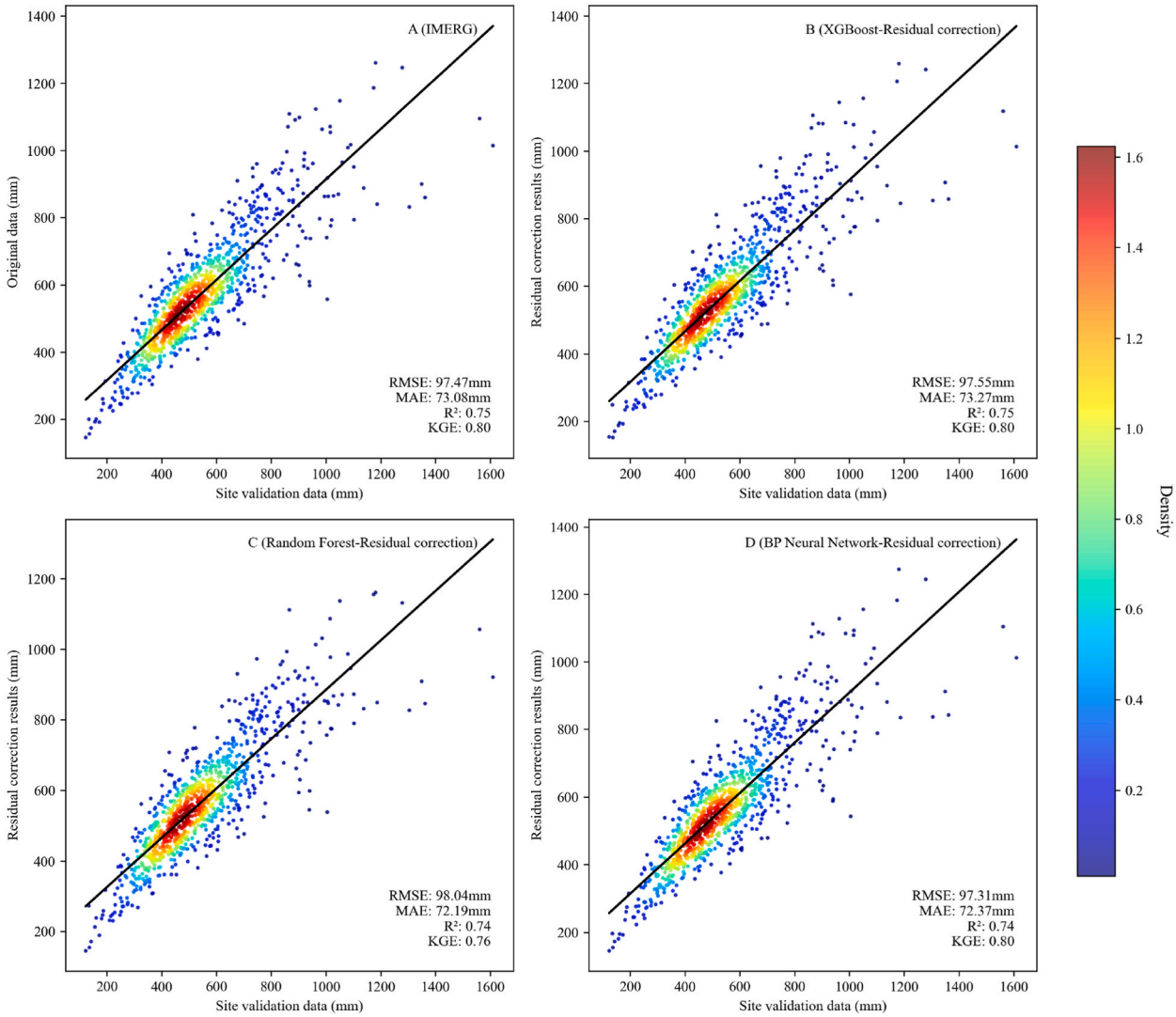


**Fig. 9.** Scatter plot of residual correction results.

**Table 4**
Comparison of the accuracy of residual correction results with the original downscaling results.

| Method | RMSE (mm) | KGE | MAE (mm) | $R^2$ |
|---|---|---|---|---|
| Random Forest | 101.95 | 0.73 | 76.46 | 0.73 |
| Random Forest(Residual correction) | 98.04 | 0.76 | 72.19 | 0.74 |
| BP Neural Network | 99.85 | 0.74 | 75.22 | 0.74 |
| BP(Residual correction) | 97.31 | 0.80 | 72.37 | 0.74 |
| XGBoost | 98.2 | 0.75 | 73.98 | 0.75 |
| XGBoost(Residual correction) | 97.55 | 0.80 | 73.27 | 0.75 |
| IMERG | 97.47 | 0.80 | 73.08 | 0.75 |

accuracy than GDA-IDW. The calibration outcomes using GDA-Kriging show a more effective integration with rain gauge data, making it a preferable approach for enhancing precision in annual-scale rain gauge calibration compared to GDA-IDW.

### *3.3. Monthly and daily scale analysis*

As illustrated in Fig. 13, the downscaled monthly data (Fig. 13B) parallels the accuracy of the original dataset (Fig. 13A), enhancing resolution while maintaining precision. From a spatial analysis perspective, as demonstrated in Fig. 14, the downscaled monthly

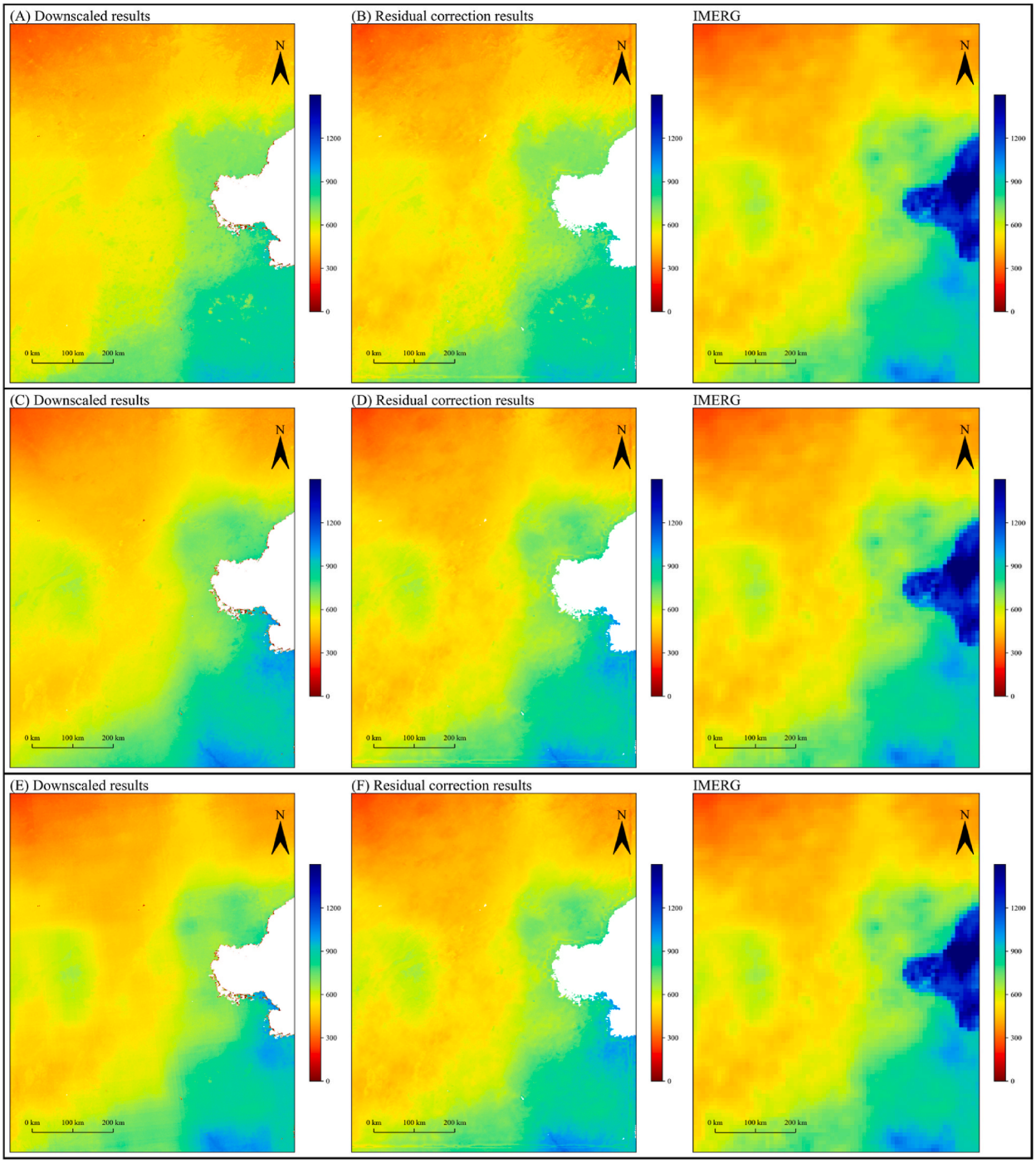


**Fig. 10.** Comparison of precipitation spatial distribution of original images , downscaled results and residual correcrtion results by (A–B) Random Forest, (C–D) BP Neural Network and (E–F) XGBoost in 2018.

results maintain the inherent characteristics of the original data. This consistency is maintained throughout the year, across both rainy (Fig. 14C) and dry seasons (Fig. 14 A B D). The spatial patterns seamlessly align with the original configuration, without anomalies, accurately reflecting the precipitation trends and providing detailed precipitation information across various regions. These results validate the effectiveness and suitability of the spatially weighted moving window technique for generating monthly data.

Fig. 15 presents the validation results for daily data. When approaching zero values in rain gauge data, it is apparent that both original data (Fig. 15A) and downscaled data (Fig. 15B) exhibit a noticeable overestimation compared to the rain gauge readings. This finding corroborates previous research that consistently identifies an overestimation bias in GPM data, which intensifies with

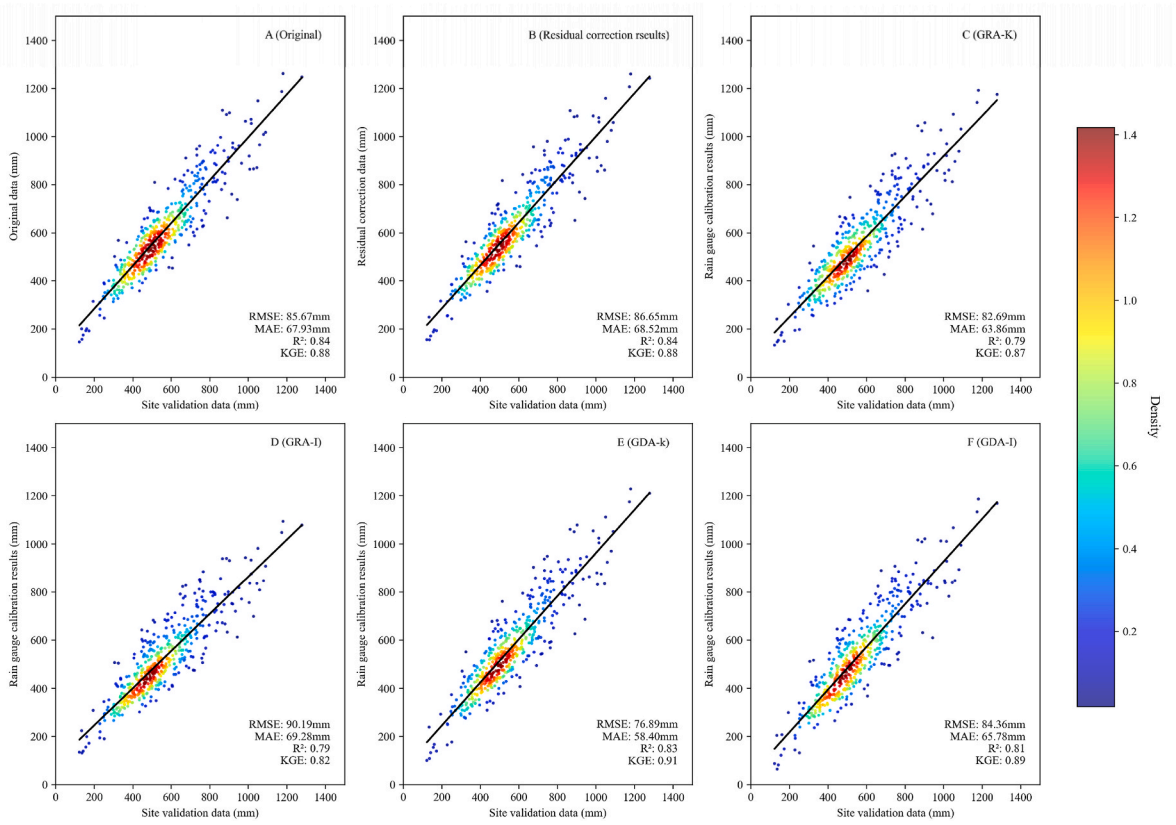


**Fig. 11.** Scatter plot of rain gauge calibration results.

increasing precipitation levels [43,55]. Therefore, the presence of this bias near zero values on the chart is expected, considering the downscaled results derive from the decomposition of original data and are influenced accordingly. Regarding precipitation detection, the Probability of Detection (POD) for the daily disaggregated results is 0.82, surpassing the original data's 0.74, which suggests that the daily scale products possess enhanced capability in capturing rainfall events. Nonetheless, the False Alarm Ratio (FAR) is higher in the downscaled data, indicating an overestimation that erroneously classifies non-rainfall events as rainfall. Moreover, examining the Composite Skill Index (CSI), both datasets perform similarly, indicating that despite discrepancies in POD and FAR, the overall ability to accurately detect rainfall events remains comparable. Consequently, considering all evaluation metrics, the downscaled daily data maintains its precision even with improved resolution.

The Hai River Basin is notably prone to frequent flooding disasters, predominantly occurring between June and September. These events are typically linked with several intense rainfall episodes, marking it as one of the most flood-prone regions in the country. The unique topography of the region, characterized by a directional tilt from the south, west, and north towards the east, historically channels all river floods to converge at Tianjin before entering the sea. This leads to a significant disparity in the river's discharge capacity, which is greater upstream and reduced downstream, especially at the estuary where the discharge volume is minimal. Consequently, even moderate floods can lead to widespread disasters. Therefore, accurately predicting and analyzing extreme weather conditions is critical. Fig. 16 displays the intense rainfall event on July 21, 2012 (Fig. 16A), and the exceptionally heavy precipitation on July 19, 2016 (Fig. 16B), clearly demonstrating that the daily decomposition results can precisely identify areas with excessive precipitation. This capability facilitates rapid identification of precipitation locations through imagery, enabling efficient analysis.

### *3.4. Analysis of environmental variables*

SHAP (Shapley additive explanations) provides a method for interpreting predictions made by machine learning models. It establishes a framework to quantify the contribution of each feature to the model's prediction, facilitating a deeper understanding of the model's decision-making process. Fig. 17 displays the SHAP values for select years. Notably, in 2004 (Fig. 17A), the most significant influences on the model were nocturnal land surface temperature and latitude, whereas in 2006 (Figs. 17B), 2010 (Fig. 17C), and 2012 (Fig. 17D), latitude and longitude emerged as the predominant factors. The magnitude of impact of each factor varies annually, corroborating findings from other literature [56]. Fig. 17E aggregates the mean SHAP values from 2003 to 2018. It reveals that latitude and longitude consistently exert the greatest average impact, while nocturnal LST_night, DEM, and PWV also contribute significantly. NDVI and Albedo have a relatively lower influence compared to other features, and LST_day shows the least impact.

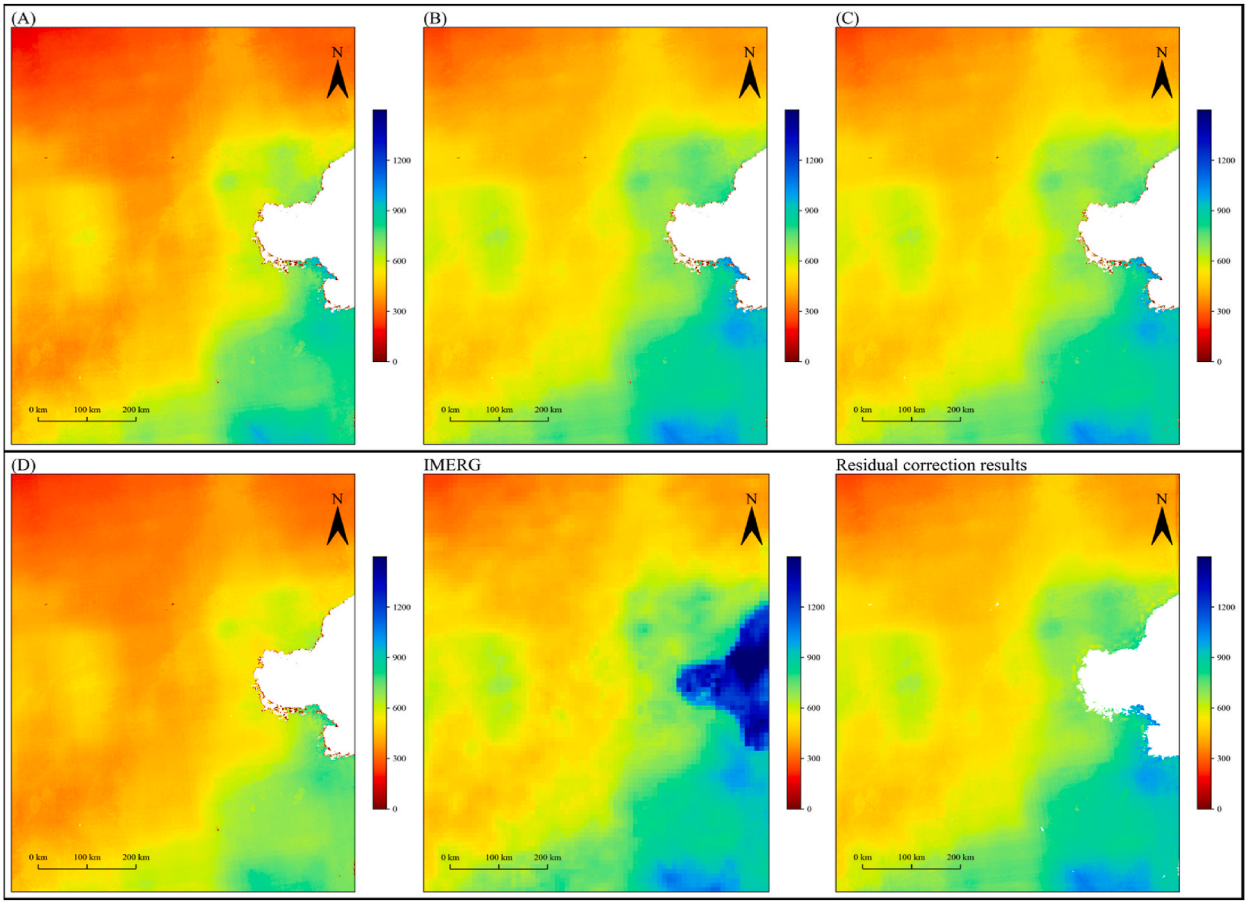


**Fig. 12.** Precipitation spatial distributions comparing rain gauge-calibrated methods (A) GDA-IDW, (B) GDA-Kriging, (C) GRA-Kriging, (D) GRA-IDW, with original and model-corrected data.

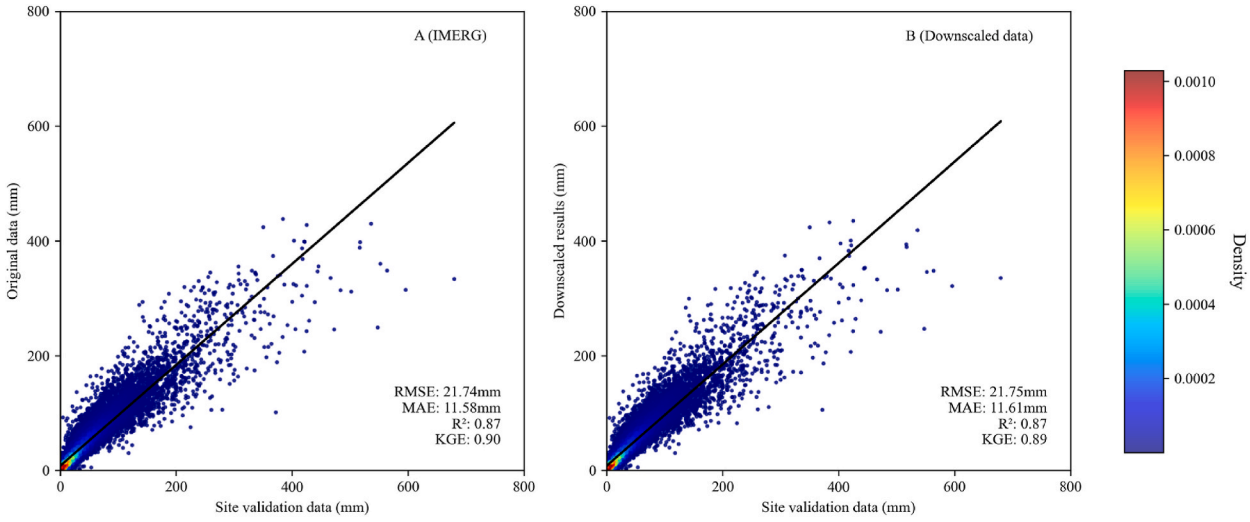


**Fig. 13.** Comparative scatter plots of monthly values in downscaled versus original data.

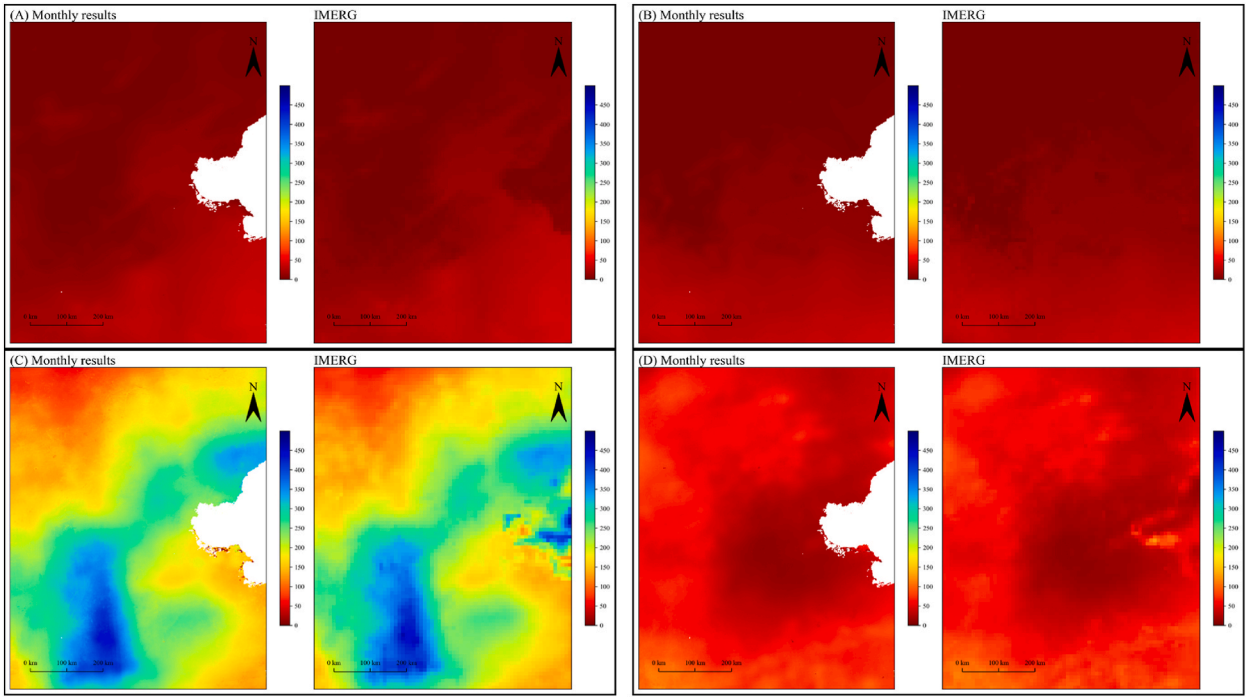


**Fig. 14.** Spatial Comparison of Precipitation Distributions in Original and Monthly results.(A) March 2009, (B) February 2014, (C) June 2016, (D) September 2018.

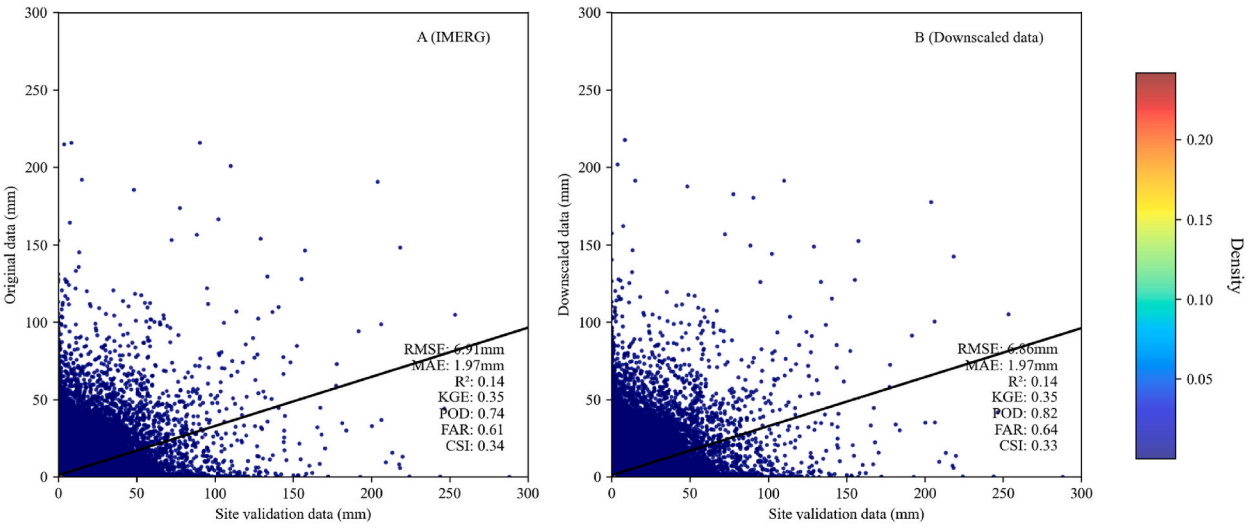


**Fig. 15.** Comparative scatter plots of daily values in downscaled versus original data.

In the majority of downscaling literature, the selection of scale factors typically focuses on variables such as elevation, latitude and longitude, and NDVI for model construction [43,49,57,58], In contrast, this study introduces eight independent variables to establish the downscaling model, exploring the impact of a broader range of scale factors on model performance. SHAP analysis identifies latitude and longitude as significant factors, primarily due to the influence of eastern coastal areas on rainfall in the Hai River Basin, which significantly affects precipitation and its spatial distribution. NDVI, a common predictive factor, influences the sensible and latent heat fluxes into the atmosphere due to vegetation characteristics, directly affecting the humidity of the lower atmosphere and consequently, the development of moist convection locally and across scales of tens of thousands of kilometers [59]. However, in the Hai River Basin, areas of higher altitude exhibit sparse vegetation, leading to lower SHAP values for NDVI at desert verification points. Elevation, another geographical spatial predictor in downscaling methods, increases the relative humidity of air masses through

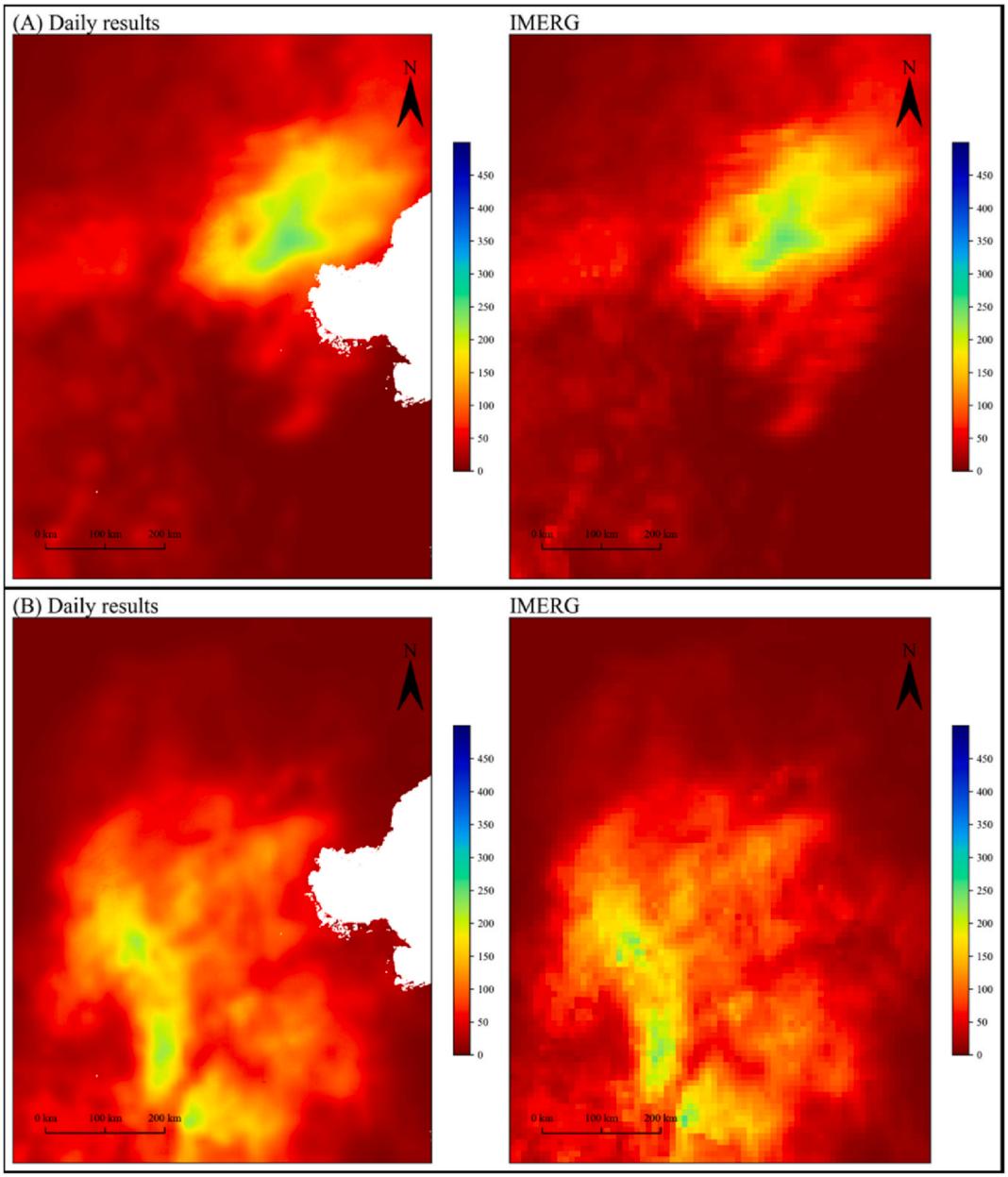


**Fig. 16.** The spatial distributions under extreme weather conditions at two different time points: (A) July 12, 2012,and (B) July 19, 2016.

expansion and cooling as altitude rises, thus inducing precipitation. The western mountainous areas of the Hai River Basin make the inclusion of the DEM a significant factor in the model. LST has also been incorporated as a suitable environmental variable [60], Studies suggest that if the ground is wet, more energy might be expended in evaporation rather than sensible heat, thus acting like an "air conditioner." In the Hai River Basin, LST-night has a higher impact than daytime LST-day, possibly due to cloud cover blocking the sun if the ground is moist from precipitation, thereby reducing initial energy input and further lowering temperatures [59,61]. PWV is a crucial parameter, providing information on the atmospheric water vapor content, and its stability makes it a reliable auxiliary factor due to its direct relation to precipitation. Albedo, was also included in this study. Precipitation can alter the surface albedo by affecting the moisture level of the soil surface. Since our study area is located at a higher elevation with sparse soil, this environmental variable has a relatively minor influence on the model.

## 4. Conclusion

In this study, we conducted a comprehensive evaluation of high spatiotemporal resolution precipitation downscaling methods in

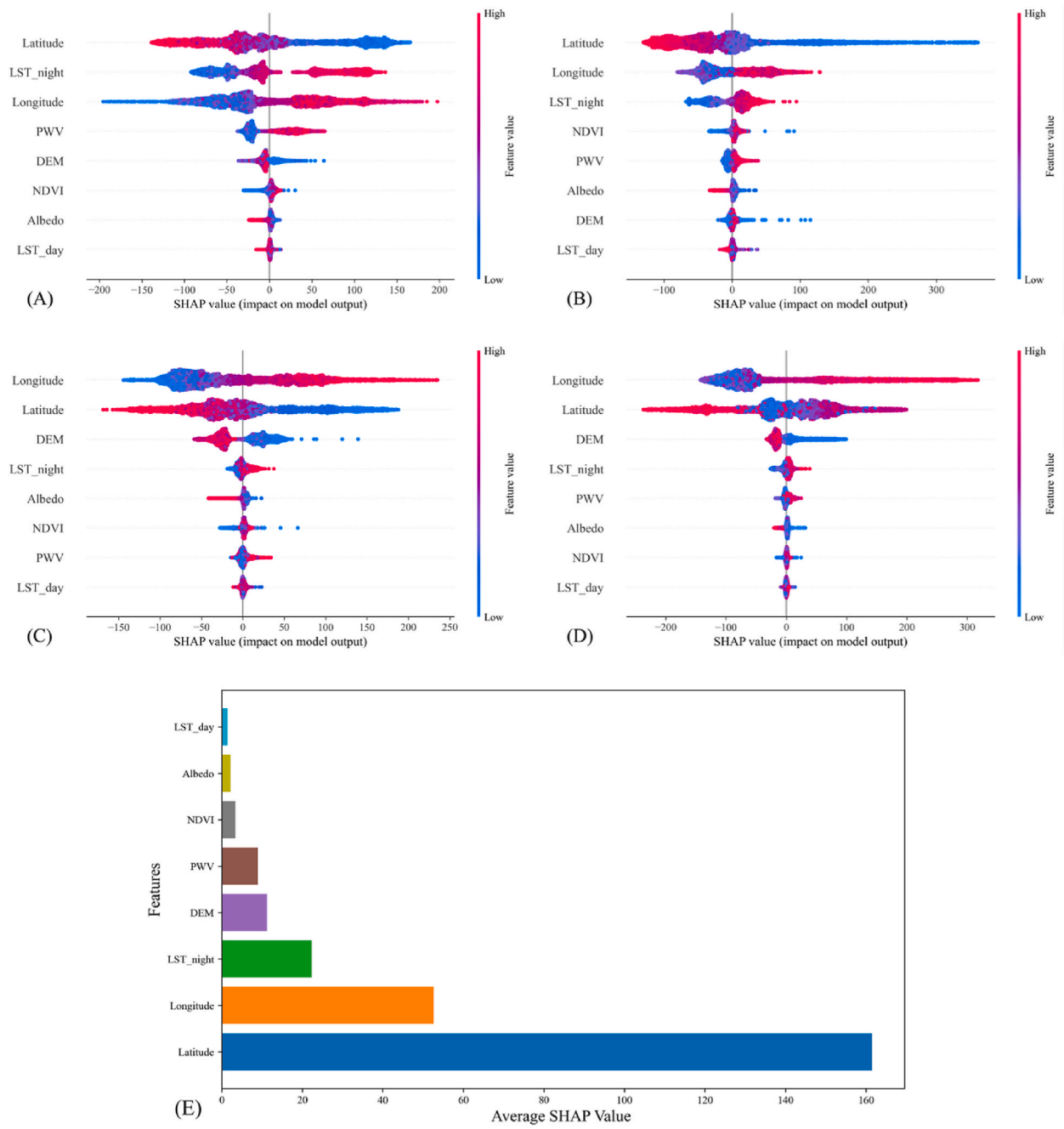


**Fig. 17.** Partial Year Model SHAP Analysis and Annual Model SHAP Value Analysis. SHAP (Shapley Additive explanations) analysis for partial years (A–D) and the average over 2003–2019 (E). (A) 2004, (B) 2006, (C) 2010, (D) 2012 show the SHAP values for individual years. (E) Depicts the average SHAP values for the years 2003–2019.

the Haihe River Basin using three machine learning algorithms: Random Forest, XGBoost, and BP neural networks. By processing and finely calibrating annual precipitation data, combined with spatial decomposition methods, we effectively achieved accurate decomposition to monthly and daily scales. Key findings of this study are as follows.

1. The XGBoost algorithm demonstrated superior precipitation capturing ability compared to Random Forest and BP Neural Networks, making it the preferred model for annual precipitation downscaling in the Haihe River Basin.
2. The application of cubic spline interpolation for residual correction significantly improved spatial heterogeneity and effectively enhanced overall accuracy.

3. The integration of GDA correction with Kriging spatial interpolation further improved the quality of the downscaling results.
4. Environmental variables such as latitude and longitude, LST-night, and DEM had significant impacts on the model's performance in reproducing atmospheric precipitation.

Based on the comprehensive results, XGBoost should be employed as the primary downscaling method. Results should be calibrated using residual correction and GDA-Kriging, and the annual data further decomposed into daily and monthly data using a spatially weighted moving window approach. The high temporal resolution downscaling framework evaluated in this study demonstrates high accuracy and temporal resolution. Moreover, constructed based on the relationship between precipitation and environmental factors, the model is applicable to regions with complex terrain and diverse climates. The high spatiotemporal resolution data provided by this framework offers robust support for hydrological and related fields.

## Funding statement

The research was supported by the National Key Research and Development Program of China (2022YFE0113900), The Science and Technology Fundamental Resources Investigation Program (Grant No. 2022FY100102), and the National Natural Science Foundation of China (Grant No. 42301420).

## Data availability statement

Data will be made available on request.

## CRediT authorship contribution statement

**Tao Sun:** Writing – original draft, Software, Conceptualization. **Nana Yan:** Project administration. **Weiwei Zhu:** Funding acquisition. **Qifeng Zhuang:** Validation, Conceptualization.

## 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.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e36368.

## References

[1] M. Li, Q. Shao, An improved statistical approach to merge satellite rainfall estimates and raingauge data, J. Hydrol. 385 (2010) 51–64, https://doi.org/10.1016/j.jhydrol.2010.01.023.
[2] C. Wu, J.M. Chen, The use of precipitation intensity in estimating gross primary production in four northern grasslands, J. Arid Environ. 82 (2012) 11–18, https://doi.org/10.1016/j.jaridenv.2012.02.014.
[3] K.K. Yilmaz, T.S. Hogue, K. Hsu, S. Sorooshian, H.V. Gupta, T. Wagener, Intercomparison of rain gauge, Radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting, J. Hydrometeorol. 6 (2005) 497–517, https://doi.org/10.1175/JHM431.1.
[4] Y. Tian, C.D. Peters-Lidard, J.B. Eylander, R.J. Joyce, G.J. Huffman, R.F. Adler, K. Hsu, F.J. Turk, M. Garcia, J. Zeng, Component analysis of errors in satellite-based precipitation estimates, J. Geophys. Res. 114 (2009) D24101, https://doi.org/10.1029/2009JD011949.
[5] Z. Ma, Y. Zhou, B. Hu, Z. Liang, Z. Shi, Downscaling annual precipitation with TMPA and land surface characteristics in China, Intl Journal of Climatology 37 (2017) 5107–5119, https://doi.org/10.1002/joc.5148.
[6] T.C. Peterson, M.P. Hoerling, P.A. Stott, S.C. Herring, EXPLAINING EXTREME EVENTS OF 2012 FROM A CLIMATE PERSPECTIVE, 2013.
[7] G.J. Huffman, R.F. Adler, D.T. Bolvin, E.J. Nelkin, The TRMM multi-satellite precipitation analysis (TMPA), in: M. Gebremichael, F. Hossain (Eds.), Satellite Rainfall Applications for Surface Hydrology, Springer Netherlands, Dordrecht, 2010, pp. 3–22, https://doi.org/10.1007/978-90-481-2915-7_1.
[8] V. Levizzani, C. Kidd, D.B. Kirschbaum, C.D. Kummerow, K. Nakamura, F.J. Turk (Eds.), Satellite Precipitation Measurement, vol. 1, Springer, Cham, 2020, https://doi.org/10.1007/978-3-030-35798-6.
[9] H. Ashouri, K.-L. Hsu, S. Sorooshian, D.K. Braithwaite, K.R. Knapp, L.D. Cecil, B.R. Nelson, O.P. Prat, PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies, Bull. Am. Meteorol. Soc. 96 (2015) 69–83, https://doi.org/10.1175/BAMS-D-13-00068.1.
[10] D.A. Sachindra, B.J.C. Perera, Statistical downscaling of general circulation model outputs to precipitation accounting for non-stationarities in predictor-predictand relationships, PLoS One 11 (2016) e0168701, https://doi.org/10.1371/journal.pone.0168701.
[11] E. Sharifi, B. Saghafian, R. Steinacker, Downscaling satellite precipitation estimates with multiple linear regression, artificial neural networks, and spline interpolation techniques, JGR Atmospheres 124 (2019) 789–805, https://doi.org/10.1029/2018JD028795.
[12] Arman Abdollahipour, Hassan Ahmadi, Babak Aminnejad, A review of downscaling methods of satellite-based precipitation estimates, Earth Sci. Inf. 15 (2022) 1–20, https://doi.org/10.1007/s12145-021-00669-4.
[13] W.W. Immerzeel, M.M. Rutten, P. Droogers, Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula, Rem. Sens. Environ. 113 (2009) 362–370, https://doi.org/10.1016/j.rse.2008.10.004.
[14] S. Jia, W. Zhu, A. Lű, T. Yan, A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China, Rem. Sens. Environ. 115 (2011) 3069–3079, https://doi.org/10.1016/j.rse.2011.06.009.

[15] S. Xu, C. Wu, L. Wang, A. Gonsamo, Y. Shen, Z. Niu, A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics, Rem. Sens. Environ. 162 (2015) 119–140, https://doi.org/10.1016/j.rse.2015.02.024.
[16] X. Li, X. He, X. Li, Y. Du, G. Yang, H. Tian, J. Yu, Construction of high-resolution precipitation dataset and its implication to drought over the Tianshan Mountains, China. https://doi.org/10.3389/feart.2023.1128990, 2023.
[17] Y. Shi, L. Song, Z. Xia, Y. Lin, R. Myneni, S. Choi, L. Wang, X. Ni, C. Lao, F. Yang, Mapping annual precipitation across mainland China in the period 2001–2010 from TRMM3B43 product using spatial downscaling approach, Rem. Sens. 7 (2015) 5849–5878, https://doi.org/10.3390/rs70505849.
[18] Z. Ma, K. He, X. Tan, J. Xu, W. Fang, Y. He, Y. Hong, Comparisons of spatially downscaling TMPA and IMERG over the Tibetan plateau, Rem. Sens. 10 (2018) 1883, https://doi.org/10.3390/rs10121883.
[19] J. Zhang, H. Fan, D. He, J. Chen, Integrating precipitation zoning with random forest regression for the spatial downscaling of satellite-based precipitation: a case study of the Lancang–Mekong River basin, Intl Journal of Climatology 39 (2019) 3947–3961, https://doi.org/10.1002/joc.6050.
[20] Zhehui Shen, Bin Yong, Downscaling the GPM-based satellite precipitation retrievals using gradient boosting decision tree approach over Mainland China, J. Hydrol. 602 (2021) 126803, https://doi.org/10.1016/j.jhydrol.2021.126803.
[21] H. Wang, Z. Li, T. Zhang, Q. Chen, X. Guo, Q. Zeng, J. Xiang, Downscaling of GPM satellite precipitation products based on machine learning method in complex terrain and limited observation area, Adv. Space Res. 72 (2023) 2226–2244, https://doi.org/10.1016/j.asr.2023.05.057.
[22] M. Kofidou, S. Stathopoulos, A. Gemitzi, Review on spatial downscaling of satellite derived precipitation estimates, Environ. Earth Sci. 82 (2023) 424, https://doi.org/10.1007/s12665-023-11115-7.
[23] W.W. Immerzeel, R.A. Quiroz, S.M. De Jong, Understanding precipitation patterns and land use interaction in Tibet using harmonic analysis of SPOT VGT-S10 NDVI time series, Int. J. Rem. Sens. 26 (2005) 2281–2296, https://doi.org/10.1080/01431160512331326611.
[24] R. Quiroz, C. Yarlequé, A. Posadas, V. Mares, W.W. Immerzeel, Improving daily rainfall estimation from NDVI using a wavelet transform, Environ. Model. Software 26 (2011) 201–209, https://doi.org/10.1016/j.envsoft.2010.07.006.
[25] C. Xu, C. Liu, W. Zhang, Z. Li, B. An, Downscaling and merging of daily scale satellite precipitation data in the three river headwaters region fused with cloud attributes and rain gauge data. https://doi.org/10.3390/w15061233, 2023.
[26] Z. Zeng, H. Chen, Q. Shi, J. Li, Spatial downscaling of IMERG considering vegetation index based on adaptive lag phase, IEEE Trans. Geosci. Rem. Sens. 60 (2022) 1–15, https://doi.org/10.1109/TGRS.2021.3070417.
[27] O. Zandi, B. Zahraie, M. Nasseri, A. Behrangi, Stacking machine learning models versus a locally weighted linear model to generate high-resolution monthly precipitation over a topographically complex area, Atmos. Res. 272 (2022) 106159, https://doi.org/10.1016/j.atmosres.2022.106159.
[28] B. Kumar, R. Chattopadhyay, M. Singh, N. Chaudhari, K. Kodari, A. Barve, Deep learning–based downscaling of summer monsoon rainfall data over Indian region, Theor. Appl. Climatol. 143 (2021) 1145–1156, https://doi.org/10.1007/s00704-020-03489-6.
[29] B. Kumar, K. Atey, B.B. Singh, R. Chattopadhyay, N. Acharya, M. Singh, R.S. Nanjundiah, S.A. Rao, On the modern deep learning approaches for precipitation downscaling, Earth Sci Inform 16 (2023) 1459–1472, https://doi.org/10.1007/s12145-023-00970-4.
[30] H. Zhu, H. Liu, Q. Zhou, A. Cui, Towards an accurate and reliable downscaling scheme for high-spatial-resolution precipitation data, Rem. Sens. 15 (2023) 2640, https://doi.org/10.3390/rs15102640.
[31] A. Karbalaye Ghorbanpour, T. Hessels, S. Moghim, A. Afshar, Comparison and assessment of spatial downscaling methods for enhancing the accuracy of satellite-based precipitation over Lake Urmia Basin, J. Hydrol. 596 (2021) 126055, https://doi.org/10.1016/j.jhydrol.2021.126055.
[32] M. Nasseri, G. Schoups, M. Taheri, A spatiotemporal framework to calibrate high-resolution global monthly precipitation products: an application to the Urmia Lake Watershed in Iran, Intl Journal of Climatology 42 (2022) 2169–2194, https://doi.org/10.1002/joc.7358.
[33] O. Zandi, M. Nasseri, B. Zahraie, A locally weighted linear ridge regression framework for spatial interpolation of monthly precipitation over an orographically complex area, Intl Journal of Climatology 43 (2023) 2601–2622, https://doi.org/10.1002/joc.7992.
[34] R. Zheng, Z. Yong, C. Ling, G. Jiaguo, W. Li, Changing properties of daily precipitation concentration in the Hai River Basin, China, Sci. Program. 2023 (2023) 1–14, https://doi.org/10.1155/2023/1149108.
[35] Z. Jia, S. Liu, Z. Xu, Y. Chen, M. Zhu, Validation of remotely sensed evapotranspiration over the Hai River Basin, China, J. Geophys. Res. 117 (2012) 2011JD017037, https://doi.org/10.1029/2011JD017037.
[36] T. Chen, T. He, xgboost: eXtreme Gradient Boosting, (n.d.).https://cran.ms.unimelb.edu.au/web/packages/xgboost/vignettes/xgboost.pdf.
[37] H. Tao, M. Habib, I. Aljarah, H. Faris, H.A. Afan, Z.M. Yaseen, An intelligent evolutionary extreme gradient boosting algorithm development for modeling scour depths under submerged weir, Inf. Sci. 570 (2021) 172–184, https://doi.org/10.1016/j.ins.2021.04.063.
[38] R. Díaz-Uriarte, S. Alvarez De Andrés, Gene selection and classification of microarray data using random forest, BMC Bioinf. 7 (2006) 3, https://doi.org/10.1186/1471-2105-7-3.
[39] S. Hu, H. Liu, W. Zhao, T. Shi, Z. Hu, Q. Li, G. Wu, Comparison of machine learning techniques in inferring phytoplankton size classes, Rem. Sens. 10 (2018) 191, https://doi.org/10.3390/rs10030191.
[40] G. Xiao, X. Zhang, Q. Niu, X. Li, X. Li, L. Zhong, J. Huang, Winter wheat yield estimation at the field scale using sentinel-2 data and deep learning, Comput. Electron. Agric. 216 (2024) 108555, https://doi.org/10.1016/j.compag.2023.108555.
[41] F. Xuan, Y. Dong, J. Li, X. Li, W. Su, X. Huang, J. Huang, Z. Xie, Z. Li, H. Liu, W. Tao, Y. Wen, Y. Zhang, Mapping crop type in Northeast China during 2013–2021 using automatic sampling and tile-based image classification, Int. J. Appl. Earth Obs. Geoinf. 117 (2023) 103178, https://doi.org/10.1016/j.jag.2022.103178.
[42] X. He, N.W. Chaney, M. Schleiss, J. Sheffield, Spatial downscaling of precipitation using adaptable random forests, Water Resour. Res. 52 (2016) 8217–8237, https://doi.org/10.1002/2016WR019034.
[43] H. Wang, Z. Li, Z. Tao, Q. Chen, X.-P. Guo, Q. Zeng, J. Xiang, Downscaling of GPM satellite precipitation products based on machine learning method in complex terrain and limited observation area, Adv. Space Res. 72 (2023) 2226–2244, https://doi.org/10.1016/j.asr.2023.05.057.
[44] Ziqiang Ma, Kang He, Xiao Tan, Yong Liu, Hanyu Lu, Zhou Shi, A new approach for obtaining precipitation estimates with a finer spatial resolution on a daily scale based on TMPA V7 data over the Tibetan Plateau, Int. J. Rem. Sens. 40 (2019) 8465–8483, https://doi.org/10.1080/01431161.2019.1612118.
[45] M. Cheema, W. Bastiaanssen, Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin, Int. J. Rem. Sens. 33 (2012) 2603–2627, https://doi.org/10.1080/01431161.2011.617397.
[46] X. Yan, H. Chen, B. Tian, S. Sheng, J. Wang, J.-S. Kim, A downscaling–merging scheme for improving daily spatial precipitation estimates based on random forest and cokriging, Rem. Sens. 13 (2021) 2040, https://doi.org/10.3390/rs13112040.
[47] W.W. Immerzeel, M.M. Rutten, P. Droogers, Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula, Rem. Sens. Environ. 113 (2009) 362–370, https://doi.org/10.1016/j.rse.2008.10.004.
[48] S. Xu, C. Wu, L. Wang, A. Gonsamo, Y. Shen, Z. Niu, A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics, Rem. Sens. Environ. 162 (2015) 119–140, https://doi.org/10.1016/j.rse.2015.02.024.
[49] Karbalaye Ghorbanpour Ali, Tim Hessels, Sanaz Moghim, Abbas Afshar, Comparison and assessment of spatial downscaling methods for enhancing the accuracy of satellite-based precipitation over Lake Urmia Basin, J. Hydrol. 596 (2021) 126055, https://doi.org/10.1016/j.jhydrol.2021.126055.
[50] B. Sun, Y. Zhang, Q. Zhou, X. Zhang, Effectiveness of semi-supervised learning and multi-source data in detailed urban landuse mapping with a few labeled samples, Rem. Sens. 14 (2022) 648, https://doi.org/10.3390/rs14030648.
[51] N. Zhao, An efficient downscaling scheme for high-resolution precipitation estimates over a high mountainous watershed, Rem. Sens. 13 (2021) 234, https://doi.org/10.3390/rs13020234.
[52] C. Chen, B. Hu, Y. Li, Easy-to-use spatial Random Forest-based downscaling-calibration method for producing high resolution and accurate precipitation data, Hydrometeorology/Modelling approaches. https://doi.org/10.5194/hess-2021-332, 2021.
[53] Z. Duan, W.G.M. Bastiaanssen, First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure, Rem. Sens. Environ. 131 (2013) 1–13, https://doi.org/10.1016/j.rse.2012.12.002.

[54] A. Arshad, W. Zhang, Z. Zhang, S. Wang, B. Zhang, M.J.M. Cheema, M.J. Shalamzari, Reconstructing high-resolution gridded precipitation data using an improved downscaling approach over the high altitude mountain regions of Upper Indus Basin (UIB), Sci. Total Environ. 784 (2021) 147140, https://doi.org/10.1016/j.scitotenv.2021.147140.
[55] C. Chen, Q. Chen, B. Qin, S. Zhao, Z. Duan, Comparison of different methods for spatial downscaling of GPM IMERG V06B satellite precipitation product over a typical arid to semi-arid area, Front. Earth Sci. 8 (2020) 536337, https://doi.org/10.3389/feart.2020.536337.
[56] W. Jing, Y. Yang, X. Yue, X. Zhao, A spatial downscaling algorithm for satellite-based precipitation over the Tibetan plateau based on NDVI, DEM, and land surface temperature, Rem. Sens. 8 (2016) 655, https://doi.org/10.3390/rs8080655.
[57] Y. Jing, L. Lin, X. Li, T. Li, H. Shen, An attention mechanism based convolutional network for satellite precipitation downscaling over China, ArXiv abs/2203.14812 (2022) 613, https://doi.org/10.48550/arXiv.2203.14812.
[58] C. Chen, Q. Chen, B. Qin, S. Zhao, Z. Duan, Comparison of different methods for spatial downscaling of GPM IMERG V06B satellite precipitation product over a typical arid to semi-arid area, Front. Earth Sci. 8 (2020) 536337, https://doi.org/10.3389/feart.2020.536337.
[59] D.V. Spracklen, S.R. Arnold, C.M. Taylor, Observations of increased tropical rainfall preceded by air passage over forests, Nature 489 (2012) 282–285, https://doi.org/10.1038/nature11390.
[60] W. Jing, Y. Yang, X. Yue, X. Zhao, A spatial downscaling algorithm for satellite-based precipitation over the Tibetan plateau based on NDVI, DEM, and land surface temperature, Rem. Sens. 8 (2016) 655, https://doi.org/10.3390/rs8080655.
[61] K.E. Trenberth, D.J. Shea, Relationships between precipitation and surface temperature: precipitation and temperature relations, Geophys. Res. Lett. 32 (2005), https://doi.org/10.1029/2005GL022760 n/a-n/a.