Theoretical and Applied Climatology (2026) 157:130
https://doi.org/10.1007/s00704-026-06041-0

RESEARCH

Comparative analysis of spatial interpolation methods for daily rainfall data in complex terrain

Lu Jiang[1,2] · Qinggaozi Zhu[2,3] · Xihua Yang[3,4] · Genghong Wu[1,2] · Buddhi Dayananda[5] · Jiaojiao Tan[2] · Qiang Yu[2]

Received: 3 September 2025 / Accepted: 14 January 2026

Abstract

The Loess Plateau in China is considered one of the most ecohydrologically sensitive regions globally, primarily due to its significant spatial and temporal variability in rainfall. Accurately obtaining the spatial distribution of precipitation is crucial for hydrological simulation, ecological restoration and disaster warning. Using the daily rainfall observation of 384 meteorological stations and SRTM elevation data in the Loess Plateau from 1980 to 2020, we systematically evaluated the performance of three typical interpolation techniques including Thin Plate Spline Interpolation (TPS), Inverse Distance Weighting (IDW), and Co-kriging (elevation as covariate) along with three machine learning methods including Random Forest (RF), Support Vector Machine (SVM) and Gaussian Process Regression (GPR). The training set and the validation set were divided using stratified sampling. We assessed the accuracy of different methods in interannual variation, seasonality and ecological zoning scale. The results show that TPS (RMSE = 2.76 mm/d, R2 = 0.71) and IDW (RMSE = 2.75 mm/d, R2 = 0.71)have the best overall performance. The accuracy of the Co-kriging method (R2 = 0.52) is notably compromised in areas of significant elevation change. Conversely, the machine learning method (with R2 ranging from 0.61 to 0.67) demonstrates an advantage in capturing the influence of elevation but tends to underestimate extreme rainfall values. The interpolation uncertainty exhibits seasonal and zonal differences; the largest errors occur in summer (mean RMSE = 5.98 mm/d) and in the gully-dominated regions of the Loess Plateau (Zone A1), while the highest accuracy observed in the sandy and irrigated agricultural areas (Zone C).

Highlights

- TPS and IDW provide the best overall rainfall interpolation accuracy across the Loess Plateau.
- Machine learning methods better capture elevation effects but underestimate extreme rainfall.
- Interpolation errors are highest during summer months in July–August.
- Interpolation accuracy varies across eco-zones, lower in gully regions and higher in agricultural areas.

✉ Qinggaozi Zhu
esther.zhu@dcceew.nsw.gov.au

✉ Qiang Yu
yuq@nwsuaf.edu.cn

[1] State Key Laboratory of Soil and Water Conservation and Desertification Control, College of Soil and Water Conservation Science and Engineering, Northwest A&F University, Yangling 712100, Shaanxi, China

[2] State Key Laboratory of Soil and Water Conservation and Desertification Control, Institute of Soil and Water Conservation, Northwest A&F University, Yangling 712100, Shaanxi, China

[3] New South Wales Department of Climate Change, Energy, the Environment and Water, Parramatta, NSW, Australia

[4] School of Life Sciences, Faculty of Science, University of Technology Sydney, Broadway, NSW, Australia

[5] School of Agriculture and Food Sciences, The University of Queensland, Brisbane, QLD 4072, Australia

Published online: 10 February 2026

1 Introduction

Rainfall is a fundamental driving factor in hydrological and soil erosion models, and its spatial distribution plays a crucial role in flood mitigation strategies, drought warning, ecological restoration and agricultural production (Lopes 2013; Zhu et al. 2020; Eugenio et al. 2014). Accurately understanding and simulating the distribution of daily rainfall is also crucial for assessing hydrological extreme events in the context of global climate change (Feng et al. 2025). However, due to constraints imposed by topography, climate and uneven distribution of meteorological stations (Zhu et al. 2025), observed rainfall data often suffer from discontinuities and insufficient local representativeness (Camera et al. 2014; Chen et al. 2025). To address these limitations, spatial interpolation methods are widely applied to derive high-resolution rainfall distribution patterns. For example, in eco-hydrologically sensitive regions that experience strong spatiotemporal variability in rainfall and severe soil erosion, accurate simulation of rainfall distribution is crucial for advancing ecological restoration objectives and promoting sustainable agricultural development strategies.

Rainfall spatial interpolation methods can be broadly classified into deterministic methods (e.g., Inverse Distance Weighting (IDW), Thin Plate Spline (TPS)) and geostatistical methods (e.g., Kriging and its variants). Among deterministic methods, the IDW technique is based on the Tobler's 'first law of geography', where weights are assigned based on a power function. This method is most effective in regions with a uniform distribution of stations and strong spatial autocorrelation, and it is favoured due to its simplicity and computational efficiency. Das et al. (2017) compared the rainfall interpolation effects of the Kriging method, IDW and spline method in West Bengal, India, and found that IDW performed best in weekly rainfall interpolation. Similarly İlker et al. (2019) demonstrated that the IDW method can provide reasonable results at most stations, especially in the Central China Sea while Fung et al. (2022) highlighted the advantages of IDW in short-time scale interpolation. Yang (2015) further confirmed that IDW performed well in high-resolution (1–8 km) rainfall interpolation scenarios. The spline function method utilises the spatial variation trend through mathematical functions, and can generate a smooth rainfall distribution surface, which is suitable for large-scale spatial interpolation. Plouffe et al. (2015) evaluated IDW, TPS, ordinary Kriging (OK) and Bayesian Kriging (EBK) based on Sri Lanka's agricultural ecological monitoring data, and found that TPS had the highest accuracy in high rainfall areas, while EBK was better in low rainfall areas. The study of Stalenberg et al. (2018) in Madagascar further verified that TPS can effectively improve the spatial consistency of climate data, especially in long-term climate trend analysis. Lyra et al. (2018) compared five interpolation methods (IDW, NRN, TLI, NN, SPT) in Rio de Janeiro, Brazil, and found that the thin plate spline method (SPT) achieved the highest accuracy during transitional seasons (summer to early autumn), effectively representing rainfall heterogeneity under complex terrain.

Geostatistical methods such as Kriging and its variants quantify spatial correlation through semivariograms, which can effectively integrate covariate information and improve interpolation accuracy. Pellicone et al. (2018) and Liu et al. (2021) demonstrated that Kriging methods often outperform deterministic approaches in rainfall spatial interpolation. For example in the Iberian Peninsula, the long-term rainfall interpolation accuracy of the UK _ gauss method is the highest (Ruiz-Ortiz et al. 2024). Zhu and Jia (2004) showed that in the Chaobai River Basin, the uncertainty of Kriging decreased significantly as the number of stations increased. In Portugal's Guadiana River Basin, Fagandini et al. (2024) found that OK outperformed the traditional FAO method for daily rainfall interpolation, while Guidoum (2025) showed that Regression Kriging (RK) provided the best performance for annual rainfall mapping in the Chott El Hodna Basin.

In recent years, machine learning techniques have demonstrated significant advantages in meteorological data interpolation due to their strong nonlinear fitting ability and adaptive learning mechanism (Zagorecki et al. 2013; Chang and Guo 2020; Park et al. 2020; Hou et al. 2025). For example, Pinthong et al. (2024) compared six machine learning algorithms and four traditional interpolation methods and found that machine learning approaches provided superior performance in estimating monthly rainfall. However, there are significant differences in the performance of various machine learning algorithms when addressing spatial prediction problems (Hou et al. 2024). The RF OK method proposed by Li et al. (2011) combined random forest with the Kriging method, leading to a substantial reduction in interpolation error. Moreover, Nobrega and Barroca Filho (2025) found that random forest (RF) had the highest interpolation accuracy for annual maximum daily rainfall ($R^2 = 0.707$) in Brazil's semi-arid region. Rodríguez-Carrillo et al. (2025) demonstrated that daily rainfall estimation in the semi-arid basin of Mexico, both random forest and artificial neural network models yielded the most accurate predictions, with data augmentation techniques further enhancing model performance. In the context of monthly rainfall prediction in the eastern Mediterranean region of Turkey, Sattari et al. (2020) found that Vector Regression (SVR) performed better than other methods. Similarly, Achite et al. (2024) analyzed the data of 150 stations in northern Algeria and found that random forest had the best training performance in annual rainfall prediction. Collectively, these studies confirm the effectiveness of machine learning approaches for

rainfall interpolation across various spatial and temporal scales and highlight their potential for simulating the spatial distribution of rainfall in environmentally complex regions.

As a typical eco-hydrological sensitive area in China, the Loess Plateau has large spatial and temporal variability of precipitation and serious soil erosion. The accurate simulation of rainfall distribution is very important for ecological restoration and sustainable agricultural development. Early research on rainfall interpolation in this area primarily relied on traditional deterministic methods such as spline, IDW, polynomial interpolation and radial basis function (RBF). However, Meng et al. (2006) found that these methods often failed to account for the spatial autocorrelation of annual rainfall interpolation, resulting in relatively high interpolation errors. In contrast, the Kriging method which incorporates spatial structure through semivariogram modeling, significantly improved accuracy. This finding was further supported by Zhao and Yang (2012) demonstrated that the ordinary Kriging method had the smallest error (RMS = 43.50 mm) when applied to data from 108 meteorological stations. Yu et al. (2015) found that the OK method (RMSE = 43.79, MAE = 34.61) was superior to IDW and RBF in the annual rainfall interpolation of the Loess Plateau, with the circular semivariogram model yielding the best performance. More recent studies have incorporated topographic factors to further enhance interpolation accuracy. For instance, Han et al. (2023), proposed the ADAGW method, which integrates aspect and elevation gradient information. This approach showed strong performance across multiple temporal scales (daily, monthly, and annual R^2 were 0.74, 0.92, and 0.87, respectively), and RMSE was 10% −30% compared to traditional interpolation methods.

This study addresses two critical scientific gaps in the spatial interpolation of rainfall over the Loess Plateau: (1) the majority of existing research emphasizes annual-scale analyses, with limited systematic evaluation of interpolation accuracy at the daily scale, particularly because of the high spatial variability, intermittency, and computational demands involved; and (2) the role of terrain variation across different ecological zones in influencing interpolation performance remains insufficiently explored. In response, the objectives of this study are threefold: 1) to compare six spatial interpolation methods for daily rainfall estimation on the Loess Plateau; 2) to assess their accuracy across seasons and ecological zones; 3) to identify optimal methods for improved rainfall mapping.This study provides distinct scientific novelty through its focus on: 1) Daily-scale systematic comparison across ecological zones; 2) Error behavior under elevation gradients.

2 Study area and dataset

2.1 Study area

The Loess Plateau is located in north-central China spanning from 33°N to 41°N latitude and 100°E to 114°E longitude, covering an area of approximately 640,000 square kilometres. The region exhibits a pronounced elevation gradient, ranging from 100 to 5,000 m, with a topography that generally slopes from the highlands in the northwest to the lowlands in the southeast. The Loess Plateau is characterized by a typical temperate continental monsoon climate, with an average annual rainfall of 150–750 mm, the majority of which occurs between July and September. This results in a distinct seasonal pattern of wet summers and autumns, contrasted with dry winters and springs. Due to its unique geographic and climatic conditions, the Loess Plateau is both a critical zone for soil erosion research and a focal area for national ecological restoration initiatives.

To reduce the prediction error associated with the'edge effect' in spatial interpolation (Bajat et al. 2013), a buffer zone with a bandwidth of 50 km around the Loess Plateau was considered. Accordingly, rainfall data were collected from 299 meteorological stations within the Loess Plateau and an additional 85 stations located in the surrounding buffer zone (Fig. 1). The spatial distribution of daily rainfall across the Loess Plateau from 1980 to 2020 was jointly interpolated (Table 1).

2.2 Dataset

This dataset was obtained from daily observation records from meteorological stations nationwide provided by the National Meteorological Science Data Center (https://data.cma.cn/). It covers the period from 1980 to 2020 and includes multiple meteorological variables, such as temperature, humidity, rainfall, wind speed, sunshine duration, and atmospheric pressure.

In addition, the digital elevation model (DEM) data were obtained from the Shuttle Radar Topography Mission (SRTM) conducted by the U.S. Space Shuttle Endeavour, with a spatial resolution of 90 m. The SRTM data was acquired through the Resource and Environment Data Center of the Chinese Academy of Sciences (http://www.resdc.cn/).

The Loess Plateau ecological zoning data were obtained from the National Earth System Science Data Center, China (http://www.geodata.cn).

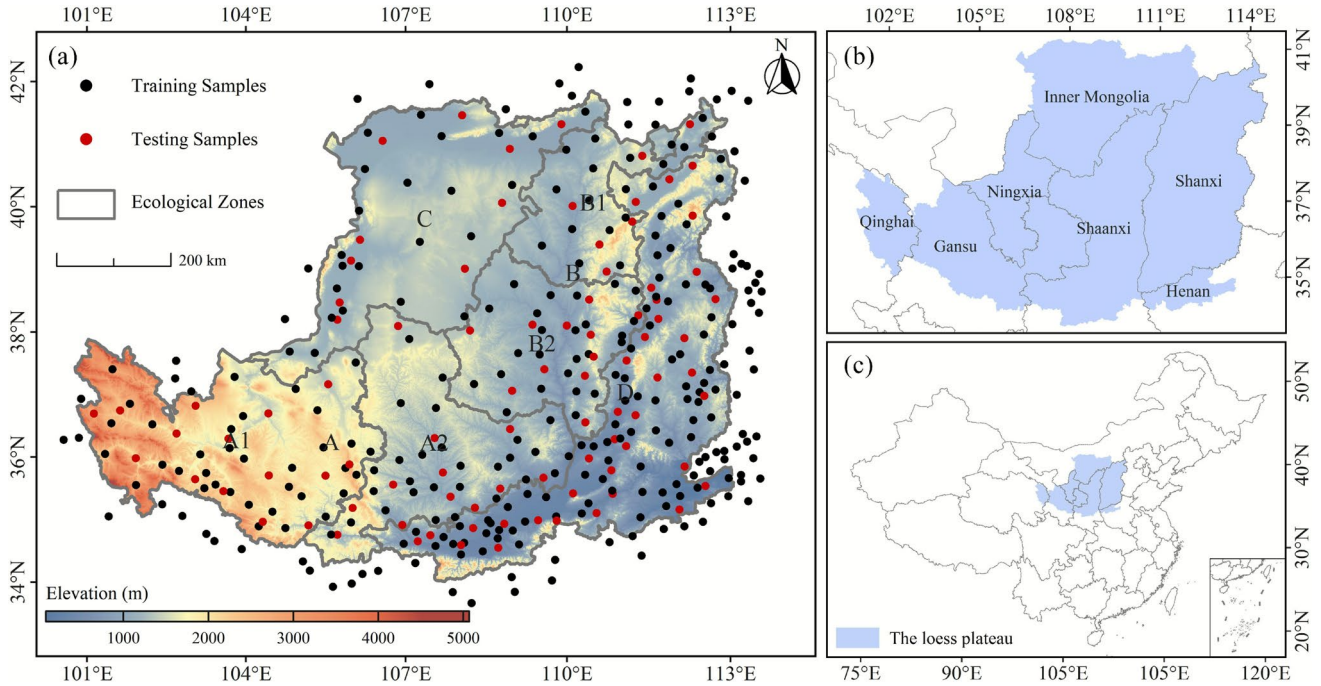

Fig. 1 Research area of the loess plateau. (**a**) Distribution of training and testing samples, ecological zones, and elevation; (**b**) Provincial context of the study area; (**c**) Geographic location of the Loess Plateau in China. Ecological zones: A: Loess Tableland and Gully Region, A1: Loess Tableland and Gully Region A1 Sub-region, A2: Loess Tableland and Gully Region A2 Sub-region, B: Loess Hill and Gully Region, B1: Loess Hill and Gully Region B1 Sub-region, B2: Loess Hill and Gully Region B2 Sub-region, C: Sandy and Agricultural Irrigation Area, D: Rocky Mountain and River Valley Plain Area

Table 1 Data sources and characteristics for the loess plateau study

Data	Years	Spatial range	Resolution/sites	Source	Uses
Meteorological observation data	1980–2020	Loess Plateau and buffer zone	Daily,384sites(299 + 85)	(https://data.cma.cn/)	Rainfall interpolation
SRTM DEM data	2000	Loess Plateau and buffer zone	90m	(http://www.resdc.cn/)	Terrain factors (elevation, slope, aspect)

2.3 Sampling strategy and data partitioning

A stratified sampling approach was applied to partition the data to ensure that both the training and validation sets were representative of the heterogeneous terrain and ecological conditions of the study area. The ecological zoning map (Fig. 1) was used as the stratification criterion. This strategy guarantees that samples from each ecological unit are proportionally represented in both sets, thereby preventing spatial bias and ensuring a robust evaluation of interpolation methods across all landscapes. For each ecological zone, 70% of the meteorological stations were randomly selected for model training, while the remaining 30% were held out for independent validation. The distribution of training and testing samples across all ecological zones is detailed in Table 2.

Table 2 Distribution of training and testing samples across ecological zones

Ecological zone	Total stations	Training stations (70%)	Testing stations (30%)
A1	51	36	15
A2	40	28	12
B1	21	15	6
B2	36	25	11
C	33	23	10
D	118	83	35
Total	299	210	89

3 Methodology

3.1 Research framework

The technical workflow of this study is illustrated in Fig. 2: Based on the meteorological stations and elevation data of the Loess Plateau and its surrounding areas, the training set and the validation set are divided by stratified sampling; six methods of TPS, IDW, Co-kriging, RF, SVM and GPR were used for spatial interpolation. Finally, the accuracy of each method was evaluated by MAE, RMSE and R^2, and its spatial and temporal distribution characteristics were analyzed.

3.2 Thin plate spline interpolation

Thin Plate Spline (TPS) is a spatial interpolation method based on the principle of minimum curvature. It constructs a smooth surface that passes through all observation points using radial basis functions (Hutchinson 1995) to simulate the spatial distribution of meteorological elements while minimizing the overall bending energy of the surface. In this study, the model is implemented using the fields package in R, with its core involving the optimization of the smooth function f via penalized least squares. The smoothing parameter ρ is automatically determined by the package's built-in generalized cross-validation (GCV) procedure to balance fitting accuracy and surface smoothness, thereby preventing overfitting (Hutchinson 1998; Liu et al. 2008, 2012). The theoretical model formula of the thin plate smooth spline model can be expressed as:

$$Z_i = f(x_i) + b^T y_i + e_i \tag{1}$$

$$\text{MIN}: \sum_{i=0}^{n} \left[\frac{Z_i - f(x_i) - b^T y_i}{W_i} \right] + \rho J_m(f) \tag{2}$$

where Z_i is the point to be interpolated at space i; x_i is the d-dimensional spline independent variable, which is the meteorological element value of the known control point around the position i; f is an unknown smooth function about x_i; y_i is an independent covariate; b is the coefficient of independent covariate; e_i is random error. $J_m(f)$ is the roughness measure function of function f, which is the m-order partial derivative of function f; ρ is a positive smoothing parameter, which balances the accuracy of data and the smoothness of the surface.

3.3 Inverse distance weighting

Inverse Distance Weighting (IDW) is a deterministic interpolation method based on spatial distance weighting. In this study, the IDW method was used for spatial interpolation of rainfall. The underlying assumption of IDW is that the value of the unknown point has a spatial correlation with the value of its adjacent known points, and this correlation decreases with the increase in distance. The IDW is a weighted average of the distance, so it is especially suitable for areas with uniform site distribution and moderate density. This study determines through sensitivity analysis (Appendix Table 5) that the parameter combination of a distance attenuation coefficient (p) of 1 and a maximum number of neighboring stations (Nmax) of 5 optimally balances the capture of local

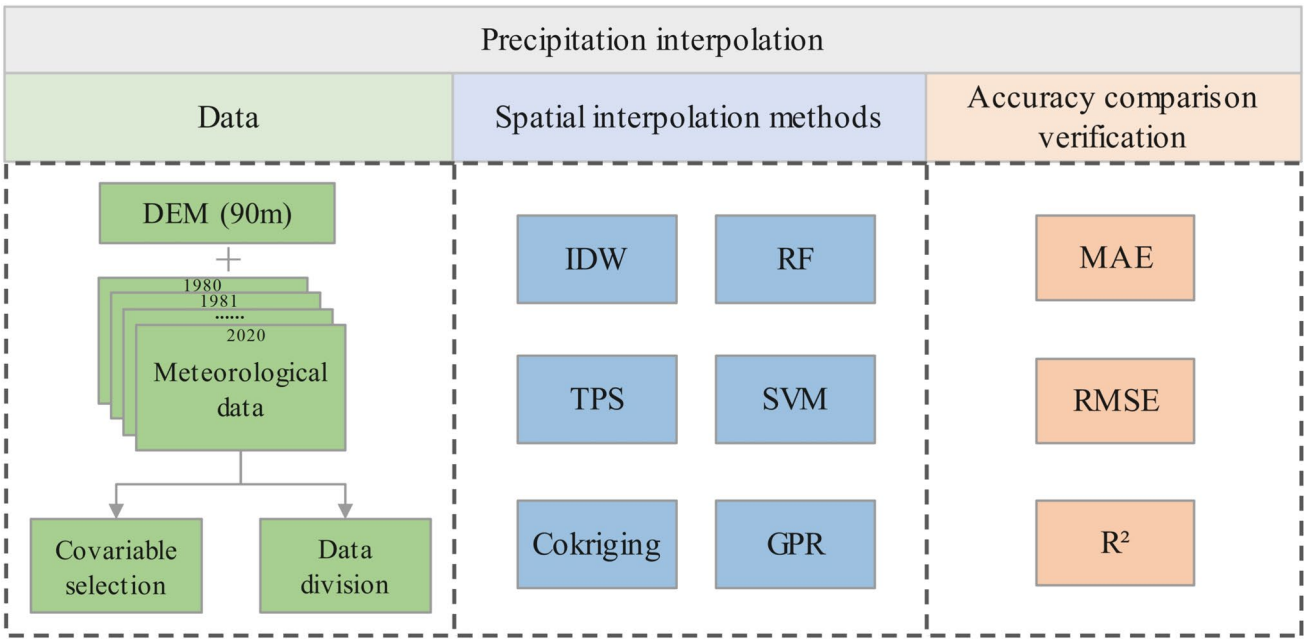

Fig. 2 Flow chart of rainfall interpolation

characteristics with the model's overall performance. The IDW algorithm is as follows, implemented using the gstat package in R:

$$Z = \frac{\sum_{i=1}^{n} \frac{1}{d_i^1} Z(x_i)}{\sum_{i=1}^{n} \frac{1}{d_i^1}} \tag{3}$$

where Z is the grid value of the interpolation point to be simulated; $Z(x_i)$ is the measured value of the $i\,(i = 1,2,3,\ldots,n)$ meteorological station; n is the number of sample points; d_i is the distance from the interpolation point to the i site.

3.4 Co-kriging

Co-kriging is a multivariate geostatistical estimation method developed on the basis of Kriging. It adds auxiliary information that is highly correlated with variables as covariates to the Kriging interpolation process. It not only considers the spatial autocorrelation of independent variables. but also considers the correlation between covariates and independent variables (Knotters et al. 1995). Consequently, it requires the estimation of both the variograms of individual variables and the cross-variograms between the target and covariates, thereby improving prediction accuracy in cases where auxiliary information is available. The method is implemented using the automap and gstat packages in R, with elevation serving as the primary covariate. The semivariogram is automatically fitted using the autofitVariogram function, which evaluates multiple models base on their goodness-of-fit to the sample data (Pebesma 2004; Hiemstra et al. 2009). Co-kriging theoretical model formula (Deutsch and Journel 1997):

$$\widehat{Z}(S_0) = \sum_{i=0}^{N} a_i Z(S_i) + \sum_{j=0}^{M} b_j Y(S_j) \tag{4}$$

where i and j denote the indices of the independent variable and covariate, respectively. S_0 represents the location of the point to be interpolated. S_i and S_j indicate the locations of the observation points for the independent variable and covariate around S_0, respectively. $Z(S_i)$ and $Z(S_j)$ are the measured values of the independent variable and covariate at points i and j, respectively. a_i and b_j are the unknown weights of the independent variable and covariate at points i and j, respectively. $\widehat{Z}$ is the interpolation result at the point S_0. N and M represent the number of measurement points for the independent variable and covariate S_0, respectively.

3.5 Random forest

The random forest spatial interpolation method is a non-parametric spatial prediction technology based on ensemble learning (Breiman 2001). It captures spatial heterogeneity and nonlinear relationships by constructing multiple regression decision trees. The model is implemented using the ranger package in R, configured with 500 trees. The core principle involves using Bootstrap resampling to generate several training subsets. Each subset independently trains the decision tree and introduces random feature selection to reduce the model variance (Breiman 2001), and finally outputs the interpolation result through the predicted mean of the whole tree. This method does not need to preset data distribution assumptions and can adaptively process complex spatial structures. It can automatically identify key environmental covariates (such as elevation, slope, etc.) through feature importance assessment, in which the built-in Out-of-Bag error estimation provides model uncertainty quantification. In mathematical expression, Cutler et al. (2012) defined n as a dimensional random vector $X = (x_1, \ldots, x_n)^T$ representing the input variable, Y as the output variable (rainfall), and its joint distribution is $E_{X,Y} = (X, Y)$. The prediction function $f(x)$ is constructed by minimizing the expected loss function $L(Y, f(x))$, which is represented by the following formula:

$$E_{X,Y}(L(Y, f(X))) \tag{5}$$

In the random forest regression model, when the least squares method is used to minimize the prediction error, its mathematical expression can be expressed as a formula (6).

$$f(X) = E(Y|X = x) \tag{6}$$

This method constructs multiple base predictors ' base learners $'h_1(x), \ldots, h_N(x)$ and integrates their outputs into a ' joint predictor $'f(x)$. Finally, the average prediction value shown in formula (7) is used as the final result for the regression problem.

$$f(x) = \frac{1}{N} \sum_{n=1}^{N} h_n(x) \tag{7}$$

3.6 Support Vector machine

Support vector regression (SVR) is a non-parametric spatial prediction method based on statistical learning theory (Drucker et al. 1996). Its core idea is to map low-dimensional space to high-dimensional feature space through kernel function, and construct the optimal regression hyperplane

in this space. This method is particularly suitable for processing spatial data with complex nonlinear relationships and can still maintain good performance in the case of small samples (Salcedo-Sanz et al. 2016). In this study, SVR is used for spatial interpolation of rainfall, which is characterized by elevation, latitude and longitude, and rainfall is the target. By optimizing the regularization parameter $C = 10$, insensitive parameter $\varepsilon = 0.1$ and the kernel parameter γ for the radial basis function (RBF) is automatically and heuristically determined by the built-in procedure of the e1071 package (Zhang 2024). the optimal parameter combination is determined by cross validation to ensure the generalization ability of the model. The mathematical model of SVR is achieved by minimizing the following objective function:

$$\min_{w} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{8}$$

$$\text{subject to} \begin{cases} y_i - w^T x_i - b & \leq \varepsilon + \xi_i \\ w^T x_i + b - y_i & \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* & \geq 0 \end{cases} \tag{9}$$

where w is the weight vector, x_i is the vector of the input value, b is the bias term, ε is the insensitive loss parameter, and y_i is the actual output value. C is the regularization parameter, and ξ_i and ξ_i^* are relaxation variables (Zhang and ODonnell 2020). SVR uses kernel trick to deal with nonlinear problems. The commonly used radial basis function (RBF) kernel is defined as:

$$K(x, x') = \begin{cases} RFB \\ exp\left(-\gamma\|x - x'\|^2\right) \end{cases} \tag{10}$$

where γ represents the distance threshold from the hyperplane to the support vector, which is used to effectively distinguish different types of samples. $\|x - x\prime\|^2$ represents the Euclidean square distance from the sample point to the hyperplane (Raghavendra and Deka 2014).

3.7 Gaussian process regression

Gaussian process regression (GPR) is a nonparametric spatial interpolation method based on Bayesian framework, which shows unique advantages in the field of meteorology and hydrology (Chilès and Delfiner 2012). This method regards the rainfall field in the target area as the realization of the Gaussian random process, and characterizes the spatial dependence by the covariance function, including the square exponential function or the Matérn function (Keriven et al. 2018). The model is implemented using the kernlab package in R, employing the Gaussian kernel with automatic parameter determination (Zhang and Liu 2025). Its core model can be expressed as:

$$Y = f(X) + \xi \tag{11}$$

where Y is the observed value; $f(X)$ is the underlying random function; X is the input vector; ξ is the Gaussian noise of variance σ_n^2 (here,$\sigma_n^2 = 0.1$),$\xi\ N\left(0, \sigma_n^2\right)$. In Gaussian process regression, each input variable x has a random variable $f(x)$, that is, the value of the random function at the x position. The posterior distribution of the observed value Y and the joint prior distribution of the observed value Y and the predicted value y are:

$$Y \sim N\left(0, K(X, X) + \sigma_n^2 I_n\right) \tag{12}$$

$$\begin{bmatrix} Y \\ y \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X,X) + \sigma_n^2 I_n & K(X, x_*) \\ K(x, X) & K(x_*, x_*) \end{bmatrix}\right) = N\left(0, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix}\right) \tag{13}$$

where $K(X, X) = (K_{ij})$ is a symmetric positive definite covariance matrix, and the matrix element (K_{ij}) measures the correlation between K_i and K_j through the kernel function.$K(x_*, X) = K(X, x_*)^T$ is the covariance matrix of test set x and training set X. $K(x_*, x_*)$ is the covariance of the test set x_* itself. I_n is an n-dimensional unit matrix. Then the posterior distribution of the predicted value y can be obtained, $\overline{y}$ and i are the mean and variance of the predicted value y, respectively.

3.8 Validation strategy

To comprehensively evaluate the performance and spatial robustness of the interpolation methods, this study adopt a two-tier validation strategy. First, a standard random splitting procedure is employed, allocating 70% of the meteorological stations to the training set and the remaining 30% to the test set, ensuring a statistically representative data partition for assessing overall predictive accuracy. In order to further evaluate whether the performance of the model is affected by the spatial clustering of the data, this study uses the global Moran 's index to perform spatial autocorrelation analysis on the interpolation residuals to verify the spatial independence of the model errors (Asokan et al. 2025), thus providing a statistical basis for evaluating the effectiveness of the random segmentation verification scheme in this study.

3.9 Evaluation

Cross-validation was employed to evaluate the accuracy of the spatial interpolation. Specifically, 87 stations from the full dataset are randomly selected as test samples, and the

remaining stations are used as training samples for spatial interpolation. The interpolation algorithms simulate the rainfall values at the test sites, assuming the true values are unknown and the performance is assessed by comparing these simulated values with the actual measurements. The evaluation metrics include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and Pearson Correlation Coefficient (CC). The formulas are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Z_{ai} - Z_{\lambda i}| \tag{14}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Z_{ai} - Z_{\lambda i}\right)^2} \tag{15}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Z_{ai} - Z_{\lambda i}\right)^2}{\sum_{i=1}^{n}\left(Z_{ai} - \overline{Z}_a\right)^2} \tag{16}$$

$$CC = \frac{\sum_{i=1}^{n}\left(Z_{ai} - \overline{Z}_a\right)\left(Z_{\lambda i} - \overline{Z}_\lambda\right)}{\sqrt{\sum_{i=1}^{n}\left(Z_{ai} - \overline{Z}_a\right)^2}\sqrt{\sum_{i=1}^{n}\left(Z_{\lambda i} - \overline{Z}_\lambda\right)^2}} \tag{17}$$

where n is the number of sites in the test sample; Z_{ai} is the measured value of the—ith site; $Z_{\lambda i}$ is the estimated value of the i—th site. $\overline{Z}_a$ is the average value of the measured values of all sites. $\overline{Z}_\lambda$ is the average value of the interpolated estimated values. n is the number of sites involved in verification.

4 Results

4.1 Comparison of interpolation accuracy of different methods

The variable selection strategy is determined through a preliminary evaluation of covariate combinations (Appendix Table 6). The results indicate that using only longitude and latitude generally yields superior statistical metrics across all methods. Nonetheless, given the significant topographic relief of the Loess Plateau, elevation is a physically critical determinant of rainfall patterns. Therefore, to ensure physical rationality and to maintain a consistent basis for comparing the ability of different methods to incorporate topographic information, we adopt a unified set of covariates (longitude, latitude, and elevation) for the Co-kriging, RF, SVM, and GPR methods. In contrast, the TPS and IDW methods maintain the traditional two-dimensional approach (longitude and latitude only). This variable selection strategy not only ensures the rationality and comparability between methods but also takes into account the actual needs of spatial interpolation in complex terrain areas.

This study systematically evaluated the performance of six spatial interpolation methods in daily rainfall prediction. Based on observed and predicted values (Table 3), the actual range of rainfall predicted by each method is TPS: 0–197.97 mm/d; IDW: 0–189.58 mm/d; Co-kriging: 0–391.19 mm/d; RF: 0–146.45 mm/d; SVM: 0–97.98 mm/d; GPR: 0–119.14 mm/d. In terms of accuracy evaluation, both MAE and RMSE indicated that IDW performed best (0.68 mm/d and 2.75 mm/d), followed by TPS (0.70 mm/d and 2.76 mm/d). Notably although the MAE of Co-kriging is at a medium level (0.80 mm/d), its RMSE is significantly higher (3.58 mm/d), which may be due to the sensitivity of the method to extreme rainfall events. The correlation analysis indicated that the traditional methods showed stronger correlation while maintaining the prediction accuracy. The correlation coefficient of TPS and IDW was the highest (CC = 0.85). In machine learning methods, RF and GPR (CC = 0.82 and 0.81) performed better than SVM (CC = 0.79). The results suggest that the traditional spatial interpolation methods (TPS and IDW) perform well in all indicators, especially in maintaining a high correlation coefficient while achieving a low prediction error. In contrast, although some machine learning methods (such as RF and GPR) have achieved acceptable results, their overall performance remained inferior to that of the traditional interpolation technique.

Table 3 Six spatial interpolation schemes for four statistical indicators (CC, R^2, RMSE and MAE)

Models	Variables	CC	MAE (mm/d)		RMSE (mm/d)		R^2	
			Value	95% CI	Value	95% CI	Value	95% CI
TPS	Lon, lat	0.845***	0.698	[0.693, 0.703]	2.758	[2.754, 2.763]	0.713	[0.704, 0.722]
IDW	Lon, lat	**0.846*** **	**0.678**	**[0.673, 0.682]**	**2.751**	**[2.746, 2.755]**	**0.715**	**[0.706, 0.724]**
Co-kriging	Lon, lat, elevation	0.752***	0.797	[0.791, 0.802]	3.584	[3.578, 3.590]	0.516	[0.495, 0.536]
RF	Lon, lat, elevation	0.820***	0.793	[0.788, 0.798]	2.974	[2.969, 2.979]	0.667	[0.657, 0.676]
SVM	Lon, lat, elevation	0.785***	0.774	[0.768, 0.779]	3.237	[3.231, 3.242]	0.605	[0.594, 0.617]
GPR	Lon, lat, elevation	0.813***	0.805	[0.799, 0.809]	2.998	[2.992, 3.002]	0.661	[0.652, 0.671]

The symbol *** indicates that the correlation between observations and predictions is statistically significant at the 0.1% significance level ($p < 0.001$). Bold indicates the lowest MAE and RMSE values for the best interpolation scheme (highest CC and R^2). Lon, lat represent longitude, latitude respectively.95% CI represents 95% confidence interval

4.2 Performance of interpolation methods on extreme rainfall indices

Based on the 11 extreme rainfall indices recommended by the World Meteorological Organization (WMO) (Wu et al. 2016; Wang et al. 2017), the prediction performance of six spatial interpolation methods for extreme rainfall indices was systematically evaluated. According to the performance of each index (Fig. 3), the traditional method is better than the machine learning method, among which TPS and IDW are the most prominent, and the average correlation coefficients (r) of the two are 0.79 and 0.78, respectively. These two traditional methods show stable high-performance in most extreme rainfall index predictions. Among the machine learning methods, GPR performed best (average r = 0.75), especially in terms of SDII (0.78) and Rx5day (0.73). RF performed best in PRCPTOT prediction (r = 0.88), and support vector machine (SVM) had a unique advantage in CWD prediction (r = 0.67). Notably, Co-kriging demonstrated the weakest overall performance, with an average correlation coefficient of 0.65.

Further analysis indicated that different interpolation methods exhibit distinct strengths depending on the type of extreme rainfall index. The predictions for CDD and R10 were relatively insensitive to the choice of method, with correlation coefficients remaining stable across models, ranging from 0.77–0.79 and 0.80–0.81, respectively. Notably, for PRCPTOT prediction, machine learning methods generally outperformed traditional approaches, with RF achieving the highest correlation coefficient of 0.88.

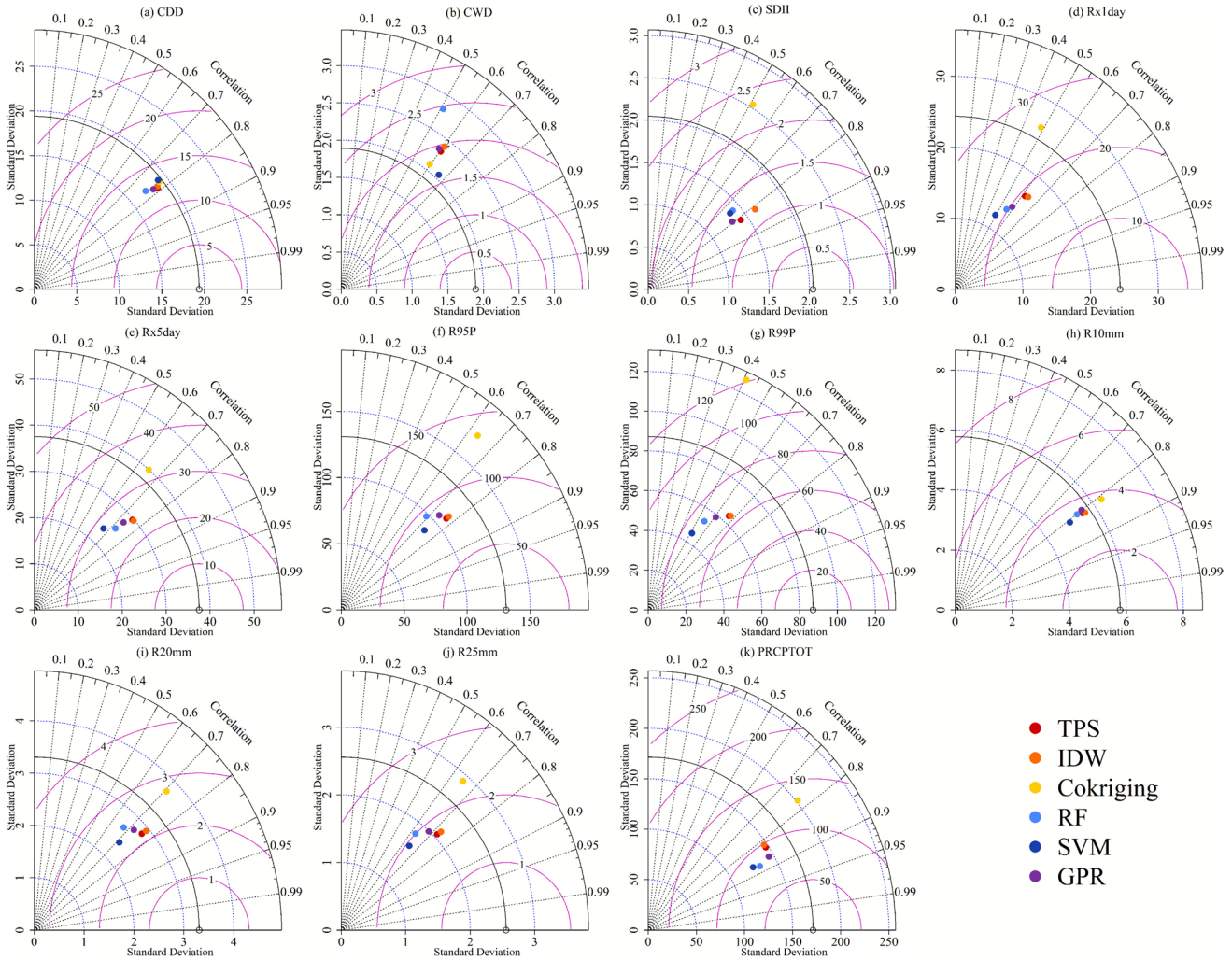

Fig. 3 The 11 extreme rainfall indices were continuous drought days CDD (d), continuous wet days CWD (d), rainfall intensity SDII (mm/d), single-day maximum rainfall Rx1day (mm), five-day maximum rainfall Rx5day (mm), strong rainfall R95P (mm), extreme rainfall R99P (mm), moderate rainfall days R10 (d), heavy rainfall days R20 (d), strong rainfall days R25 (d), and annual total rainfall PRCPTOT (mm), r is Correlation

4.3 Comparison of verification results of rainfall in different ecological zones

The Fig. 4 shows that the interpolation accuracy of the C ecological area is the best, and the average values of MAE and RMSE are 0.49±0.006 mm/d and 2.23±0.006 mm/d, respectively, which are significantly lower than those of other ecological areas. The average MAE and RMSE were 0.87±0.007 mm/d and 3.25±0.008 mm/d, respectively, and the average R^2 was only 0.45±0.03. The C ecological zone is located in the northwest of the Loess Plateau. The annual average rainfall is only 100–400 mm, and the spatial variability of rainfall events is small. Secondly, the high altitude but flat terrain characteristics of the area reduce the difficulty of spatial interpolation. In most ecoregions (A1, A2, B1, C, D), TPS and IDW performed best, MAE and RMSE in these ecoregions were 7% −50% lower than other methods, and R^2 increased by 0.07% −40%. TPS (MAE=0.71±0.007mm/d, RMSE=2.94±0.007mm/d, R^2=0.73±0.010) was better than IDW (MAE=0.71±0.008mm/d, RMSE=3.05±0.008mm/d, R^2=0.70±0.014) in B2 subregion of loess hilly-gully region. It can be seen from the figure that the Co-kriging method has regional limitations. In the D ecological area (earth-rock mountain area and valley plain area), the Co-kriging method performs the poorest performance. MAE=0.88±0.006 mm/d (22% higher than TPS), RMSE=4.19±0.006 mm/d (50% higher than TPS), R^2=0.45±0.019 (40% lower than TPS). Based on the comprehensive analysis results, it is recommended to adopt differentiated interpolation strategies in different ecological regions: TPS or IDW with higher computational efficiency can be used in (A1, A2, B1, C, D) ecological regions; tPS was preferred in B2 ecological area. For the Co-kriging method, it should be avoided in the D ecological region, but in the C ecological region, its interpolation accuracy is comparable to that of TPS and IDW.

4.4 Seasonal variation of spatial interpolation accuracy of daily rainfall

The Fig. 5 shows that the error has seasonal characteristics. From the trend of MAE and RMSE, it can be seen that the average error (MAE=2.25mm/d, 2.10mm/d) and root mean square error (RMSE=6.12mm/d, 5.84mm/d) of each interpolation method in July and August are significantly larger than those in other months, showing a unimodal distribution. The error gradually increased from January to July, peaked in July, and decreased from August to December. The error in December was the smallest (MAE=0.06 mm/d), which was consistent with the seasonal fluctuation of rainfall. This seasonal pattern is primarily driven by two factors: first, the summer monsoon season brings intense convective rainfall events that are highly localized and spatially heterogeneous, challenging the representativeness of station-based observations; second, these months experience the highest frequency of extreme rainfall, which significantly increases error variance, as evidenced by the upper outliers in the July–August boxplots (July MAE maximum=3.97 mm/d, August MAE maximum=3.63 mm/d).

Comparison of the accuracy indicators of different methods, it can be found that there are differences in the performance of different interpolation methods. The MAE and RMSE medians of TPS and IDW are the lowest throughout the year (such as TPS, IDW July MAE median=2.11mm/d, 2.02mm/d), and R^2 is stable at a high level (TPS, IDW annual R^2 median>0.6). However, the RMSE median of Co-kriging was higher throughout the year, peaking in July (7.15 mm/d). The R^2 median of all methods in July and August decreased by about 0.2 compared with the dry season (TPS decreased from 0.81 in April to 0.61 in July), but TPS and IDW still maintained the best fitting performance (R^2 median 0.63 and 0.64 in August).

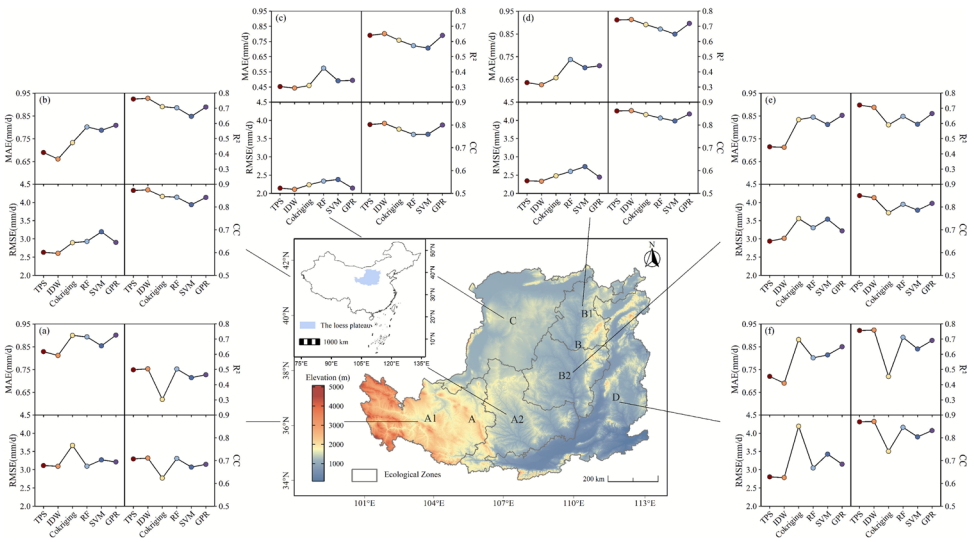

Fig. 4 Comparison of verification results of rainfall in different ecological areas, **A**: Loess Plateau gully area, A1: Loess Plateau gully area A1 sub-area, A2: Loess Plateau gully area A2 sub-area, **B**: Loess Hilly and gully area. B1: Loess Hilly and gully area B1 sub-area, B2: Loess Hilly and gully area B2 sub-area, **C**: Sandy land and agricultural irrigation area, **D**: Earth-rocky mountain area and valley plain area

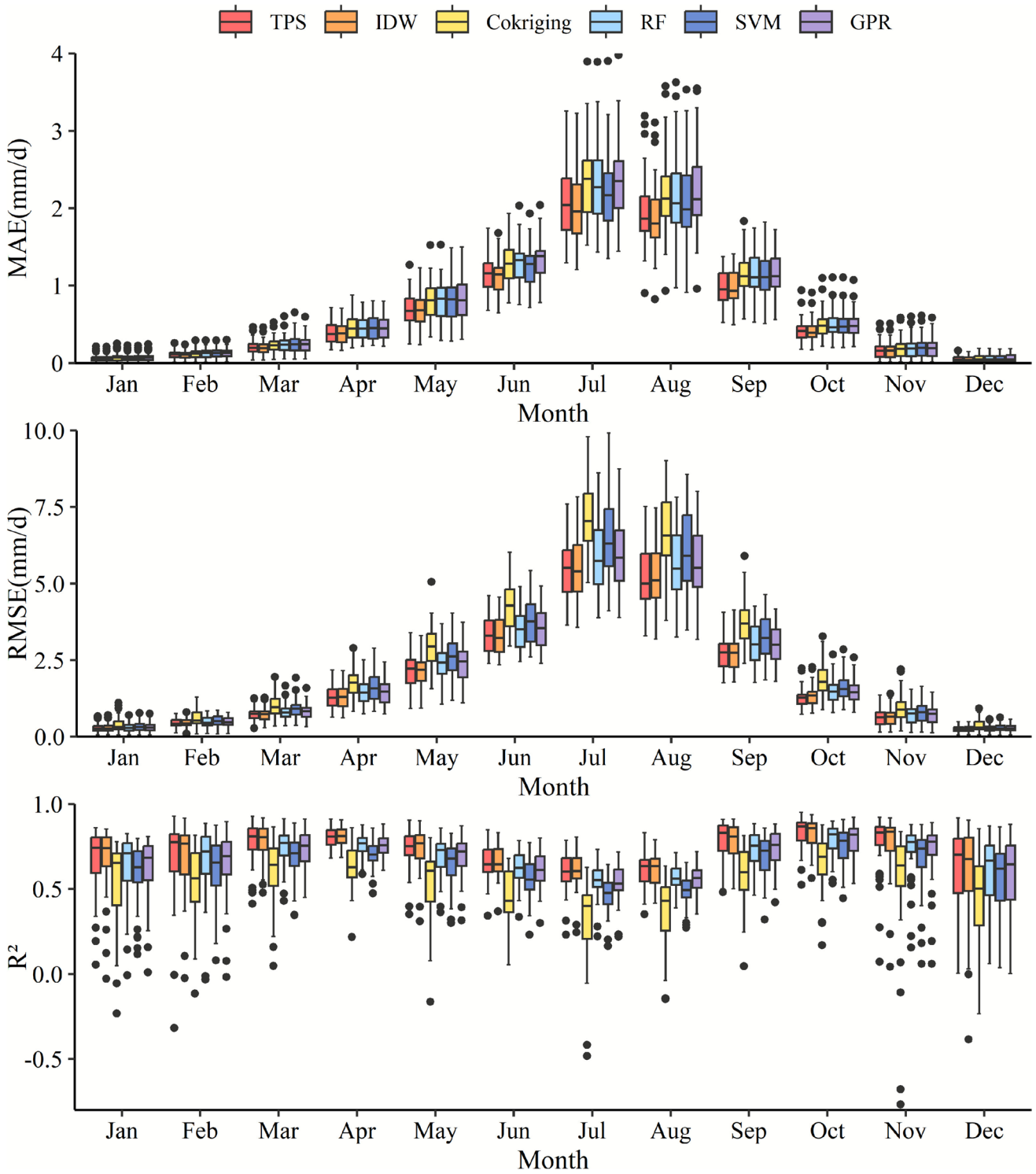

Fig. 5 Seasonal variation of spatial interpolation accuracy of daily rainfall

Collectively, these results indicate that TPS and IDW achieve stronger error control (lower MAE and RMSE) and greater model stability (higher R^2), yielding more reliable interpolation outcomes. Conversely, the reduced accuracy of Co-kriging, particularly during heavy rainfall periods, may reflect its strong reliance on covariate data, which limits its robustness under conditions of high rainfall variability.

4.5 Interannual variation of spatial interpolation accuracy of daily rainfall

Based on the accuracy evaluation of daily rainfall spatial interpolation results from 1980 to 2020 (Fig. 6), the performance differences of six interpolation methods (TPS, IDW, Co-kriging, RF, SVM, GPR) on long time scales were analyzed. Comparative results indicated that TPS and IDW performed best in the spatial interpolation of daily rainfall, and their MAE decreased by 15.9% compared with other methods. The RMSE is reduced by 0.21 ~ 0.47 mm/d compared with other methods, 0.83 mm/d compared with Co-kriging,

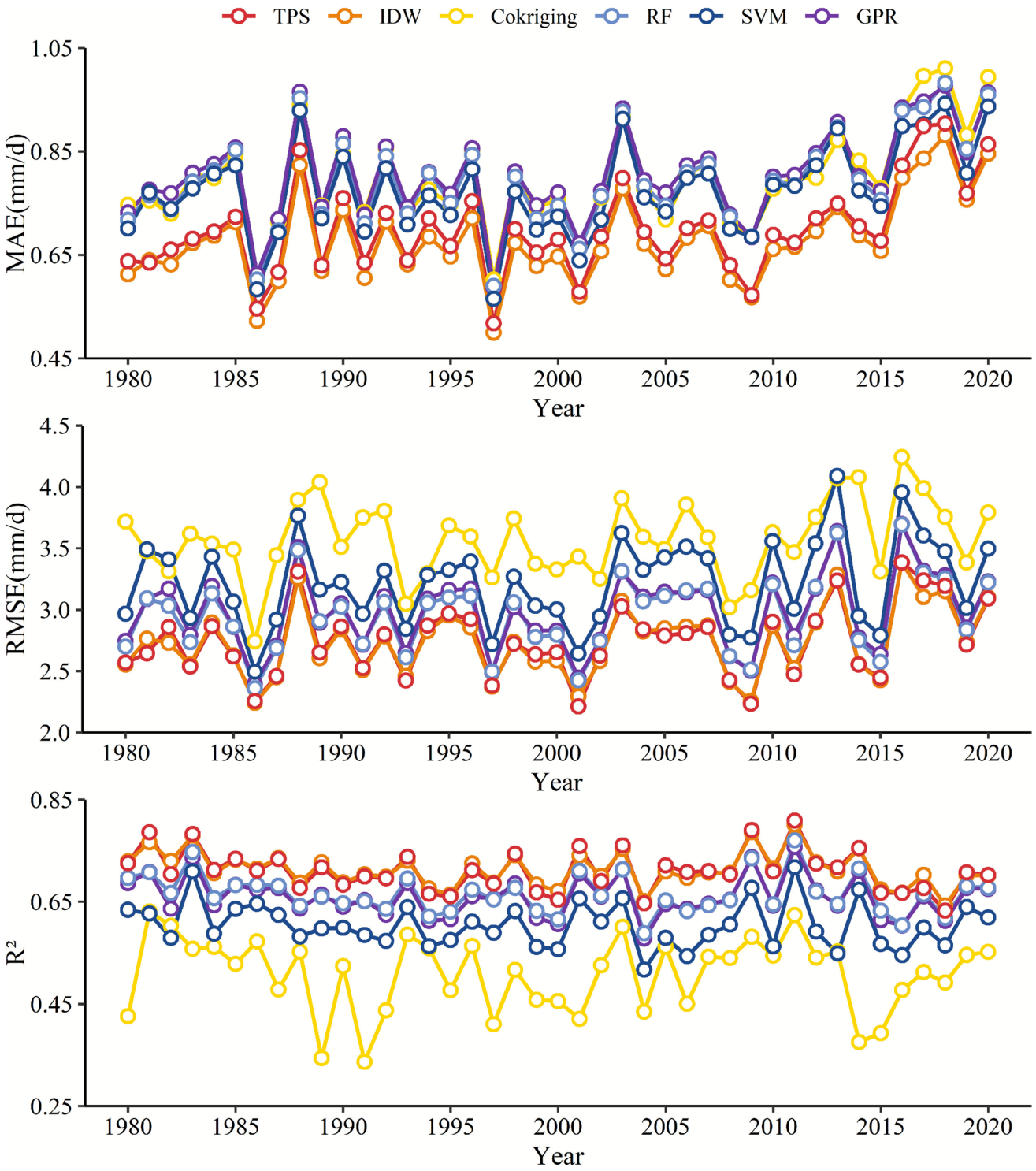

Fig. 6 Interannual variation of spatial interpolation accuracy of daily rainfall

and the performance is improved by 30.3% compared with Co-kriging. From 1980 to 2016, the R^2 of TPS and IDW was basically stable at 0.71 mm/d, and TPS was slightly higher than IDW in a few years. The error indices of RF and GPR were in the middle, the multi-year average values of MAE were 0.79 mm/d and 0.80 mm/d, the multi-year average values of RMSE were 2.95 mm/d and 2.98 mm/d, and the multi-year average values of R^2 were 0.67 and 0.66, respectively. The RMSE of Co-kriging and SVM was high, especially in 1990 and 2014; the RMSE of Co-kriging showed obvious peaks, reaching 4.04 mm/d, 3.75 mm/d and 4.08 mm/d, respectively. Correspondingly, R 2 decreased to 0.34, 0.34 and 0.38, which was the lowest among all methods, indicating poor stability in these approaches.

4.6 Extreme rainfall events and spatial interpolation

In order to systematically evaluate the performance of different spatial interpolation methods during heavy rainfall events on the Loess Plateau, this study selected rainfall data on July 29, 2011 (daily rainfall measured range of 0–111.9 mm), and simulated the regional rainfall distribution with a spatial resolution of 0.001°. The interpolation results of the six methods (Fig. 7) show significant differences. The prediction range of TPS (0–75.53 mm) was lower than the measured range (ratio of predicted-to-observed maxima = 0.68), which may weaken the characteristics of local heavy rainfall due to its smoothing effect. However, its spatial distribution pattern is in good agreement with the measured rainfall field. Although the predicted range of IDW (0–111.9 mm) is completely consistent with the measured value, showing good adaptability to the spatial heterogeneity of rainfall, the interpolation results show a typical " bull 's eye pattern, " that is, a concentric circle of high/low value bands is formed around the site, reflecting the sensitivity of this method to the weight distribution of discrete points. The machine learning methods (RF, SVM, GPR) showed smooth spatial distribution characteristics, and the prediction ranges were 0–70.57 mm, 0–55 mm, and 0–97.63 mm, respectively. Although there is no extreme value deviation, the prediction of heavy rainfall is still underestimated. Co-kriging showed a significant overestimation in the southeastern mountainous areas (A2, B2, D ecological areas) (the black part is the predicted rainfall > 112 mm), and the predicted value in some areas exceeded the measured maximum value by 416.99mm/d. Notably, these abnormally high value areas only account for less than 1% of the total area of the study area, and their spatial distribution is highly consistent with the terrain mutation area (elevation gradient > 176.5 m/km), suggesting that the synergistic effect of terrain factors and extreme rainfall may lead to model failure. It is speculated that the model fails under the combined action of elevation mutation and extreme rainfall.

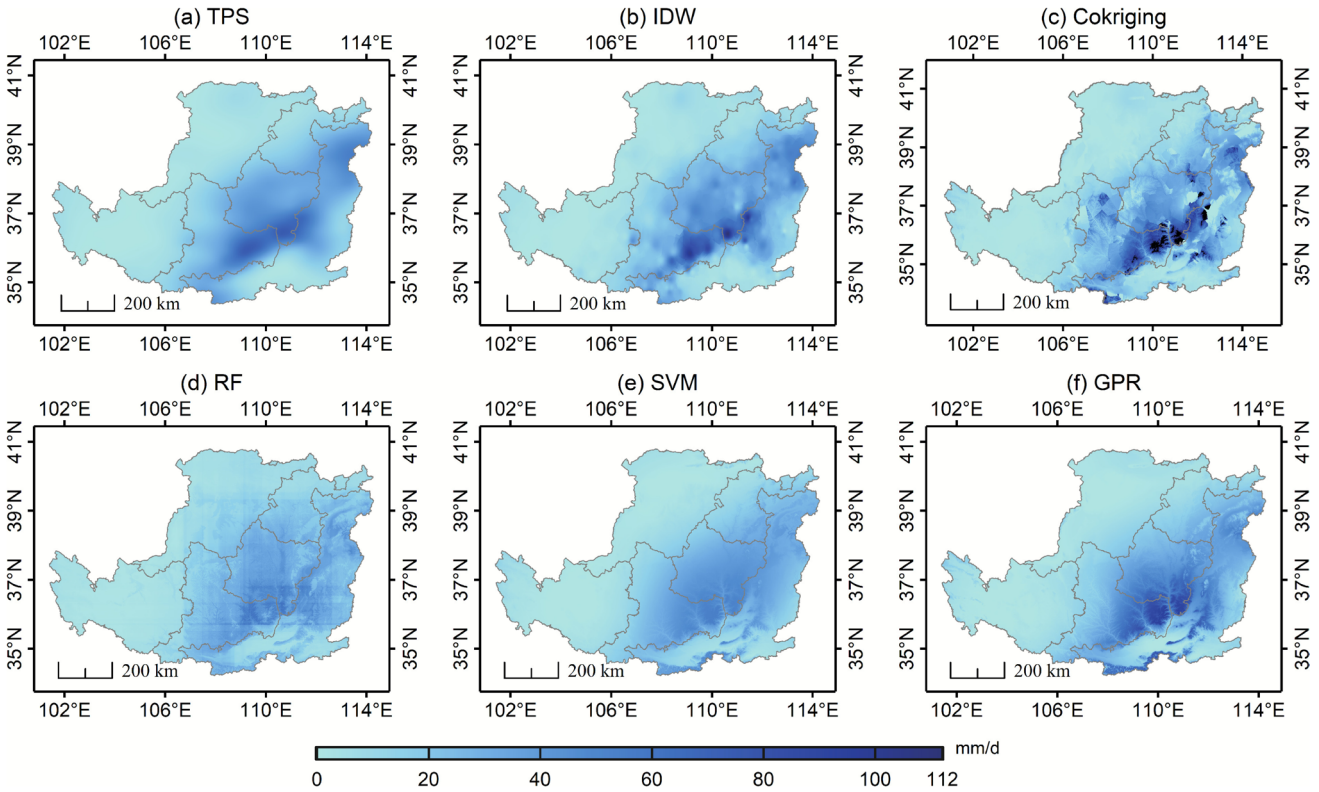

Fig. 7 Spatial distribution of rainfall on July 29, 2011 interpolated by different methods

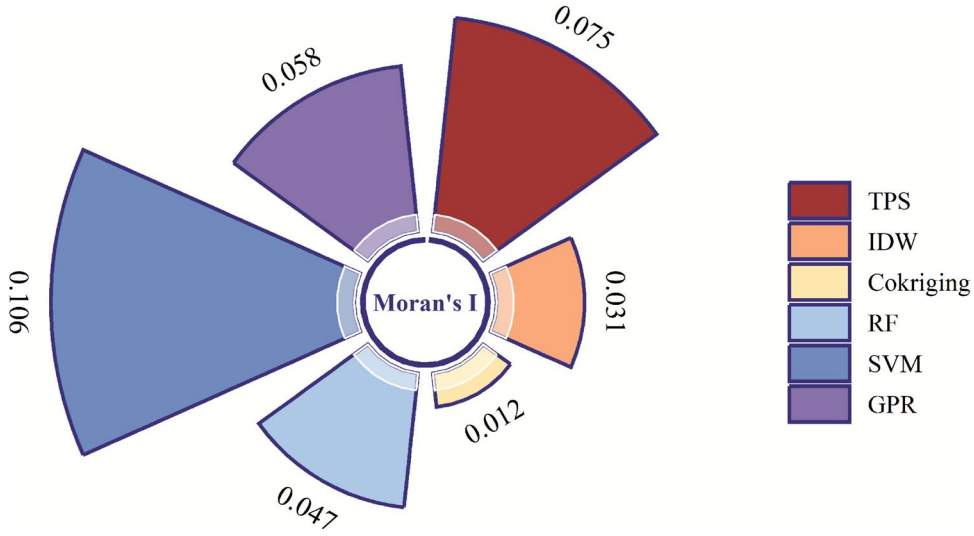

Fig. 8 Spatial autocorrelation of interpolation errors measured by Moran's I

4.7 Spatial autocorrelation analysis of interpolation errors

The analysis demonstrates that all interpolation methods (TPS, IDW, Co-kriging, RF, SVM, and GPR) show non-significant spatial autocorrelation in their errors (global Moran's I: 0.012–0.106, $P>0.25$), indicating spatially random error distributions and strong spatial robustness across all approaches. Among these, SVM exhibited relatively higher spatial dependency (global Moran's I: 0.106), yet still within the non-significant range, suggesting consistent spatial randomness in error patterns (Fig. 8).

The spatially random distribution of errors demonstrates that model performance is not artificially inflated by spatial clustering in the data, thereby supporting the validity of the conventional random split validation for this specific study.

5 Discussion

5.1 Uncertainty analysis and abnormal point recognition of Co-kriging

The comparative analysis of the six interpolation methods revealed that the Co-kriging approach exhibited notable deficiencies in the spatial interpolation of daily rainfall. This method not only produces a higher average error compared to traditional methods such as TPS and IDW, but also has the maximum root mean square error (RMSE = 3.58 mm/d) and the lowest coefficient of determination ($R^2 = 0.51$) among the six evaluated methods. Further analysis found that the six abnormal points with predicted rainfall greater than 250 mm/d were all from the prediction results of Huashan station (SiteID: 57046) (Table 4). Specifically, on July 26, 1991, the measured rainfall was only 0.0 mm, while the predicted value of Co-kriging was as high as 308.74 mm, and the relative error was 308.74 mm. On July 29, 2014, the measured value was 0.0 mm, and the predicted value was 391.19 mm. The predicted values of the other four abnormal points were also between 250–400 mm/d, while the measured values did not exceed 64 mm/d. This serious overestimation phenomenon directly led to: the RMSE before and after 1990 and 2010 reached 4.08 mm/d, which was significantly higher than in other years; the coefficient of determination of the same period decreased to 0.34, forming a significant time series peak. The reason is that Huashan station is an elevation anomaly point in the study area, and its altitude (2064.9 m) is significantly higher than that of the surrounding stations (300–600 m), forming a significant topographic mutation.

Table 4 Co-kriging interpolation of predicted rainfall exceeding 250 mm (Station: Huashan)

Date	Observed rainfall (mm)	Predicted rainfall (mm)	Longitude (°)	Latitude (°)	Elevation (m)
1989/07/10	28.3	333.60	110.08	34.48	2064.9
1989/08/16	64.00	326.14	110.08	34.48	2064.9
1991/07/26	0.0	308.74	110.08	34.48	2064.9
2001/08/15	0.0	314.19	110.08	34.48	2064.9
2012/08/13	0.0	269.85	110.08	34.48	2064.9
2014/07/29	0.0	391.19	110.08	34.48	2064.9

This systematic overestimation is a manifestation of methodological limitation rather than an issue of input data quality. The pronounced elevation differences present a significant challenge for Co-kriging in modeling spatial correlation. Since Co-kriging method relies on elevation as a key covariate, and the elevation difference between Huashan and the surrounding stations exceeds 1000 m, it is difficult for the semivariogram to accurately describe this non-stationary spatial relationship. During heavy rainfall

events, the algorithm tends to overemphasize the correlation between elevation and rainfall, inappropriately extrapolate the high rainfall characteristics of Huashan to low altitude areas, resulting in a serious overestimation of the predicted values.

This phenomenon highlights the limitations of the Co-kriging method in extreme terrain areas, that is, when there is significant heterogeneity in the elevation distribution of sample points, the interpolation method based on the stationary hypothesis may produce bias. In contrast, methods such as TPS and IDW demonstrate greater stability under these conditions, as they do not explicitly incorporate elevation as a covariate.

The results of this study show that the Co-kriging method performs worst among the six interpolation methods, and its accuracy is even lower than that of the ordinary Kriging (OK) method. This finding is consistent with the research conclusion of Liu et al. (2021), that is, when using the gridded rainfall data set, the performance of the multivariate method CoK is not as good as the univariate method ordinary Kriging. The core assumption of the CoK method is that there is a strong correlation between rainfall and altitude although the study of Israelsson et al. (2020) shows that CoK can indeed improve the interpolation accuracy based on point data in dense rain gauge networks. However, the low correlation between rainfall and altitude is the key factor leading to the poor performance of CoK in this study. Many scholars (He et al. 2019; Herrera et al. 2019) have confirmed that when the correlation between rainfall and elevation is weak, the introduction of elevation covariates will reduce the interpolation accuracy. Especially in the Loess Plateau, complex topographic features may further weaken the spatial correlation between rainfall and elevation.

5.2 Adaptability of machine learning methods

The overall performance of the machine learning methods RF, SVM and GPR falls between traditional methods and geostatistical techniques. Among these, RF exhibits the best performance, with the lowest RMSE and highest R^2, followed by GPR, while SVM yielded the poorest performance. These findings are consistent with the findings of Chutsagulprom et al. (2022). Moreover, the study also highlights that in comparison to traditional interpolation methods, machine learning approaches such as ANN demonstrate relatively poor accuracy in rainfall estimation.

The advantage of machine learning methods lies in their ability to capture complex nonlinear relationships between rainfall and topographic variables. For instance, GPR can adaptively adjust the spatial correlation by automatically learning the covariance function, thereby eliminating the dependence of traditional geostatistical methods on semivariograms. However, this method exhibits a systematic underestimation issue in heavy rainfall events. The maximum rainfall measured on July 29, 2011 was 111.9 mm, and the predicted maximum values of RF, SVM and GPR were 70.57 mm, 55 mm and 97.63 mm, respectively. The underestimation is 13% −51%. This is because machine learning smoothes the real value of the variable to find the minimum variance, thereby promoting a more uniform surface. The smoothing effect causes the small value to be overestimated and the large value to be underestimated (Yamamoto 2005; De Almeida et al. 2025). The systematic underestimation of extreme rainfall by machine learning models (RF, SVM, GPR) stems from inherent learning mechanisms operating under imbalanced data distributions. Daily rainfall data is dominated by zero or low-intensity values, with extreme events representing rare occurrences in the distribution tail. This skewed structure induces two mechanistic biases: First, models optimized to minimize global loss functions such as Mean Squared Error prioritize prediction accuracy for the majority of low-to-medium intensity rainfall, as the optimization process penalizes errors on frequent small values more severely than those on rare extremes, pulling predictions toward the conditional mean. Second, the inherent smoothing effects of the algorithms themselves—such as prediction averaging in RF, smooth posterior means in GPR, and smooth outputs from SVM-RBF—act as implicit regularization that suppresses overfitting but also dampens responsiveness to true high-intensity signals (Yamamoto 2005; De Almeida et al. 2025). Consequently, to maintain stability across the overall dataset, the models sacrifice accuracy in the distribution tails, leading to significant underestimation of extreme events.

5.3 Temporal resolution effects in rainfall interpolation

The spatial interpolation of daily rainfall provide essential high-precision data for applications such as hydrological simulation, ecological restoration and disaster warning. However, the results of this study show that the correlation coefficient of daily rainfall interpolation is generally low, which is mainly due to the combined effect of two factors. From the perspective of time scale, the spatial variability of daily rainfall is significantly stronger than that of the monthly scale and the annual scale. Liu et al. (2020) demonstrated that the flow simulation is more sensitive to the rainfall spatial interpolation scheme on the short time scale, whereas differences in interpolation schemes at the monthly and annual scales have a relatively minor impact. This finding aligns with the results of Liao and Li (2024) who applied radial basis function network and BP neural network. Both show that the simulation accuracy of daily rainfall data is

significantly lower than that of monthly average data. From a regional perspective, the rainfall in the Loess Plateau shows an obvious decreasing trend from southeast to northwest. The spatial distribution of daily rainfall is strongly influenced by both topographic uplift and patterns of water vapour transport, resulting in pronounced regional heterogeneity (Zhao et al. 2018). This complex spatial heterogeneity makes the adaptability of different interpolation methods in local areas significantly different, thus further amplifying the overall error. Although in theory, daily rainfall interpolation offers higher temporal resolution and the potential for more accurate data, its practical application remains challenging due to the combined constraints of time-scale effects and pronounced regional variability. Thus, it is necessary to develop a new interpolation method that is more suitable for the spatial variability of daily rainfall patterns.

5.4 Implications of interpolation uncertainty for hydrological and erosion modeling

The error characteristics exhibited by different interpolation methods significantly impact hydrological and erosion modeling on the Loess Plateau. In terms of capturing extreme rainfall, systematic underestimation directly leads hydrological models to underestimate peak flow and total runoff volume. For the erosion-sensitive Loess Plateau, this subsequently causes soil erosion models to significantly underestimate erosion amounts from individual rainfall events, ultimately affecting the design standards and effectiveness evaluation of soil and water conservation measures (Yang et al. 2020). Conversely, the systematic overestimation by Co-kriging in steep terrain creates the opposite problem (Jin et al. 2016), potentially leading to overestimated runoff generation and erosion risk in mountainous watersheds, resulting in misallocation of disaster prevention resources. Spatial distribution errors—the "bull's eye" effect of IDW and the spatial autocorrelation of SVM errors—distort the true spatial pattern of rainfall input (Righi and Basso 2016). In distributed hydrological models, this disrupts the accuracy of runoff concentration path simulations and leads to misidentification of erosion hotspot areas.

Therefore, selecting interpolation methods that achieve a better balance between extreme value capture and spatial distribution fidelity is crucial for enhancing the reliability of hydrological and erosion modeling. The evaluation framework established in this study provides scientific basis for such selection.

Although this study takes the Loess Plateau as a case study, the methodology holds potential for cross-regional application. When applied to different regions in the future, it should be validated and adjusted according to local topographic and climatic characteristics to ensure its scientific validity and reliability (Amini et al. 2019). Complex terrain often leads to strong spatial heterogeneity in environmental factors such as climate, soil, and hydrology, thereby imposing higher demands on spatial interpolation methods (e.g., Kriging, IDW, etc.). Compared to flat and homogeneous areas, complex terrain provides a more rigorous testing ground for method comparison, better examining their robustness and applicability (Casellas et al. 2020; Lussana et al. 2018). Successful validation in such regions can offer more universal and valuable scientific references for other areas, including flat regions.

6 Conclusion

This study systematically evaluated the performance of six spatial interpolation methods for daily rainfall in the Loess Plateau. The results indicate that TPS and IDW exhibit the highest overall accuracy and stability. While TPS achieves the best interpolation effect, IDW produces comparable accuracy but introduces a pronounced “bull’s eye” pattern. The Co-kriging method performs adequately in flat areas but shows significant overestimation in regions with abrupt elevation changes. Overall, machine learning methods underperform relative to traditional approaches, with even the better-performing algorithms, RF and GPR, as well as TPS, systematically underestimating extreme rainfall events.

Interpolation accuracy exhibits pronounced spatial and temporal heterogeneity. Temporally, errors peak during the summer months (July–August), yet traditional methods maintain good stability during the rainy season. Spatially, the highest accuracy is observed in sandy land and agricultural irrigation areas (C area), whereas the loess plateau gully region (A1 area) demonstrates the lowest accuracy. TPS and IDW consistently provide superior performance across most ecological zones (A1, A2, B1, C, D), with TPS being particularly suitable for the B2 sub-area of the loess hilly and gully region. In contrast, Co-kriging is limited in applicability in rocky mountain and valley plain regions (D area) due to overestimation in areas with strong elevation gradients.

These findings provide a robust basis for selecting appropriate rainfall spatial interpolation methods under the unique complex terrain conditions of the Loess Plateau, contributing to improved hydrological simulations and informed ecological management.

Appendix 1

Table 5 Sensitivity analysis of IDW parameters: performance metrics (CC, MAE, RMSE, R^2) under different combinations of power (P) and maximum neighbors (Nmax)

	P = 1				*P* = 2				*P* = 3			
	Nmax = 5	Nmax = 10	Nmax = 15	Nmax = 20	Nmax = 5	Nmax = 10	Nmax = 15	Nmax = 20	Nmax = 5	Nmax = 10	Nmax = 15	Nmax = 20
CC	**0.846*****	**0.846*****	0.840***	0.834***	0.839***	0.843***	0.843***	0.843***	0.829***	0.833***	0.833***	0.834***
MAE (mm/d)	**0.678**	0.702	0.731	0.758	0.679	0.683	0.693	0.703	0.690	0.689	0.691	0.693
RMSE (mm/d)	**2.751**	2.757	2.807	2.856	2.808	2.769	2.771	2.779	2.892	2.857	2.849	2.847
R^2	**0.715**	0.713	0.703	0.693	0.703	0.711	0.711	0.709	0.685	0.692	0.694	0.694

The symbol *** indicates that the correlation between observations and predictions is statistically significant at the 0.1% significance level ($p < 0.001$). Optimal values for each metric are highlighted in bold

Table 6 Six spatial interpolation schemes for four statistical indicators (CC, R^2, RMSE and MAE)

Variables	TPS				IDW				kriging				RF				SVM				GPR			
	CC	MAE	RMSE	R^2	CC	MAE	RMSE	R^2	CC	MAE	RMSE	R^2	CC	MAE	RMSE	R2	CC	MAE	RMSE	R^2	CC	MAE	RMSE	R^2
Lon, lat	0.84***	0.70	2.76	0.71	0.85***	0.68	2.75	0.71	0.84***	0.75	2.92	0.71	0.83***	0.74	2.87	0.69	0.79***	0.76	3.20	0.61	0.83***	0.75	2.86	0.69
Lon, lat, elevation	0.84***	0.73	2.84	0.70	-	-	-	-	0.75***	0.80	3.58	0.52	0.82***	0.79	2.97	0.67	0.79***	0.77	3.24	0.61	0.81***	0.80	3.00	0.66
Lon, lat, slope	0.81***	0.79	2.99	0.81	-	-	-	-	0.33***	1.25	13.80	-6.18	0.80	0.87	3.13	0.63	0.76***	0.82	3.36	0.57	0.79***	0.86	3.16	0.62
Lon, lat, aspect	0.78***	0.89	3.20	0.78	-	-	-	-	0.79***	0.80	3.25	0.60	0.79	0.89	3.17	0.62	0.77***	0.82	3.36	0.57	0.78***	0.91	3.24	0.60
Lon, lat, elevation, slope	0.77***	0.87	3.30	0.59	-	-	-	-	0.32***	1.47	15.18	-7.68	0.81***	0.81	3.01	0.66	0.77***	0.82	3.34	0.58	0.79***	0.87	3.17	0.62
Lon, lat, elevation, aspect	0.76***	0.93	3.33	0.76	-	-	-	-	0.69***	0.97	4.23	0.33	0.81	0.82	3.02	0.66	0.76***	0.83	3.38	0.57	0.77***	0.93	3.30	0.59
Lon, lat, slope, aspect	0.74***	0.98	3.48	0.74	-	-	-	-	0.33***	1.46	14.17	-6.57	0.80	0.87	3.12	0.63	0.74***	0.89	3.50	0.54	0.75***	0.99	3.44	0.55
Lon, lat, elevation, slope, aspect	0.74***	0.98	3.47	0.55	-	-	-	-	0.29***	1.88	17.64	-10.73	0.80***	0.87	3.11	0.63	0.74***	0.88	3.46	0.55	0.75***	0.97	3.41	0.56

The symbol *** indicates that the correlation between observations and predictions is statistically significant at the 0.1% significance level ($p < 0.001$). Lon, lat, elevation, slope, aspect represent longitude, latitude, elevation, slope, aspect respectively

Acknowledgements We acknowledge the National Meteorological Science Data Center of China for providing meteorological data and the Resource and Environment Data Center of Chinese Academy of Sciences for supplying DEM data. We also thank Prof. Yanfen Yang's team (Northwest A&F University) as well as the National Earth System Science Data Center for providing the ecological zoning data of the Loess Plateau.

Author contributions L.J.: Data curation, Formal analysis, Methodology, Visualization, Writing—original draft. Q.Z.: Conceptualization, Formal analysis, Methodology, Supervision, Visualization, Writing – review & editing. X.Y.: Conceptualization, Writing – review & editing. G.W.: Conceptualization, Writing – review & editing. B.D.: Conceptualization, Writing – review & editing. J.T.: Writing – review & editing. Q.Y.: Data curation, Supervision.

Data availability The datasets generated and/or analysed during the current study are available from the following public repositories: • Daily meteorological data: National Meteorological Science Data Center https://data.cma.cn/ • SRTM DEM data: Resource and Enviro nment Data Cloud Platform, Chinese Academy of Sciences http://ww w.resdc.cn/ • Ecological zoning data: National Earth System Science Data Center, China http://www.geodata.cn.

Declarations

Competing interest The authors declare no competing interests.

References

Achite M, Tsangaratos P, Pellicone G, Mohammadi B, Caloiero T (2024) Application of multiple spatial interpolation approaches to annual rainfall data in the Wadi Cheliff basin (north Algeria). Ain Shams Eng J 15(3):102578. https://doi.org/10.1016/j.asej.2023.102578

Amini MA, Torkan G, Eslamian S, Zareian MJ, Adamowski JF (2019) Analysis of deterministic and geostatistical interpolation techniques for mapping meteorological variables at large watershed scales. Acta Geophys 67(1):191–203. https://doi.org/10.1007/s11600-018-0226-y

Asokan D, Ramalingam A, Nadeem T, Mall A, Pardeshi G (2025) Spatio-temporal patterns and spatial autocorrelation of chandipura virus outbreaks in India (2009–2023): a one health perspective. Microbe 9:100567. https://doi.org/10.1016/j.microb.2025.100567

Bajat B, Pejović M, Luković J, Manojlović P, Ducić V, Mustafić S (2013) Mapping average annual precipitation in Serbia (1961–1990) by using regression kriging. Theor Appl Climatol 112:1–13. https://doi.org/10.1007/s00704-012-0702-2

Breiman L (2001) Statistical modeling: the two cultures (with comments and a rejoinder by the author). Stat Sci 16(3):199–231. https://doi.org/10.1214/ss/1009213726

Camera C, Bruggeman A, Hadjinicolaou P, Pashiardis S, Lange MA (2014) Evaluation of interpolation techniques for the creation of gridded daily precipitation (1×1 km2); Cyprus, 1980–2010. J Geophys Res Atmos 119(2):693–712. https://doi.org/10.1002/2013JD020611

Casellas E, Bech J, Veciana R, Ramon Miro J, Sairouni A, Pineda N (2020) A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. Atmos Res 246:105103. https://doi.org/10.1016/j.atmosres.2020.105103

Chang FJ, Guo S (2020) Advances in hydrologic forecasts and water resources management. Water 12(6):1819. https://doi.org/10.3390/w12061819

Chen C, Hao J, Yang S, Li Y (2025) Blending daily satellite precipitation product and rain gauges using stacking ensemble machine learning with the consideration of spatial heterogeneity. J Hydrol 658:133223. https://doi.org/10.1016/j.jhydrol.2025.133223

Chilès JP, Delfiner P (2012) Geostatistics: Modeling spatial uncertainty, 2nd edn. Wiley. https://doi.org/10.1002/9781118136188

Chutsagulprom N, Chaisee K, Wongsaijai B, Ditthakit P, Pinthong S, Salaeh N, Adunyanon P (2022) Spatial interpolation methods for estimating monthly rainfall distribution in Thailand. Theor Appl Climatol 148:317–328. https://doi.org/10.1007/s00704-022-03927-7

Cutler A, Cutler DR, Stevens JR (2012) Random forests. In: Zhang C, Ma Y (eds) Ensemble machine learning: methods and applications. Springer, New York, pp 157–175. https://doi.org/10.1007/978-1-4419-9326-7_5

Das M, Hazra A, Sarkar A, Bhattacharya S, Banik P (2017) Comparison of spatial interpolation methods for estimation of weekly rainfall in West Bengal, India. Mausam 68(1):41–50. https://doi.org/10.54302/mausam.v68i1.407

De Almeida LT, Cecílio RA, Pruski FF (2025) Method to establish intense rainfall equations based in geoprocessing. Environ Model Assess 30:141–155. https://doi.org/10.1007/s10666-024-09999-1

Deutsch CV, Journel AG (1997) GSLIB: Geostatistical software library and user's guide, 2nd edn. Oxford University Press, New York

Drucker H, Burges CJC, Kaufman L, Smola A, Vapnik V (1996) Support vector regression machines. Adv Neural Inf Process Syst 9:155–161. https://doi.org/10.5555/2998981.2999003

Eugenio FC, Peluzio TMO, Pereira AAB, Santos AR, Peluzio JBE, Bragança R, Fiedler NC, de Paula ENSO (2014) Zoning agroclimatological *Coffea canephora* for Espírito Santo by spatial interpolation. Coffee Sci 9(3):319–328. https://doi.org/10.25186/cs.v9i3.649

Fagandini C, Todaro V, Tanda MG, Pavan G, Zanini A, Secci D, Montisci A (2024) Missing rainfall daily data: a comparison among gap-filling approaches. Math Geosci 56:191–217. https://doi.org/10.1007/s11004-023-10113-6

Feng ZQ, Tan ML, Juneng L, Tye MR, Xia LL, Zhang F (2025) Effects of solar radiation modification on precipitation extremes in Southeast Asia: insights from the GeoMIP G6 experiments. Adv Clim Change Res 16:591–605. https://doi.org/10.1016/j.accre.2025.04.009

Fung KF, Chew KS, Huang YF, Ahmed AN, Teo FY, Ng JL, Elshafie A (2022) Evaluation of spatial interpolation methods and spatio-temporal modeling of rainfall distribution in Peninsular Malaysia. Ain Shams Eng J 13(2):101571. https://doi.org/10.1016/j.asej.2021.09.001

Guidoum A (2025) Statistical modeling and mapping of rainfall in the endorheic basins of Northern Algeria: a comparison of spatial interpolation methods. Acta Geophys 73:1679–1699. https://doi.org/10.1007/s11600-024-01392-6

Han JT, Hou Y, Wang ZR, Yang YT (2023) Spatial interpolation method of precipitation based on terrain revising and application. Yangtze River 54(7):75–80

Herrera S, Kotlarski S, Soares PMM, Cardoso RM, Jaczewski A, Gutiérrez JM, Maraun D (2019) Uncertainty in gridded precipitation products: Influence of station density, interpolation method and grid resolution. Int J Climatol 39(9):3717–3729. https://doi.org/10.1002/joc.5878

He S, Yang J, Bao Q, Wang L, Wang B (2019) Fidelity of the observational/reanalysis datasets and global climate models in representation of extreme precipitation in East China. J Clim 32(1):195–212. https://doi.org/10.1175/JCLI-D-18-0104.1

Hiemstra PH, Pebesma EJ, Twenhöfel CJW, Heuvelink GBM (2009) Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. Comput Geosci 35(8):1711–1721. https://doi.org/10.1016/j.cageo.2008.10.011

Hou CL, Tan ML, Li LH, Zhang F (2024) Comparison of three machine learning algorithms for retrieving soil moisture information from Sentinel-1A SAR data in northwest Shandong plain, China. Adv Space Res 74:75–88. https://doi.org/10.1016/j.asr.2024.03.047

Hou CL, Tan ML, Yaseen ZM, Zhang F (2025) Advancing soil moisture retrieval from SAR data using meta-heuristic optimization xgboost and SHAP analysis. Int J Remote Sens 46:5354–5383. https://doi.org/10.1080/01431161.2025.2518507

Hutchinson MF (1995) Interpolating mean rainfall using thin plate smoothing splines. Int J Geogr Inf Syst 9(4):385–403. https://doi.org/10.1080/02693799508902045

Hutchinson MF (1998) Interpolation of rainfall data with thin plate smoothing splines. Part I: two dimensional smoothing of data with short range correlation. J Geogr Inf Decis Anal 2(2):139–151. https://doi.org/10.1080/10824009809480509

İlker A, Terzi Ö, Şener E (2019) Comparing of interpolation methods in mapping spatial distributions of rainfall: case study of Mediterranean Region. Teknik Dergi 30(3):9213–9219. https://doi.org/10.18400/tekderg.334186

Israelsson J, Black E, Neves C, Torgbor FF, Greatrex H, Tanu M, Lamptey PNL (2020) The spatial correlation structure of rainfall at the local scale over southern Ghana. Journal of Hydrology: Regional Studies 31:100720. https://doi.org/10.1016/j.ejrh.2020.100720

Jin Q, Zhang J, Shi M, Huang J (2016) Estimating loess plateau average annual precipitation with multiple linear regression kriging and geographically weighted regression kriging. Water 8(6):266. https://doi.org/10.3390/w8060266

Keriven N, Garreau D, Poli I (2018) Testing for differences in Gaussian process regression. In: 2018 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2661–2665. https://doi.org/10.1109/ICASSP.2018.8461792

Knotters M, Brus D, Oude Voshaar J (1995) A comparison of kriging, cokriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. Geoderma 67(3):227–246. https://doi.org/10.1016/0016-7061(95)00011-C

Li J, Heap AD, Potter A, Daniell JJ (2011) Application of machine learning methods to spatial interpolation of environmental variables. Environ Model Softw 26(12):1647–1659. https://doi.org/10.1016/j.envsoft.2011.07.004

Liao Z, Li M (2024) Comparing rainfall prediction at various time scales and rainfall interpolation at the regional scale using artificial neural networks. Theor Appl Climatol 155:9929–9940. https://doi.org/10.1007/s00704-024-05205-0

Liu ZH, Li LT, McVicar TR (2008) Introduction of the professional interpolation software for meteorology data: ANUSPLIN. Meteorol Mon 34(2):92–100. https://doi.org/10.7519/j.issn.1000-0526.2008.02.013

Liu ZJ, Yu XX, Wang SS (2012) Comparative analysis of three covariates methods in thin-plate smoothing splines for interpolating precipitation. Prog Geogr 31(1):56–62. https://doi.org/10.11820/dlkxjz.2012.01.008

Liu D, Zhao Q, Fu D, Guo S, Liu P, Zeng Y (2020) Comparison of spatial interpolation methods for the estimation of precipitation patterns at different time scales to improve the accuracy of discharge simulations. Hydrol Res 51(4):583–601. https://doi.org/10.2166/nh.2020.146

Liu Y, Liu Z, Wang W, Zhao L, Li S, Yang X, Wang L, Zhu Y (2021) An assessment of statistical interpolation methods suited for gridded rainfall datasets. Int J Climatol 42(5):2754–2772. https://doi.org/10.1002/joc.7389

Lopes JRRFM (2013) Methodologies for the assessment of climate change impacts on irrigated agriculture and irrigation systems. Dissertation, Universidade Técnica De Lisboa

Lussana C, Saloranta T, Skaugen T, Magnusson J, Tveito OE, Andersen J (2018) Senorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day. Earth Syst Sci Data 10(1):235–249. https://doi.org/10.5194/essd-10-235-2018

Lyra GB, Correia TP, de Oliveira-Júnior JF (2018) Evaluation of methods of spatial interpolation for monthly rainfall data over the state of Rio de Janeiro, Brazil. Theor Appl Climatol 134:955–965. https://doi.org/10.1007/s00704-017-2322-3

Meng QX, Liu GB, Yang QK (2006) Research on spatial interpolation methods of precipitation on Loess Plateau. J Northwest Sci-Tech Univ Agric For (Nat Sci Ed) 34(3):83–88

Nobrega AEL, Barroca Filho IM (2025) Analysis of machine learning performance in spatial interpolation of rainfall data. IEEE Access 13:84727–84737. https://doi.org/10.1109/ACCESS.2025.3569261

Park J, Park JH, Choi JS, Joo JC, Park K, Yoon HC, Park CY, Lee WH, Heo TY (2020) Ensemble model development for the prediction of a disaster index in water treatment systems. Water 12(11):3195. https://doi.org/10.3390/w12113195

Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. Comput Geosci 30(7):683–691. https://doi.org/10.1016/j.cageo.2004.03.012

Pellicone G, Caloiero T, Modica G, Guagliardi I (2018) Application of several spatial interpolation techniques to monthly rainfall data in the Calabria region (southern Italy). Int J Climatol 38(9):3651–3666. https://doi.org/10.1002/joc.5525

Pinthong S, Ditthakit P, Salaeh N, Phetprayoon T, Wongsaijai B, Putthividhya A, Jirakittayakorn K, Torsri K (2024) Imputation of missing monthly rainfall data using machine learning and spatial interpolation approaches in Thale Sap Songkhla River Basin, Thailand. Environ Sci Pollut Res 31:54044–54060. https://doi.org/10.1007/s11356-022-23022-8

Plouffe CCF, Robertson C, Chandrapala L (2015) Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: a case study of Sri Lanka. Environ Model Softw 67:57–71. https://doi.org/10.1016/j.envsoft.2015.01.011

Raghavendra NS, Deka PC (2014) Support vector machine applications in the field of hydrology: a review. Appl Soft Comput 19:372–386. https://doi.org/10.1016/j.asoc.2014.02.002

Righi E, Basso LA (2016) Application and analysis of rainfall spatial interpolation techniques. Ambiência 12(1):45–62. https://doi.org/10.5935/ambiencia.2016.01.06

Rodríguez-Carrillo JA, González-Trinidad J, Silva-Avalos RU, Júnez-Ferreira HE, Moreno-Chavez G, Robles-Rovelo CO, Contreras-Rodríguez AR (2025) Spatial rainfall estimation applying machine learning techniques in a semi-arid basin. J Hydroinform 27(1):69–87. https://doi.org/10.2166/hydro.2024.253

Ruiz-Ortiz V, Fernandez HM, Granja-Martins FM, Vélez-Nicolás M, Isidoro JMGP, García-López S (2024) A multimethod interpolation approach for mapping the spatial distribution of rainfall in southwest Iberian Peninsula. Hydrol Sci J 69(13):1736–1749. https://doi.org/10.1080/02626667.2024.2387805

Salcedo-Sanz S, Rojo-Álvarez JL, Martínez-Ramón M, Camps-Valls G (2016) Support vector machines in engineering: an overview. Wiley Interdiscip Rev Data Min Knowl Discov 4(3):234–267. https://doi.org/10.1002/widm.1125

Sattari MT, Falsafian K, Irvem A, Qasem SN (2020) Potential of kernel and tree-based machine-learning models for estimating missing data of rainfall. Eng Appl Comput Fluid Mech 14(1):1078–1094. https://doi.org/10.1080/19942060.2020.1803971

Stalenberg E, Hutchinson MF, Foley WJ (2018) Using historical normals to improve modern monthly climate normal surfaces for Madagascar. Int J Climatol 38(15):5746–5765. https://doi.org/10.1002/joc.5776

Wang H, Pan Y, Chen Z (2017) Spatial and temporal patterns of temperature and precipitation extremes in the Huaihe River Basin, China during 1960–2014. Sci Geogr Sin 37(12):1900–1908. https://doi.org/10.13249/j.cnki.sgs.2017.12.014

Wu W, You Q, Wang D (2016) Characteristics of extreme precipitation in China based on homogenized precipitation data. J Nat Resour 31(6):1015–1026

Yamamoto JK (2005) Correcting the smoothing effect of ordinary kriging estimates. Math Geol 37:69–94. https://doi.org/10.1007/s11004-005-8748-7

Yang M (2015) Evaluation of rainfall interpolation methods in the Netherlands. Master's thesis, Chang'an University

Yang X, Zhang X, Lv D, Yin S, Zhang M, Zhu Q, Yu Q, Liu B (2020) Remote sensing estimation of the soil erosion cover-management factor for China's Loess Plateau. Land Degrad Dev 31(15):1942–1955. https://doi.org/10.1002/ldr.3577

Yu Y, Wei W, Chen LD, Yang L, Zhang HD (2015) Comparison on the methods for spatial interpolation of the annual average precipitation in the Loess Plateau region. Chin J Appl Ecol 26(4):999–1006

Zagorecki AT, Johnson DE, Ristvej J (2013) Data mining and machine learning in the context of disaster and crisis management. Int J Emerg Manag 9(4):351–365. https://doi.org/10.1504/ijem.2013.059879

Zhang F, O'Donnell LJ (2020) Support vector regression. In: Zhou ZH (ed) Machine Learning. Academic Press, pp 123–140. https://doi.org/10.1016/B978-0-12-815739-8.00007-9

Zhang YK (2024) Spatial interpolation of precipitation in Tangshan City based on support vector machine regression algorithm. Jilin Water Resour 02:23–25+78. https://doi.org/10.15940/j.cnki.22-1179/tv.2024.02.006

Zhang HL, Liu JL (2025) Jointly stochastic fully symmetric interpolatory rules and local approximation for scalable Gaussian process regression. Pattern Recognition 159:111125. https://doi.org/10.1016/j.patcog.2024.111125

Zhao T, Yang XY (2012) Research on spatial interpolation methods of annual average precipitation on Loess Plateau. Ground Water 34(2):189–191

Zhao G, Zhai J, Tian P, Liu X, Franz K, Xia J, Wang L (2018) Variations in extreme precipitation on the Loess Plateau using a high-resolution dataset and their linkages with atmospheric circulation indices. Theor Appl Climatol 133:1235–1247. https://doi.org/10.1007/s00704-017-2251-1

Zhu H, Jia S (2004) Uncertainty in the spatial interpolation of rainfall data. Prog Geogr 23(2):35–42. https://doi.org/10.11820/dlkxjz.2004.02.005

Zhu Q, Yang X, Ji F, Liu DL, Yu Q (2020) Extreme rainfall, rainfall erosivity, and hillslope erosion in Australian Alpine region and their future changes. Int J Climatol 40(2):1213–1227. https://doi.org/10.1002/joc.6266

Zhu Q, Yang X, Ji F, Du Z (2025) Rainfall erosivity projection in South-East Australia using the improved regional climate simulations. Int J Climatol 45(2):e8702. https://doi.org/10.1002/joc.8702