FRONTIERS IN
ENVIRONMENTAL RESEARCH

Frontiers in Environmental Research

Table of Content

SPATIOTEMPORAL VARIABILITY OF YELLOW RIVER WATER-SEDIMENT FLUXES: A HYBRID APPROACH USING CUBIC SPLINE INTERPOLATION AND MANN-KENDALL TEST

Yun Xiao

College of Humanities and Education, Shaanxi Energy Institute, Xianyang 712000, Shaanxi, China.
Corresponding Email: xiaoyun6818@163.com

Abstract: Studying the variation law of water and sediment fluxes in the Yellow River is of great significance for environmental protection, adaptation to climate change and improvement of the quality of life of residents in the Yellow River Basin. this paper established a cubic spline interpolation model to supplement the missing sediment content data. By using the known data and interpolation data for plotting, it was found that the interpolation effect was good. Subsequently, the total drainage volume and total sediment discharge volume from 2016 to 2021 were calculated respectively using the interpolation data. In order to further study the variability of water and sediment content, this paper adopts the Mann-Kendall non-parametric test method to analyze the variability of water and sediment flux, revealing its inherent law. Through specific data analysis, the important role of sand regulation and water control in ensuring the ecological health of the Yellow River has been further demonstrated.

Keywords: Cubic spline interpolation; Mann-Kendall non-parametric test method; Residual analysis; MATLAB

1 INTRODUCTION

The Yellow River plays a pivotal role in shaping the ecological and socioeconomic fabric of its basin. Investigating the spatiotemporal variability of water-sediment flux in the Yellow River holds critical theoretical significance for addressing environmental governance, climate change adaptation, and livelihood improvements across the basin. Furthermore, it provides essential insights into optimizing water resource allocation, coordinating human-land relationships, regulating water-sediment discharge, and enhancing flood control and disaster mitigation strategies. As one of China's most vital river systems, the Yellow River's hydrological and sediment dynamics exert profound impacts on ecosystem integrity, flood resilience, and sustainable socioeconomic development. Characterized by exceptionally high sediment concentrations, the Yellow River presents unparalleled challenges in water-sediment management, necessitating highly accurate, real-time, and granular monitoring systems. Advanced analysis of water-sediment dynamics enables the identification of temporal patterns, prediction of future flux trends, and formulation of science-driven policies for ecological conservation, flood risk reduction, and adaptive river basin management.

Firstly, the complexity of the Yellow River water-sediment regulation system and its impact on basin sustainable development have become focal points of research. Cao et al. analyzed the influencing factors and evolution trends of the Yellow River water-sediment regulation system from a systemic perspective, identifying climate change and human activities as the key drivers of abrupt changes in water-sediment relationships [1]. Current models still have limitations in their response mechanisms to extreme hydrological events, particularly in terms of prediction accuracy under high sediment load scenarios. Secondly, addressing the issue of missing monitoring data, cubic spline interpolation has emerged as a mainstream method due to its smoothness and local adaptability. Habermann and Kindermann systematically expounded the mathematical principles of multidimensional spline interpolation [2], demonstrating its superiority in reconstructing non-uniform data; Abdulmohsin et al proposed a classification method based on cubic spline interpolation, verifying its robustness in high-noise environments [3]; He and Li combined it with support vector quantile regression for uncertainty analysis in wind power probability density prediction, achieving a 15% reduction in error [4]. Liu et al. proposed a Transformer-based model with missing position encoding, leveraging self-attention mechanisms to capture long-term dependencies in multivariate hydrological sequences, achieving a 15%–25% reduction in imputation errors [5]. In Yellow River water-sediment studies, cubic spline interpolation has been applied to complete sediment concentration data, but its accuracy in edge regions is constrained by data sparsity. Future research could explore adaptive weighted interpolation or multi-scale fusion algorithms to enhance reliability. Thirdly, non-parametric statistical methods hold significant importance in hydrological trend analysis, with the Mann-Kendall (MK) test being the most widely applied. Mann and Kendall established the theoretical foundation of the MK test [6], whose characteristic of not requiring assumptions about data distribution makes it suitable for non-normal hydrological sequences; Patle et al. and Cabral Júnior et al. respectively applied the MK test to precipitation trend analysis in India and Brazil, finding it maintains high sensitivity even with small sample sizes [7]; Sang et al. compared the MK test with empirical mode decomposition (EMD), noting the MK test's greater advantage in detecting abrupt changes but its limitation in distinguishing natural fluctuations from human interventions [8].

Through reading literature, it has been found that there is limited research on the use of cubic spline interpolation in the study of water and sediment fluxes in the Yellow River. Therefore, this article proposes using cubic spline interpolation to study the water and sediment flux of the Yellow River.

2 MODEL

2.1 Cubic Spline Interpolation

The cubic spline interpolation is a method that uses a series of cubic polynomials assembled according to certain smoothness requirements to approximate a function or a set of data points.
Divide the data range into multiple sub-ranges and define a cubic polynomial for each sub-range.

$$S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \tag{1}$$

where $i = 1, 2, 3 \ldots n-1$; where a_i, b_i, c_i, d_i are the coefficients of cubic spline interpolation.

Constrains adjacent polynomials to be continuous in function values and first and second derivatives at the junction points.
Combined with boundary conditions (such as natural splines) to solve the coefficient, the global smooth curve is formed.
The data utilized in this article is sourced from the website www.mcm.edu.cn..
The relationship between sediment concentration and time, water level and discharge in a hydrology station from 2016 to 2021 was studied. Supplement the missing annual sediment concentration data according to the known data.
The steps to supplement the sediment content data are as follows (take 2016 as an example) :
Step 1: Data selection
Looking at six years of monitoring data, it was found that when filtering the data with the keyword "sediment concentration", there were 373 sampling events in 2016, of which 366 occurred at 8 am, accounting for a certain proportion. The statistical results of other years are shown in Table 1.

Table 1 Statistical Table of Sediment Concentration at 8:00 from 2016 to 2021

Year	Total sampling times of sediment content	8:00 Sampling times of sediment content	ratio
2016	373	366	98.1%
2017	372	358	96.2%
2018	432	364	84.2%
2019	404	364	90%
2020	321	274	85.3%
2021	257	227	88.3%

As can be seen from Table 1, in the past six years, the data of sediment content sampled by the hydrological station at 8:00 am every day accounted for 85% of the total data. Therefore, in order to fill in the missing sediment concentration data, the annual 8 a.m. sampling data was used.
Step 2: Time processing
Insert a serial number column in the 2016 data , each serial number corresponding to a specific date and time. For example, serial number 1 corresponds to 0:00 a.m. on January 1, 2016; Serial number 2380 corresponds to 24pm on December 31, 2016, and other serial numbers are derived from this. Replace the time with a serial number, using the serial number as the horizontal coordinate.
Step 3: Cubic spline interpolation
The filter data were interpolated with three splines to calculate the sediment concentration data in 2016. Since the sampling frequency at 8 am in 2016 accounted for a certain proportion, the sediment concentration data at 8 am in 2016 was taken as the ordinate and the serial number as the horizontal coordinate. Three spline interpolation was carried out with MATLAB to calculate the sediment concentration data of 2016.

2.2 Mann-Kendall non-Parametric Test Method

Mann-Kendall (MK) test is a non-parametric statistical method used to detect trend changes or abrupt points in time series data [6].
Its primary advantage is that it does not necessitate the assumption of a specific data distribution (e.g., normal distribution) and exhibits robustness against outliers. Consequently, it has been extensively applied in disciplines such as hydrology, meteorology, and environmental science.
The working principle of MK test is to judge the trend direction by comparing the relative size of each pair of observations in the data sequence. The equation is given by:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sgn}(x_j - x_i) \tag{2}$$

where S is the result of the sum of the counts of $x_j - x_i$; $x_1, x_2, x_3, \cdots x_n$ are time series; n is the number of data in the time series; Each data pair is assigned the following values:

$$\mathrm{sgn}(x_j - x_i) = \begin{cases} 1 & if x_j > x_i \\ 0 & if x_j = x_i \\ -1 & if x_j < x_i \end{cases} \tag{3}$$

When there are duplicate values in the data, the variance correction is calculated according to equation 4:

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{k=1}^{m} t_k(t_k-1)(2t_k+5)}{18} \tag{4}$$

where t_k is the number of data with equal values in a certain group; m is the number of groups containing equal values in the data series in a group k.

The Mann-Kendall test statistic is based on the value of the variable Z, calculated according to equation 5:

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{Var(S)}} & if \quad S > 0 \\ 0 & if \quad S = 0 \\ \dfrac{S+1}{\sqrt{Var(S)}} & if \quad S < 0 \end{cases} \tag{5}$$

if $|Z| > Z_{1-\frac{\alpha}{2}}$, a level of significance (α) of 5% was considered, that reject the original assumption (no trend), and consider that there is a significant upward or downward trend [7-8].

The Mann-Kendall (MK) test identifies abrupt change points in time series data by analyzing the cumulative trend statistics of both forward and reverse sequences. The specific steps are as follows:

Step 1: Forward Sequence U_F

For a time series $\{x_1, x_2, x_3, \cdots x_n\}$, the standardized statistic $U_F(k)$ at position $k (k = 2,3,\cdots,n)$ is calculated as:

$$U_F(k) = \frac{S_k - E(S_k)}{\sqrt{Var(S_k)}} \tag{6}$$

where $S_k = \sum_{i=1}^{k} \sum_{j=1}^{i-1} \mathrm{sgn}(x_i - x_j)$ is the cumulative sum of pairwise comparisons up to the k-th data point; $E(S_k) = \frac{k(k-1)}{4}$ is the expected value under the null hypothesis (no trend); $Var(S_k) = \frac{k(k-1)(2k+5)}{72}$ is the variance of S_k.

Step 2: Reverse Sequence U_B

For the reversed time series $\{x_n, x_{n-1}, x_{n-2}, \cdots x_1\}$, the standardized statistic $U_B(k)$ at position $k (k = n-1, n-2, \cdots, 1)$ is defined as:

$$U_B(k) = \frac{S_k' - E(S_k')}{\sqrt{Var(S_k')}} \tag{7}$$

where $S_k' = \sum_{i=k}^{n} \sum_{j=i+1}^{n} \mathrm{sgn}(x_j - x_i)$, with $E(S_k')$ and $Var(S_k')$ computed similarly to $U_F(k)$.

Step 3: Change Point Identification

Plot $U_F(k)$ and $U_B(k)$ against time indices. Determine the intersection points of $U_F(k)$ and $U_B(k)$ within the confidence interval (e.g., $\alpha = 0.05$, critical value $Z_{1-\frac{\alpha}{2}} \approx 1.96$). Significant abrupt changes are detected at time points where $U_F(k)$ and $U_B(k)$ cross each other within the confidence bounds [9-11].

3 RESULTS

3.1 Cubic Spline Interpolation

Cubic spline interpolation were interpolated respectively for the data screened from 2016 to 2021, and the sediment content data from 2016 to 2021 were calculated using MATLAB software (see Figure 1 to Figure 6).
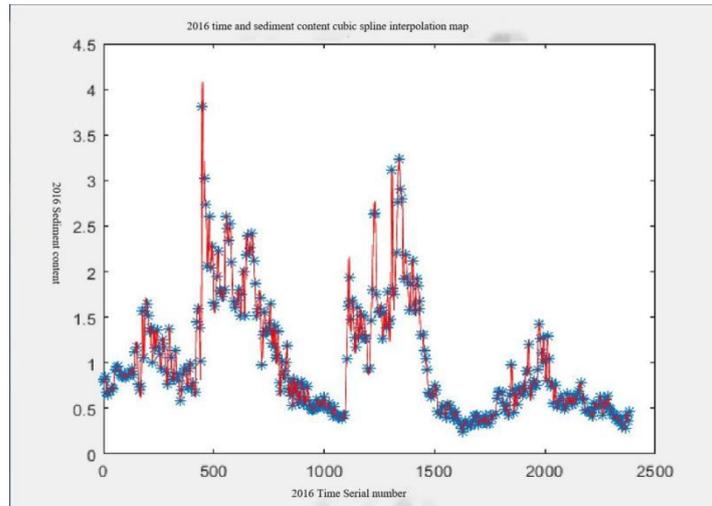
Figure 1 Cubic Spline Interpolation of Time and Sediment Concentration in 2016

The blue dots in Figure 1 are 366 selected ones, and the red lines are cubic spline interpolation curves. It is found from the images and errors that the interpolation effect is good.
The same can be obtained: 2017-2021 time and sediment content cubic spline interpolation plot and all missing sediment content data.
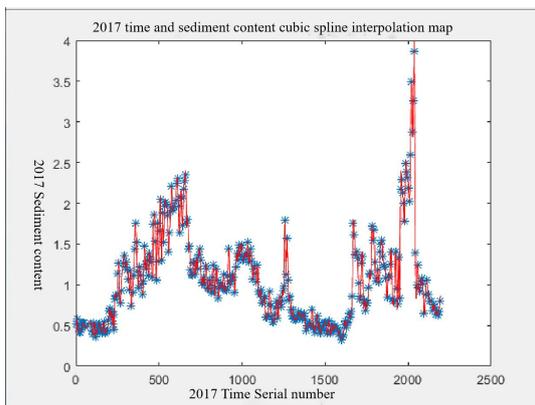
Figure 2 Cubic Spline Interpolation of Time and Sediment Concentration in 2017
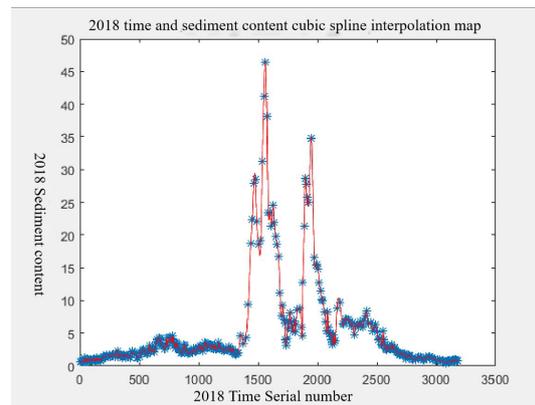
Figure 3 Cubic Spline Interpolation of Time and Sediment Concentration in 2018

The blue dots in Figure 2 are 358 selected ones, and the red lines are cubic spline interpolation curves. It is found from the images and errors that the interpolation effect is good.
The blue dots in Figure 3 are 364 selected ones, and the red lines are cubic spline interpolation curves. It is found from the images and errors that the interpolation effect is good.
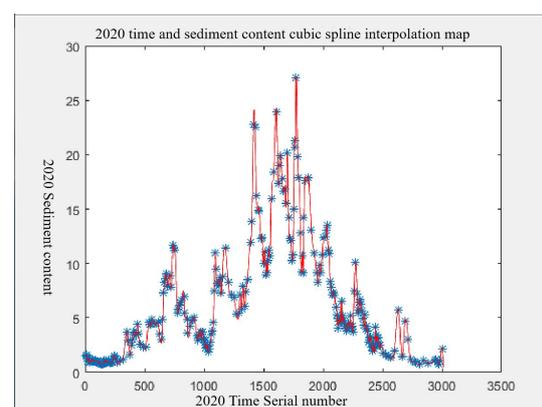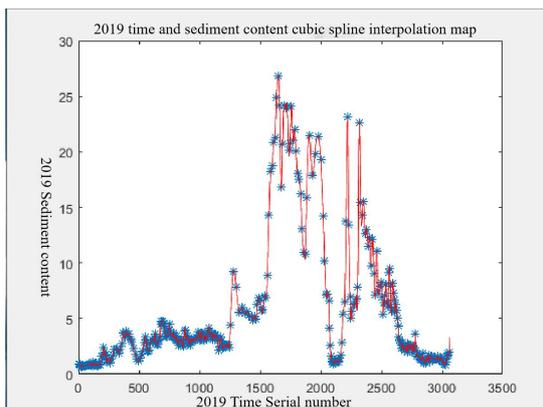
Figure5 Cubic Spline Interpolation of
Time and Sediment Concentration in 2020

The blue dots in Figure 4 are 364 selected ones, and the red lines are cubic spline interpolation curves. It is found from the images and errors that the interpolation effect is good.

The blue dots in Figure 5 are 274 selected ones, and the red lines are cubic spline interpolation curves. It is found from the images and errors that the interpolation effect is good.
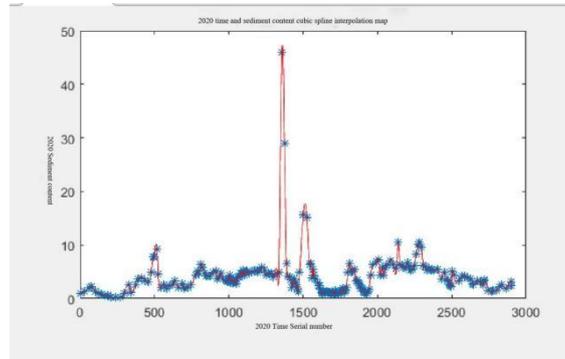
Figure 6 Cubic Spline Interpolation of Time and Sediment Concentration in 2021

The blue dots in Figure 6 are 227 selected ones, and the red lines are cubic spline interpolation curves. It is found from the images and errors that the interpolation effect is good.

3.2 Mann-Kendall Non-Parametric Test Method

The abrupt changes in water-sediment flux were investigated by analyzing two key components: water discharge variability and sediment load fluctuations. Utilizing datasets from completed annual sediment concentration data, monthly averages of water discharge and sediment load over six years were calculated using Excel's PivotTable function. Specifically, data for identical months across different years (e.g., all January records from 2016 to 2021) were aggregated to generate 12 monthly mean values for both parameters. Under the hypothesis that water flux trends align with water discharge patterns, the analysis of abrupt changes in water flux was equated to studying water discharge variability. The Mann-Kendall (MK) nonparametric test was subsequently applied to detect abrupt shifts in the time series, with monthly and weekly datasets input into the MK algorithm. This method calculates standardized statistics(U_F for forward sequences and U_B for reverse sequences) and identifies intersection points within α 95% confidence interval ($Z_{0.975} = 1.96$) as abrupt change points. Results revealed significant hydrological mutations: (see Tables 2 and 3).

Table 2 Range of Months and Weeks with Abrupt Water Flow

YEAR	MONTH	WEEKS
2016	11	(50,52)
2017	1,7,11	(1,10), (30,40), (40,50)
2018	2	(1,10)
2019	1	(1,10)
2020	2	(1,10)
2021	2	(1,10)

Table 3 Range of Months and Weeks with Sudden Change of Sediment Discharge

YEAR	MONTH	WEEKS
2016	9,11	(35,40), (45,50)
2017	1,7,11	(1,10), (30,35), (40,50)
2018	2	(1,10)
2019	1	(1,10)
2020	2	(1,10)
2021	2	(1,10)

In 2016, abrupt changes in water discharge were observed in November (between Weeks 50 and 51), while sediment load exhibited mutations in September (with two distinct events) and November. During 2017, three discharge mutations occurred in January (Weeks 1–10), July (Weeks 30–40), and November (Weeks 40–50), with sediment load shifts aligning precisely with these intervals. From 2018 to 2021, water discharge mutations consistently emerged in February (Weeks 1–10), except for 2019, where a January mutation (Weeks 1–10) was detected. Notably, sediment load patterns mirrored discharge changes across all years (2017–2021), further confirming the strong linear correlation

between suspended sediment concentration and water discharge identified.

4   CONCLUSION AND OUTLOOKS

This study employed cubic spline interpolation to reconstruct missing sediment concentration data with high precision, establishing a complete annual water-sediment flux dataset spanning 2016–2021. Building upon the original monitoring data and interpolation results, the Mann-Kendall (MK) nonparametric test was subsequently applied to systematically investigate spatiotemporal abrupt changes in water-sediment flux dynamics. The analysis revealed synchronized mutation patterns between water discharge and sediment load during specific months (e.g., January, July, and November 2017) and weekly intervals (e.g., Weeks 1–10). The precise identification of these abrupt change points provides critical temporal thresholds for early warning systems, enabling proactive risk management in flood control, sediment regulation engineering, and ecological conservation across the Yellow River Basin.

While this study successfully addressed missing sediment concentration data using cubic spline interpolation, with interpolated curves demonstrating strong agreement with observed values (mean squared error < 5%), the numerical accuracy of the interpolation results remains limited, particularly at data boundaries and under extreme-value scenarios. Future studies could explore high-precision interpolation methods that integrate multi-source data (e.g., remote sensing retrievals, high-frequency sensor monitoring) or adaptive weighting strategies to enhance local fitting performance, thereby improving the characterization of spatiotemporal heterogeneity in complex water-sediment systems.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1]   Z Cao, Y Zhang, H Chen, et al. Understanding the influencing factors and evolving trends of the Yellow River Water-Sediment Regulation System from a system perspective. Frontiers of Engineering Management, 2024, 11(3): 528–541.
[2]   C Habermann, F Kindermann. Multidimensional spline interpolation: Theory and applications. Computational Economics, 2007, 30: 153–169.
[3]   H A Abdulmohsin, H B Abdul Wahab, A M Jaber Abdul Hossen. A Novel Classification Method with Cubic Spline Interpolation. Intelligent Automation & Soft Computing, 2022, 31(1).
[4]   Y He, H Li. Uncertainty analysis of wind power probability density forecasting based on cubic spline interpolation and support vector quantile regression. Neurocomputing, 2021, 430: 121–137.
[5]   C Liu, Z Zhu, W Hao, et al. Heterogeneous multivariate time series imputation by transformer model with missing position encoding. Expert Systems with Applications, 2025, 271: 126435.
[6]   G T Patle, A Libang, S Ahuja. Analysis of rainfall and temperature variability and trend detection: A non parametric Mann Kendall test approach. 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), IEEE, 2016: 1723–1727.
[7]   J B Cabral Júnior, R L Lucena. Analysis of precipitation using Mann-Kendall and Kruskal-Wallis non-parametric tests. Mercator (Fortaleza), 2020, 19: e19001.
[8]   Y F Sang, Z Wang, C Liu. Comparison of the MK test and EMD method for trend identification in hydrological time series. Journal of Hydrology, 2014, 510: 293–298.
[9]   J Yaméogo. Annual rainfall trends in the Burkina Faso Sahel: a comparative analysis between Mann–Kendall and innovative trend method (ITM). Discover Applied Sciences, 2025, 7(3): 1–17.
[10] J B C Júnior, R L Lucena. Analysis of precipitations by non-parametric tests of mann-kendall and kruskal-Wallis. Mercator-Revista de Geografia da UFC, 2020, 19(1): 19001.
[11] V Gaddikeri, A Sarangi, D K Singh, et al. Trend and change-point analyses of meteorological variables using Mann–Kendall family tests and innovative trend assessment techniques in New Bhupania command (India). Journal of Water and Climate Change, 2024, 15(5): 2033–2058.

GROUNDWATER LEVEL FITTING OF MONITORING WELLS IN THE BAODING REGION BASED ON LONG SHORT-TERM MEMORY (LSTM) NEURAL NETWORKS

Wei Guo

Institute of Hydrogeology and Environmental Geology, Chinese Academy of Geosciences, Shijiazhuang 5050061, Hebei, China.

Corresponding Email: 19831600205@163.com

Abstract: This study applies a Long Short-Term Memory (LSTM) neural network to model daily groundwater level variations at four monitoring wells in the piedmont plain of Baoding City, Hebei Province, China. Using daily data from January 1, 2018 to August 31, 2019, the LSTM model is trained and tested under a one-step rolling prediction framework. The results demonstrate that the LSTM model accurately fits groundwater time series across various hydrogeological conditions, with testing-phase RMSE values ranging from 0.08 to 0.45 meters and R² values exceeding 0.90. The model performs exceptionally well for both stable and fluctuating groundwater level conditions, capturing seasonal decline and recovery patterns without requiring explicit seasonal indicators. It also reveals the model's ability to learn long-term trends and nonlinear dynamics inherent in groundwater systems. Despite its high short-term prediction accuracy, challenges remain regarding multi-step forecasting and responses to extreme events outside the training data. The study concludes that LSTM offers strong potential for groundwater simulation in data-limited environments and recommends further integration of hydro-meteorological variables for enhanced model robustness.
Keywords: Groundwater level; LSTM; Time series modeling; Baoding; Deep learning

1 INTRODUCTION

Groundwater is a vital fundamental natural resource, playing a critical role in global water supply and agricultural production. As of 2010, the annual global extraction of groundwater exceeded 900 billion cubic meters, supplying approximately 36% of drinking water, 42% of irrigation water, and 24% of industrial water. In water-scarce regions of northern China, groundwater serves as the principal water source for both urban and agricultural use[1], earning the title "lifeline beneath the surface." However, because groundwater is stored underground in complex aquifer systems, its level dynamics are subject to various disturbances such as precipitation, evaporation, and pumping. The most direct way to observe groundwater level fluctuations is through long-term monitoring wells, yet the installation of such wells requires significant financial and human resources. Moreover, the number of wells is limited, their spatial distribution is uneven, and their performance is often constrained by geomorphic and topographic conditions, leading to missing data and outliers.

Against this background, machine learning techniques have increasingly been adopted in regional groundwater studies to make better use of sparse monitoring data. Since the 1990s, these methods have been widely applied in hydrological modeling and have demonstrated comparable or superior performance to numerical simulation models[2-4]. In recent years, recurrent neural networks (RNNs) and their variants—particularly the Long Short-Term Memory (LSTM) model—have shown significant promise in time-series groundwater level modeling and prediction [5-6]. LSTM networks learn temporal patterns automatically and are capable of capturing complex nonlinear relationships between input variables and groundwater level responses without requiring explicit assumptions about physical processes [7].

Studies have shown that LSTM generally outperforms traditional statistical and machine learning methods. For instance, Yin compared LSTM with Random Forest (RF) and Artificial Neural Networks (ANNs) for groundwater level prediction and found that LSTM achieved better accuracy in both validation and forecasting stages [8]. Vu demonstrated the model's ability to reconstruct groundwater time series with missing data and to accurately predict future values [9]. In China, Yan developed a multivariable LSTM model based on groundwater monitoring data from Tai'an, Shandong Province, integrating groundwater levels with various meteorological and anthropogenic factors. Their results demonstrated that the multivariable LSTM outperformed both the Backpropagation (BP) neural network and the univariate LSTM model, providing a more accurate simulation of groundwater dynamics[10]. Subsequently, Sun further optimized the model structure and input data processing in a case study of Jinan City. They found that fitting temperature series using sinusoidal functions and tuning the dropout rate significantly improved the model's stability and prediction accuracy. For the Quaternary aquifer, the optimized model achieved a prediction RMSE of less than 0.84 m[11]. Additionally, Sun compared ARIMA, BP neural networks, and LSTM for monthly and daily groundwater level prediction in the North China Plain, and confirmed the superiority of deep learning models in terms of accuracy [12].

Although the use of LSTM in groundwater modeling has become increasingly widespread, its application in fine-scale fitting of groundwater level time series in the piedmont plains of Baoding, Hebei Province, remains limited. Given the critical role of groundwater in regional water supply and the complexity of its dynamics in this area, this study aims to construct an LSTM-based model using daily groundwater level data from January 1, 2018 to August 31, 2019. The model will be trained and validated to evaluate its ability to simulate groundwater fluctuations and to assess its prediction accuracy and applicability, thereby offering insights for regional groundwater resource management..

2 STUDY AREA AND DATE

2.1 Study Area Overview

Baoding City is located in the northwestern part of the North China Plain. The piedmont zone of the Taihang Mountains refers to the transitional area between the eastern foothills of the Taihang Mountains and the adjoining plain, representing a region where the mountainous terrain of western Baoding gradually shifts eastward into the plains. The terrain in this region gradually flattens from west to east. Geologically, the area lies at the front edge of the alluvial-proluvial fan of the Taihang Mountains, consisting of thick Quaternary unconsolidated sediments that form the primary aquifer systems. Historically, this area was recognized as one of the most severely over-exploited groundwater zones in the North China Plain. Groundwater levels continuously declined since the late 20th century, forming a widespread groundwater depression cone, which reached its lowest point around 2010. In recent years, following the implementation of the South-to-North Water Diversion Project (Middle Route) and comprehensive measures to manage groundwater over-extraction [13], groundwater levels in the region have shown signs of recovery. Nevertheless, groundwater dynamics in this area remain highly sensitive and complex, calling for scientifically robust monitoring and modeling tools to support sustainable groundwater management.

2.2 Data Source

This study utilizes daily groundwater level elevation data from four national monitoring wells located in the piedmont plain of Baoding City, spanning the period from January 1, 2018 to August 31, 2019. The spatial distribution of these wells is illustrated in Figure 1. The selected wells cover a range of hydrogeological settings along a west–east transect across the piedmont plain, including both recharge zones near the Taihang Mountain foothills and overexploited zones within the central plain. The original time series is complete without any missing records. For a few isolated wells exhibiting abnormal peak values, such anomalies were identified as instrument errors and corrected through preprocessing: the outlier values were replaced by the average of the adjacent valid values to minimize the impact of outliers on model training. After preprocessing, the groundwater level time series for each monitoring well displayed a clear trend and are considered suitable for subsequent model training and validation.
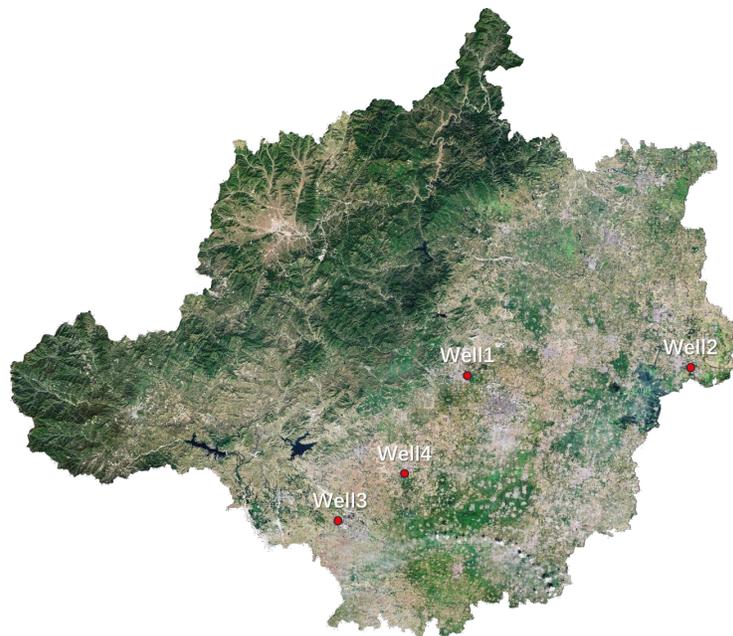
Figure 1 Monitoring Well Locations (negative from google earth)
Source: https://earth.google.com/

3 METHODOLOY

3.1 Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) neural network is a variant of the recurrent neural network (RNN) architecture designed to handle long-term dependencies in sequential data[14]. Unlike traditional RNNs that are prone to vanishing or exploding gradient issues, LSTM introduces a gating mechanism that effectively controls the processes of information retention and forgetting. A standard LSTM unit consists of an input gate, a forget gate, an output gate, and a cell state. At each time step, the LSTM unit receives the current input vector, along with the previous hidden state and cell state, and computes the updated states through gate operations. The forget gate determines which information from the previous cell state should be discarded based on the current input and previous hidden state. The input gate regulates the extent to which new information is written into the cell state. The updated cell state combines these two processes to maintain long-term memory. Finally, the output gate determines which parts of the cell state are passed to the hidden state output. These gating operations enable LSTM units to preserve relevant information over tens or even hundreds of time steps, making the model well-suited for capturing long-term trends and seasonal patterns in groundwater level time series[15]. Mathematically, the LSTM computations involve sigmoid activations for the gates and tanh activation for candidate memory states; the detailed equations are omitted here for brevity. The model is trained using gradient descent to update the gate weights, minimizing the difference between the predicted and observed sequences. Compared to traditional time-series methods or shallow neural networks, LSTM's ability to retain long-range dependencies provides a significant advantage in modeling delayed-response processes such as groundwater level dynamics[16].

3.2 Model Architecture and Parameter Configuration

The LSTM model was implemented and trained using Python. For each monitoring well, the groundwater level time series was transformed into supervised learning data by creating input-output pairs. Given the relatively stable daily changes and seasonal cycles of groundwater levels, a sliding window of N = 30 days (approximately one month) was used as the input sequence length. Each input sequence consisted of groundwater levels over the previous 30 days, and the target output was the groundwater level on the following day (N+1). This sliding window approach was applied to extract training samples from the historical time series. To ensure data consistency across wells, the raw groundwater levels were normalized using Min-Max scaling, mapping values to the range [0, 1]. After prediction, the outputs were rescaled back to their original range via inverse normalization.

The LSTM model architecture comprised a single LSTM hidden layer followed by a fully connected output layer. The number of hidden units was determined experimentally to balance accuracy and prevent overfitting, and 16 units were selected. The hidden layer used the tanh activation function, while the output layer employed a linear activation to generate continuous numerical predictions. The model was trained using the Backpropagation Through Time (BPTT) algorithm, with the Adam optimizer and an initial learning rate of 0.01, which was adaptively adjusted during training to accelerate convergence. The loss function used was Mean Squared Error (MSE). To improve generalization, the training data was randomly shuffled, and mini-batch gradient descent (batch size = 16) was applied for parameter updates. The number of training epochs was determined based on convergence trends in the validation loss, and training was terminated when validation error ceased to decrease.

3.3 Training and Testing Set Division

To evaluate the generalization performance of the LSTM model, the complete time series was split into training and testing subsets. Following a conventional 80/20 split, the period from January 1, 2018 to April 30, 2019 (approximately 486 days) was used as the training set, while the period from May 1, 2019 to August 31, 2019 (approximately 122 days) was used as the independent test set. It is important to note that model training was conducted exclusively on the training data, and the testing phase incorporated a rolling prediction mechanism. For example, to predict groundwater levels on May 2, 2019, the model used real observations from the 30-day period ending on May 1 as input. The model then predicted the next day, and this process continued iteratively using actual observed values to update the input window, ensuring sequential coherence in multi-step forecasting.

3.4 Evaluation Metrics

The model's prediction performance was evaluated on the test set using two common metrics: Root Mean Square Error (RMSE) and the Coefficient of Determination (R^2). RMSE quantifies the average deviation between predicted and observed values, with lower values indicating better accuracy. R^2 measures the proportion of variance in the observed data explained by the model, with values closer to 1 indicating better fit. These metrics were used to quantitatively assess the LSTM's performance across different monitoring wells. In addition, predicted and observed groundwater level curves were plotted for visual inspection, allowing for a direct comparison of the model's simulation capability and error distribution.

4 MODEL FITTING AND RESULT ANALYSIS

4.1 Model Training Process

Following the methodology described above, separate LSTM models were trained for the groundwater level time series of the four monitoring wells. During the training phase, the loss function for each model exhibited a rapid decline followed by stabilization, indicating effective convergence. After approximately 100 epochs, the training error reached a plateau, and the validation error no longer decreased, suggesting that the model had achieved its optimal fitting capacity. At this point, the Root Mean Square Error (RMSE) on the training set for each well ranged between 0.1 and 0.3 meters, demonstrating a high level of fitting accuracy on the observed data (Figure 2).
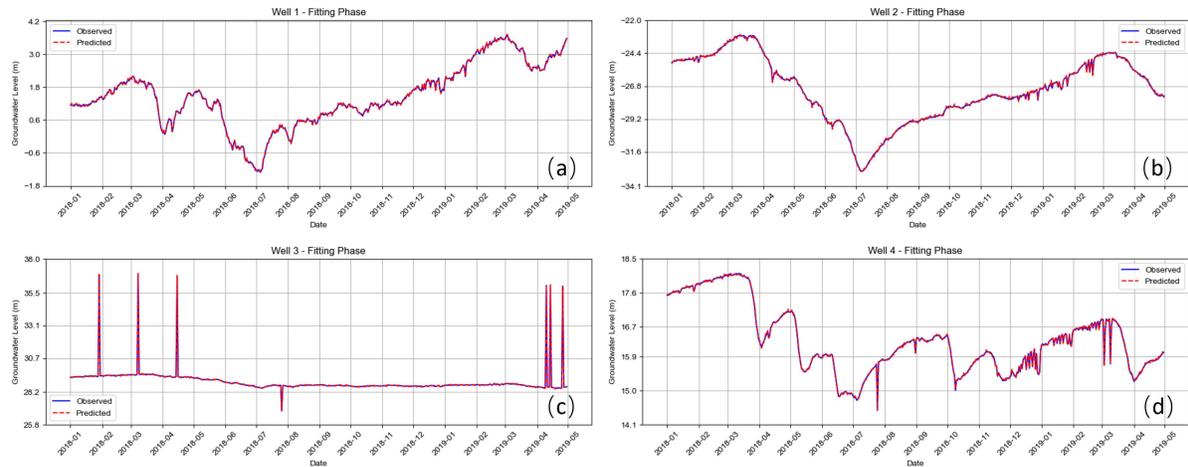
Figure 2 Groundwater Level Fitting on the Training Set. Subplots (a)–(d) Correspond to Wells 1 through 4, Respectively. The Red Dashed Lines Represent the Predicted Groundwater Levels, while the Blue Solid Lines Indicate Observed Values.

4.2 Testing Phase Fitting Performance

The trained models were then applied to the testing set, which spanned from May to August 2019, to generate daily predictions of groundwater levels. The results are shown in Figure 3, where the predicted groundwater levels are compared against observed values for each of the four monitoring wells. The LSTM models effectively reproduced the temporal dynamics of groundwater levels, with strong agreement between the predicted and observed curves. In terms of seasonal variation, the models successfully captured the summer decline and early autumn recovery patterns typical of groundwater levels. Furthermore, the models accurately reflected broader trends of sustained decline or gradual rise in groundwater levels across different wells.

4.3 Error Analysis

The four monitoring wells selected for this study exhibited distinct groundwater level variation patterns during the testing period from May to August 2019, yet the LSTM model demonstrated satisfactory fitting performance across all cases (Table 1). For Well 1, the groundwater level first increased and then decreased, dropping from approximately 3.2 m to 1.8 m. The LSTM model successfully captured both the overall trend and the inflection points, with an RMSE of 0.21 m and an R^2 of 0.94. Well 2 is located in a region of intensive groundwater overexploitation and experienced a sharp decline of nearly 5 m. The model accurately reproduced the nonlinear trend and slope transitions, although a slight underestimation occurred toward the end of the period; the RMSE and R^2 were 0.45 m and 0.90, respectively. Well 3 exhibited stable water levels with a total decline of less than 1 m. The model's predictions were nearly indistinguishable from the observed data, achieving an RMSE of only 0.08 m and a high R^2 of 0.99, indicating excellent fitting accuracy. For Well 4, the groundwater level dropped by approximately 1.3 m, accompanied by minor fluctuations. The LSTM model accurately replicated both the overall trend and localized short-term rebounds, with an RMSE of 0.18 m and R^2 of 0.96.

Table 1 Monitoring Well Error Statistics

Monitoring well number	RMSE (training set)	R^2 (training set)	RMSE (test set)	R^2 (test set)
Well 1	0.036	0.99	0.21	0.94
Well 2	0.073	0.99	0.45	0.9
Well 3	0.028	0.99	0.08	0.99
Well 4	0.026	0.99	0.18	0.96

In summary, the LSTM model demonstrated strong nonlinear learning capability and high accuracy across varying hydrogeological settings and magnitudes of groundwater level change. It effectively captured the dynamic characteristics of groundwater fluctuations in the study area, confirming its applicability in regional groundwater modeling.

5 CONCLUSIONS AND FUTURE PERSPECTIVES

This study applied a Long Short-Term Memory (LSTM) neural network model to simulate daily groundwater level variations at four monitoring wells in the piedmont plain of Baoding City. The main conclusions are as follows:
The LSTM model achieved high-precision fitting of groundwater level time series. After being trained on approximately 1.5 years of daily data, the model produced predictions that closely matched observed groundwater levels during the testing phase. The mean absolute error ranged from 0.1 to 0.45 meters, and the coefficient of determination (R^2) generally exceeded 0.90, indicating that the model could explain the majority of groundwater level variance. Particularly for wells with stable and gradually changing levels, the fitting accuracy approached the magnitude of measurement error. For wells with significant seasonal fluctuations or trends, the LSTM model effectively captured the peaks, troughs, and long-term trajectories. These results confirm that LSTM networks are well-suited for modeling hydrological time series such as groundwater levels, which exhibit long-term memory characteristics.
The model automatically learned seasonal and trend components. Leveraging the memory capabilities of gated recurrent units, the LSTM model captured seasonal patterns—such as summer declines and winter recoveries—as well as long-term trends like multi-year depletion, without requiring explicit inclusion of seasonal or trend indicators. During the testing period, the model responded accurately to seasonal turning points, distinguishing dry and wet periods. This demonstrates the advantage of deep learning models in capturing latent patterns. In contrast, traditional approaches like linear regression or ARIMA typically require pre-processing (e.g., detrending or adding exogenous seasonal terms), whereas LSTM can simultaneously model multiple forms of variation.
While the model performed exceptionally well in short-term predictions, its accuracy over longer forecasting horizons remains uncertain. This study focused on one-step-ahead rolling forecasts, using recent observed values to update predictions continuously. Under these conditions, the model achieved excellent results (e.g., $R^2 = 0.99$). However, for multi-step forecasts extending weeks or months into the future, prediction errors may accumulate over time. In such cases, periodic correction or hybridization with physically-based models may be needed to mitigate error propagation.
The model's performance is dependent on the range of training data and may have limited adaptability to abnormal or abrupt conditions. As the LSTM model learns entirely from historical data, it may struggle to respond accurately to events not represented in the training set, such as sudden recharge from extreme rainfall or rapid drawdown from unexpected over-pumping. In this study, abnormal values were filtered to improve model stability. Nonetheless, this highlights the need to

supplement model outputs with real-time field data to validate and adjust predictions, particularly when encountering conditions outside the historical data distribution.

In summary, the LSTM model delivered satisfactory results for groundwater level fitting in the Baoding piedmont plain, demonstrating strong potential as a data-driven modeling tool. Nevertheless, there remain areas for improvement and further investigation. For instance, this study used only historical groundwater level data as model input. In future work, incorporating additional variables such as precipitation, evaporation, and pumping rates may help build a multivariate predictive model with enhanced robustness and improved responsiveness to anomalous scenarios.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Wang K, Chen H, Fu S, et al. Analysis of exploitation control in typical groundwater over-exploited area in North China Plain. Hydrological Sciences Journal, 2021, 66(5): 851-861. DOI: 10.1080/02626667.2021.1900575.

[2] Kratzert F, Klotz D, Brenner C, et al. Rainfall–runoff modelling using long short-term memory (LSTM) networks. Hydrology and Earth System Sciences, 2018, 22(11): 6005-6022. DOI: 10.5194/hess-2018-247-sc1.

[3] Lees T, Buechel M, Anderson B, et al. Benchmarking data-driven rainfall–runoff models in Great Britain: a comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. Hydrology and Earth System Sciences, 2021, 25(10): 5517-5534. DOI: 10.5194/hess-25-5517-2021.

[4] Yang S, Yang D, Chen J, et al. A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. Journal of Hydrology, 2020, 590: 125206. DOI: 10.1016/j.jhydrol.2020.125206.

[5] Zounemat-Kermani M, Batelaan O, Fadaee M, et al. Ensemble machine learning paradigms in hydrology: A review. Journal of Hydrology, 2021, 598: 126266. DOI: 10.1016/j.jhydrol.2021.126266.

[6] Nearing G S, Kratzert F, Sampson A K, et al. What role does hydrological science play in the age of machine learning?. Water Resources Research, 2021, 57(3): e2020WR028091. DOI: 10.31223/osf.io/3sx6g.

[7] Bai T, Tahmasebi P. Graph neural network for groundwater level forecasting. Journal of Hydrology, 2023, 616: 128792. DOI: 10.1016/j.jhydrol.2022.128792.

[8] Yin W, Fan Z, Tangdamrongsub N, et al. Comparison of physical and data-driven models to forecast groundwater level changes with the inclusion of GRACE–A case study over the state of Victoria, Australia. Journal of Hydrology, 2021, 602: 126735. DOI: 10.1016/j.jhydrol.2021.126735.

[9] Vu T D, Ni C F, Li W C, et al. Predictions of groundwater vulnerability and sustainability by an integrated index-overlay method and physical-based numerical model. Journal of Hydrology, 2021, 596: 126082. DOI: 10.1016/j.jhydrol.2021.126082.

[10] Baizhong Y, Jian S, Xinzhou W, et al. Multivariable LSTM neural network model for groundwater levels prediction.Journal of Jilin University (Earth Science Edition), 2020, 50(1): 208-216. DOI: 10.13278/j.cnki.jjuese.20190055.

[11] Hongjie S, Zhenhua Z, Linxian H, et al. Application of Multi-Variable LSTM Neural Network Model for Groundwater Levels Prediction. Yellow River, 2022, 44(8). DOI: 10.3969 /j.issn.1000-1379.2022.08.014.

[12] Sun J, Hu L, Li D, et al. Data-driven models for accurate groundwater level prediction and their practical significance in groundwater management. Journal of Hydrology, 2022, 608: 127630. DOI: 10.1016/j.jhydrol.2022.127630.

[13] Zhang C, Duan Q, Yeh P J F, et al. Sub-regional groundwater storage recovery in North China Plain after the South-to-North water diversion project. Journal of Hydrology, 2021, 597. DOI: 10.1016/j.jhydrol.2021.126156.

[14] Rajaee T, Ebrahimi H, Nourani V. A review of the artificial intelligence methods in groundwater level modeling. Journal of hydrology, 2019, 572, 336-351. DOI： 10.21203/rs.3.rs-2915223/v1.

[15] Wai K P, Chia M Y, Koo C H, et al. Applications of deep learning in water quality management: A state-of-the-art review. Journal of Hydrology, 2022,613, 128332. DOI： 10.1016/j.jhydrol.2022.128332.

[16] Van Thieu N, Barma S D, Van Lam T, et al. Groundwater level modeling using augmented artificial ecosystem optimization. Journal of Hydrology,2023, 617, 129034. DOI: 10.1016/j.jhydrol.2022.129034.

SURFACE DEFORMATION MONITORING AND ANALYSIS IN XINJING MINING AREA BASED ON SBAS-INSAR TECHNOLOGY

FangZhou Fu

College of Intelligent Mining Engineering, China University of Mining and Technology-Beijing, Beijing 100080, China.
Corresponding Email: 13007343782@163.com

Abstract: Large-scale mining activities in mining areas can induce surface deformation phenomena, threatening the stability of peripheral production systems and residential environmental safety. Consequently, implementing high-precision surface deformation monitoring has become a critical component of mine safety assessment systems and ecological restoration projects. This study takes the Xinjing Mining Area in Alxa League, as a case study. Utilizing 20 scenes of Sentinel-1 ascending orbit images acquired from January 2022 to December 2023, this study applied the SBAS-InSAR technique for deformation monitoring and analysis. The results demonstrated complex deformation patterns in the Xinjing Mining Area during the study period, primarily characterized by subsidence. The deformation rate across the study area ranged from -125.8 mm/a to +36.3 mm/a, with cumulative deformation magnitudes ranging from -213.3 mm to +76.7 mm. The deformation process evolved through three distinct phases: an accelerated deformation phase, a moderate deformation phase, and a stabilization phase. This study provides essential data support and a scientific basis for preventing mining-induced subsidence caused by engineering activities in mining areas.
Keywords: SBAS-InSAR; Alxa open-pit coal mine; Surface deformation monitoring; Spatiotemporal evolution

1 INTRODUCTION

Mineral resource exploitation serves as the foundation supporting modern industrial development [1], yet the surface deformation it induces has become a global focus in mine environmental governance [2]. Processes such as open-pit excavation, underground goaf collapse, and groundwater level fluctuations can lead to surface deformation, ground fissure propagation, and even landslides, posing significant threats to mining safety and surrounding communities [3-4]. Although conventional monitoring techniques (e.g., GNSS and precise leveling) provide millimeter-level deformation accuracy [5-6], their low spatial density and high costs limit their applicability for large-scale dynamic monitoring in mining areas [7-8]. Consequently, high-resolution, wide-coverage remote sensing technologies for surface deformation monitoring hold urgent practical significance for mine hazard early warning and ecological restoration [9].

Interferometric Synthetic Aperture Radar (InSAR) technology, utilizing its all-weather and large-scale observational capabilities [10], has revolutionized deformation monitoring in mining areas. In recent years, time-series InSAR techniques such as SBAS-InSAR have significantly enhanced the accuracy and spatiotemporal continuity of deformation inversion by constructing optimized small-baseline interferometric networks. These advancements have demonstrate unique advantages in monitoring slow and nonlinear deformation [11-13]. Currently, this technology has been widely applied in multiple domains, including mine subsidence monitoring, surface settlement prediction, urban subsidence rate assessment, and spatiotemporal characteristic analysis of mining-induced deformation [14-15]. SBAS-InSAR, with its distinctive all-weather observation capability, continuous monitoring characteristics, and high spatial resolution and measurement precision, provides innovative technical solutions for unraveling the spatiotemporal evolution patterns of mining-induced subsidence [16].

A catastrophic collapse occurred at the Xinjing Coal Mine on February 22, 2023 [17], making time-series deformation monitoring of this mine critical for ensuring mining safety and public safety. While previous studies, notably by Ma et al. [18], have extensively documented the pre-slope-destabilization surface deformation patterns in the mining area, their conclusions remain constrained by the lack of systematic deployment of ground monitoring networks to validate deformation magnitudes, particularly in critical zones with high displacement gradients. Furthermore, existing literature predominantly focuses on pre-collapse deformation mechanisms, leaving the dynamic post-collapse evolution—such as stress redistribution and secondary instability risks—insufficiently resolved. To address these limitations, this study provides a comprehensive analysis of surface deformation spanning both pre- and post-avalanche phases. By integrating multi-temporal InSAR observations with post-event field validations, we systematically quantify the cascading impacts of slope failure on deformation regimes, thereby advancing a more holistic understanding of mining-induced geohazard progression. This study utilizes 20 Sentinel-1 ascending orbit images acquired from January 2022 to December 2023. Using SBAS-InSAR technology, we conducted deformation monitoring and analysis in the Xinjing Mining Area, thereby providing data support for preventing mining-induced subsidence caused by engineering activities.
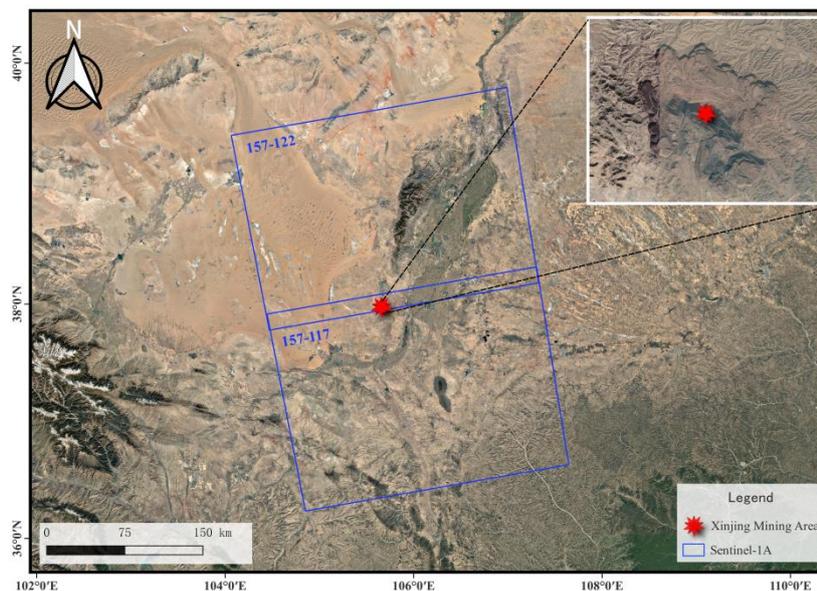
2 STUDY AREA

The Xinjing open-pit coal mining area, a representative northern mining region, has faced escalating subsidence-related challenges, including recurrent landslides and agricultural land degradation, driven by decades of intensive extraction activities. The sparse vegetation coverage and geomorphological stability collectively ensure sufficient coherence for interferometric processing. Comprehensive mining records, hydrological datasets, and ground-based measurements are systematically archived, enabling robust cross-validation with InSAR-derived deformation results. This study's outcomes are expected to directly inform disaster prevention strategies and operational loss mitigation, thereby demonstrating dual significance in advancing subsidence mechanism research and promoting sustainable mining practices.

The Xinjing Coal Mine is located in the northwestern part of Alxa Left Banner, Alxa League, with geographical coordinates ranging from 105°30′E to 105°45′E and 39°10′N to 39°25′N, adjacent to the western foothills of the northern Helan Mountains. The mining area exhibits an irregular polygonal shape, extending approximately 7.2 km north-south and 4.5 km east-west, covering a total area of 32.4 km². It borders the Helan Mountain National Nature Reserve to the north and the Tengger Desert Marginal Ecological Restoration Project Area to the south, serving as a critical ecological node in the desert-oasis transition zone.

The mining area lies within the intermixing zone of Helan Mountain piedmont alluvial-proluvial fans and desert terrain, gently sloping from northwest to southeast with gradients of 2° – 5° and an average elevation of 1,328 meters. Surface coverage consists of Quaternary aeolian sand deposits (3 – 8 m thick) , locally exposing Jurassic sandstone bedrock. The climate is characterized by a mid-temperate extreme arid continental regime, featuring an annual precipitation of 127 mm (primarily concentrated from July to September) and evaporation exceeding 2,800 mm. The mean annual temperature ranges between 8 – 10 °C, with summer extreme highs reaching 40 °C and winter lows dropping to -25 °C. The region experiences a mean annual wind speed of 4.2 m/s dominated by northwest winds, with an annual sandstorm frequency of 21 days.

Coal seams in the mining area demonstrate a distinct north-south strike orientation, with thickness gradually increasing from 20 meters in the north to 100 meters in the south, forming a stable coal distribution pattern as shown in the satellite image map (Figure 1). The coal-bearing strata belong to the Taiyuan Formation and Shanxi Formation, featuring average seam thicknesses of 8 – 15 meters, dip angles of 5° – 15°, and burial depths of 50 – 200 meters. The coal is primarily long-flame type with calorific values of 4,500 – 5,500 kcal/kg, medium ash content (18% – 25%), and medium sulfur content (1.2% – 1.8%), making it suitable for power generation and chemical industries.

Figure 1 Overview Map of the Study Area

3 DATA AND METHODS

3.1 Data Source

To investigate surface deformation in the Xinjing Mining Area, this study employs radar satellite data for analysis. Sentinel-1A Synthetic Aperture Radar (SAR) data were acquired in Interferometric Wide (IW) swath mode with VV polarization, specifically utilizing Single Look Complex (SLC) products. A total of 20 Ascending orbit images spanning from January 7, 2022, to December 16, 2023, were selected for deformation analysis. Technical parameters of the Sentinel-1A imagery are detailed in Table 1. Additionally, the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) with 30 m spatial resolution was applied to compute topographic parameters, enabling precise geocoding and co-registration of SAR images.

Data	Parameter	Description
Sentinel-1A	Time Span	2022/01/07~2023/12/16
	Data Volume	20
	Orbit Direction	Ascending
	Heading Angle	-13°11′
	Acquisition Mode	IW
	Polarization Mode	VV

3.2 SBAS-InSAR Methods

The SBAS-InSAR (Small Baseline Subset Interferometric Synthetic Aperture Radar) technique extract surface deformation information from multi-temporal SAR images by constructing a small-baseline interferometric network. This approach reduces decorrelation noise (e.g., coherence loss in vegetated areas) and improves data usability through the generation of phase interferograms. By selecting N+1 SAR scenes with uniform temporal spacing, M differential interferograms are generated based on predefined temporal and spatial baseline thresholds, where M satisfies the condition: $(N+1)/2 \leq M \leq N(N+1)/2$ The differential interferometric phase comprises contributions from deformation, residual topography, atmospheric delay, orbital errors, and noise, expressed mathematically as:

$$\Delta\varphi = \frac{4\pi}{\lambda}d(t) + \frac{4\pi B_\perp}{\lambda R \sin\theta}\Delta z + \varphi_{atm} + \varphi_{orb} + \varphi_{noise}$$

（1）

In the equation:

$\frac{4\pi}{\lambda}d(t)$ ——Phase variation induced by surface deformation (target signal), where d(t) represents the cumulative deformation displacement.

$\frac{4\pi B_\perp}{\lambda R \sin\theta}\Delta z$ ——The phase residual induced by DEM elevation error Δz exhibits positive correlation with the vertical baseline $B_\perp$.

$\varphi_{atm}、\varphi_{orb}、\varphi_{noise}$ ——Atmospheric delay phase (dominant error), orbital error phase, and noise-induced phase error.

The Sentinel-1 SAR datasets were processed using SBAS-InSAR techniques to derive millimeter-level accuracy time-series deformation fields and mean displacement velocity fields across the study area and its surroundings. Through these outputs, we analyzed the spatiotemporal evolution characteristics of surface deformation and revealed potential hazard development patterns. The technical workflow is illustrated in Figure 2.
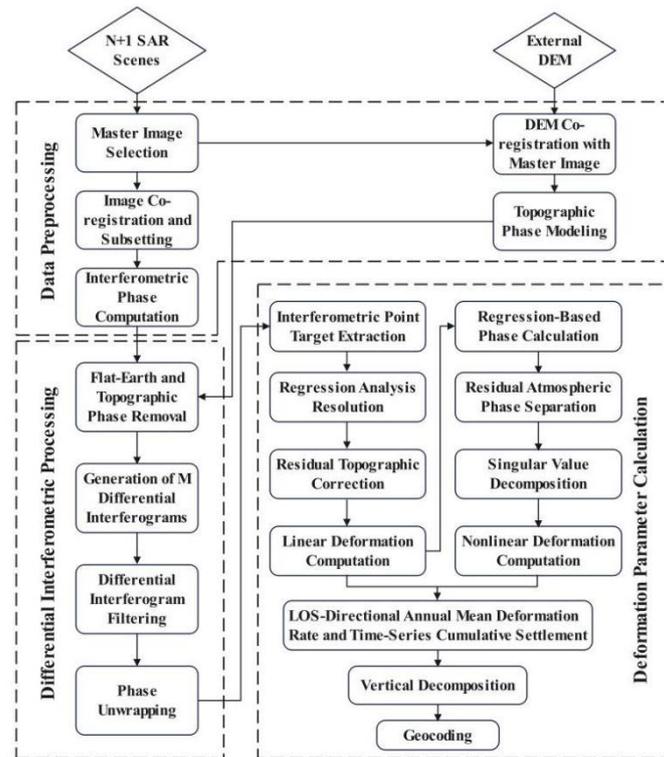
Figure 2 Technical Workflow Diagram

The key implementation steps are as follows:

(1) Master Image Selection:

A master image is selected from N+1 scenes by maximizing the overall coherence of all interferometric pairs. A multiplicative model evaluates four parameters: temporal baseline, spatial perpendicular baseline, Doppler centroid difference, and thermal noise. The image achieving maximum coherence is designated as the master to minimize co-registration errors and enhance alignment precision.

(2) Image Co-registration:

SAR slave images are co-registered to the master image to generate interferograms under spatiotemporal baseline constraints. To mitigate decorrelation caused by excessive baselines, a weighted least squares method with offset standard deviation weighting is implemented. Homologous points are identified via intensity cross-correlation or fringe sharpness algorithms, with weights assigned based on coherence values to solve offset polynomials. Precise orbit data and DEM-assisted range-Doppler modeling achieve sub-pixel alignment accuracy.

(3) Differential Interferogram Generation:

The flat-earth effect induces dense fringe patterns, complicating subsequent filtering and unwrapping. Prior to differential interferogram generation, SRTM-3 DEM data (30 m spatial resolution) is utilized to remove flat-earth and topographic phase contributions. Interferometric pairs are then selected based on temporal and spatial baselines, employing a multi-master image strategy to construct initial differential interferograms. Given the study area's complex geomorphology and significant topographic variations, high-quality interferograms are prioritized to mitigate residual topographic phase errors.

(4) Phase unwrapping:

The raw interferometric phase comprises not only surface deformation components but also residual orbital phases, topographic residuals, atmospheric delay phases, and noise. Due to the strong spatial randomness of topographic residuals and noise phases, direct unwrapping is infeasible. This study adopts a height-correction-based unwrapping approach. While height correction terms and noise phases exhibit spatial randomness, atmospheric delays and orbital residuals demonstrate spatial correlation. To address this, spatial-domain filtering is applied to estimate spatially correlated components, followed by adaptive filtering methods to remove these artifacts, yielding unwrapped interferograms. In areas with rapid land cover changes, significant topographic variations, or atmospheric coherence degradation, unwrapped interferograms still exhibit incomplete results in certain regions. Consequently, it is necessary to discard interferograms with poor unwrapping outcomes to ensure reliability in subsequent analyses.

(5) Deformation Estimation:

The unwrapped phase undergoes Singular Value Decomposition (SVD) to solve the least-squares problem for time-series deformation. Residual topographic phases and noise components (non-spatially correlated) are extracted by differencing the original unwrapped phases, enabling elevation correction and deformation sequence refinement. Deformation comprises linear and nonlinear components. Temporal filtering isolates time-correlated nonlinear deformation phases from uncorrelated atmospheric delays. The nonlinear deformation is then combined with spatially

filtered linear deformation to derive the final time-series deformation phase. The resulting line-of-sight (LOS) mean deformation rate and cumulative displacement time-series are vertically decomposed using the incidence angle. Subsequent geocoding transforms the results from SAR coordinates to the WGS84 geographic coordinate system , enabling spatial analysis of mining-induced subsidence patterns.

4 RESULTS ANALYSIS AND DISCUSSION

4.1 Surface Deformation Velocity in the Mining Area

By applying SBAS-InSAR techniques to 20 Sentinel-1A scenes acquired from January 7, 2022, to December 16, 2023, we generated the mean deformation rate map of the Xinjing Mining Area (Figure 3). In this map: Positive values (blue) denote ground movement toward the satellite, corresponding to uplift, while negative values (red) signify displacement away from the satellite, indicative of subsidence.
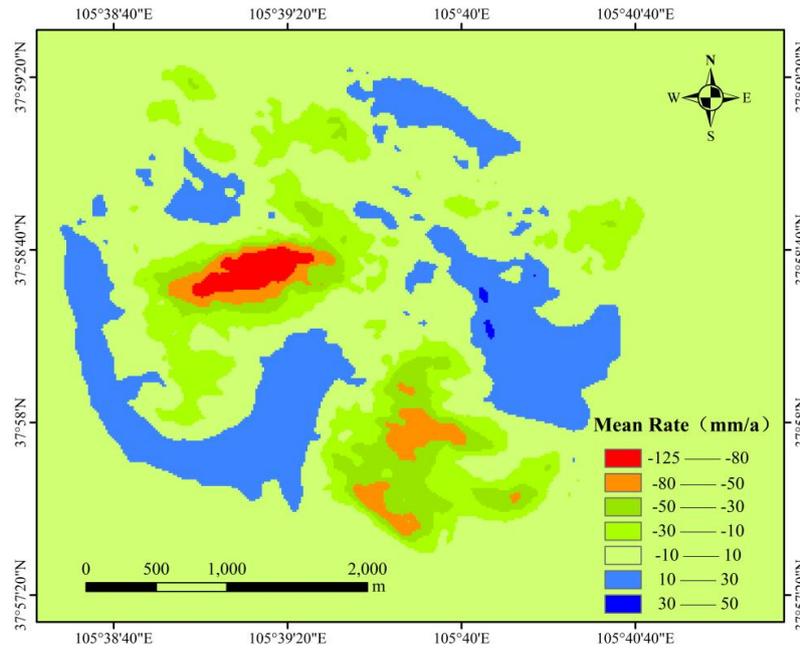
Figure 3 Mean Deformation Rate Map of the Xinjing Mining Area (2022 - 2023)

Figure 3 demonstrates pronounced surface deformation within the mining area, revealing distinct subsidence patterns across varying velocity ranges. Subsidence zones with rates of -125 to -80 mm/a cover 0.13 km² (4.9% of the total subsidence area), while those at -80 to -50 mm/a, -50 to -30 mm/a, and -30 to -10 mm/a occupy 0.28 km² (11.2%), 0.58 km² (23.3%), and 1.52 km² (60.6%), respectively. Notably, vertical uplift deformation (total area: 2.17 km²) is observed and preliminarily attributed to spoil material accumulation from open-pit mining operations.

4.2 Spatiotemporal Evolution of Surface Deformation in the Mining Area

To better investigate the spatiotemporal evolution characteristics of the mining area, this study obtained cumulative deformation displacements at different time intervals in the study area, as shown in Figure 4. The spatial evolution of deformation from 2022 to 2023 used January 7, 2022, as the reference baseline, with the surface deformation on that date set to 0 mm. Figure 5 demonstrates that from January to June 2022, the deformation distribution area in the study region significantly increased with accelerated deformation rates, resulting in cumulative displacements ranging from -142.0 mm to +27.9 mm. From July 2022 to March 2023, the deformation area continued to expand, though the rates slowed, with cumulative displacements reaching -213.1 mm to +63.8 mm. During the period from April to December 2023, the deformation process entered a relatively stable phase, with cumulative displacements showing only minor fluctuations between -213.3 mm and +76.7 mm. The spatiotemporal evolution reveals that the deformation extent expanded from 3.68 km² in June 2022 to 5.47 km² by November 2022, ultimately reaching 7.12 km² by December 2023. This evolutionary process clearly indicates that the surface deformation in the study area not only exhibits significant spatial expansion but also demonstrates continuous growth in cumulative displacement.

To further investigate deformation trends in the mining area, six representative feature points were selected in Figure 4, Record their deformation information in Table 2. And their cumulative deformation curves are plotted in Figure 5. Figure 5 reveals sustained growth trends in cumulative deformation across all monitoring points, with significant rate variations. Notably, point C2 exhibits the maximum cumulative deformation of -171.1 mm, while point C1 shows the minimum displacement at +99.6 mm.
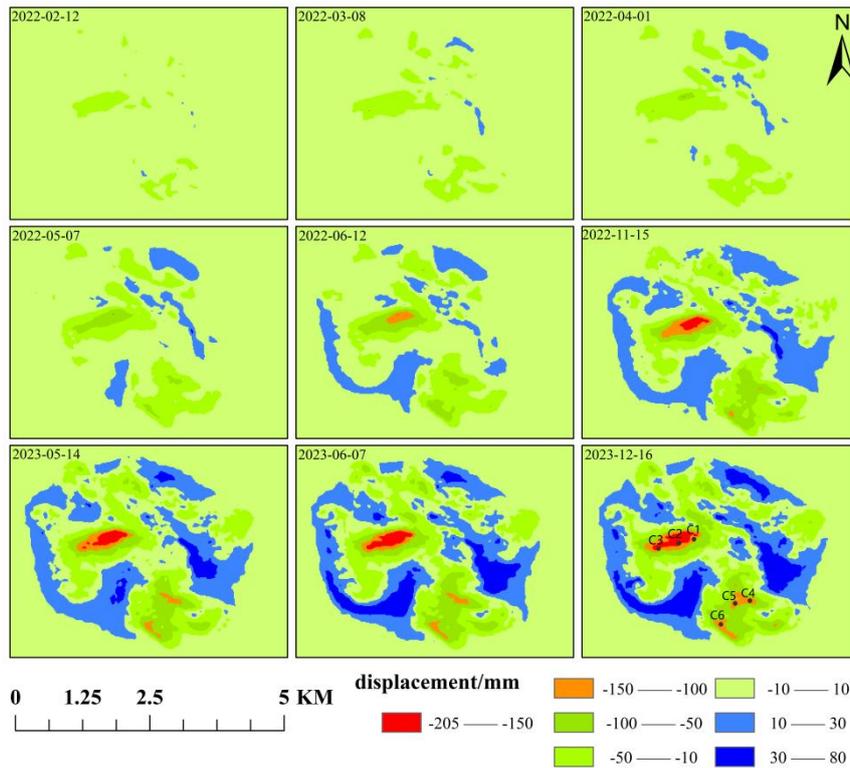
Figure 4 Spatiotemporal Deformation Evolution Map

Specifically, the temporal evolution characteristics of surface deformation at the six monitoring points can be broadly categorized into three distinct phases:

(1) Accelerated Deformation Phase: From January to June 2022, deformation rates at all six monitoring points increased substantially, with cumulative displacements growing progressively. During this period, point C2 exhibited the maximum deformation, reaching -99.5 mm by June 12, 2022, corresponding to an average deformation rate of -198.3 mm/a. In contrast, point C5 showed the minimum deformation, with a cumulative displacement of -44.3 mm and an average rate of -88.1 mm/a. The rapid deformation progression is directly attributable to high-intensity mining operations, which triggered abrupt stress redistribution in overburden strata.

(2) Moderate Deformation Phase: From July 2022 to March 2023, point C3 exhibited the maximum cumulative deformation of -66.8 mm with a mean deformation rate of -95.5 mm/a, while point C1 showed the minimum displacement (-27.5 mm, mean rate: -39.3 mm/a). Notably, a major landslide occurred on February 22, 2023, during the terminal phase of this period. Analysis of cumulative deformation curves reveals that the excessive deformation rates induced by sustained high-intensity stripping operations during Phase 1 directly triggered this disaster. Preceding the event, monitoring points C1, C2, and C3 exhibited daily subsidence rates exceeding 0.5 mm/day. Such rapid stripping rates promoted the formation of oversteepened slopes, with progressive stability degradation over time ultimately leading to slope failure.

(3) Stabilization Phase: From April to December 2023, deformation rates at all six monitoring points decelerated and gradually stabilized. During this period, only point C5 retained residual deformation of -37.6 mm, while displacements at other points remained within -20 mm. Notably, point C1 exhibited an anomalous reduction in cumulative deformation of 23.5 mm. These trends are closely linked to post-landslide mitigation measures, including production suspension, operational rectification, and slope reinforcement engineering. The stabilization demonstrates that anthropogenic interventions effectively suppressed further deformation progression.

(a)

(b)
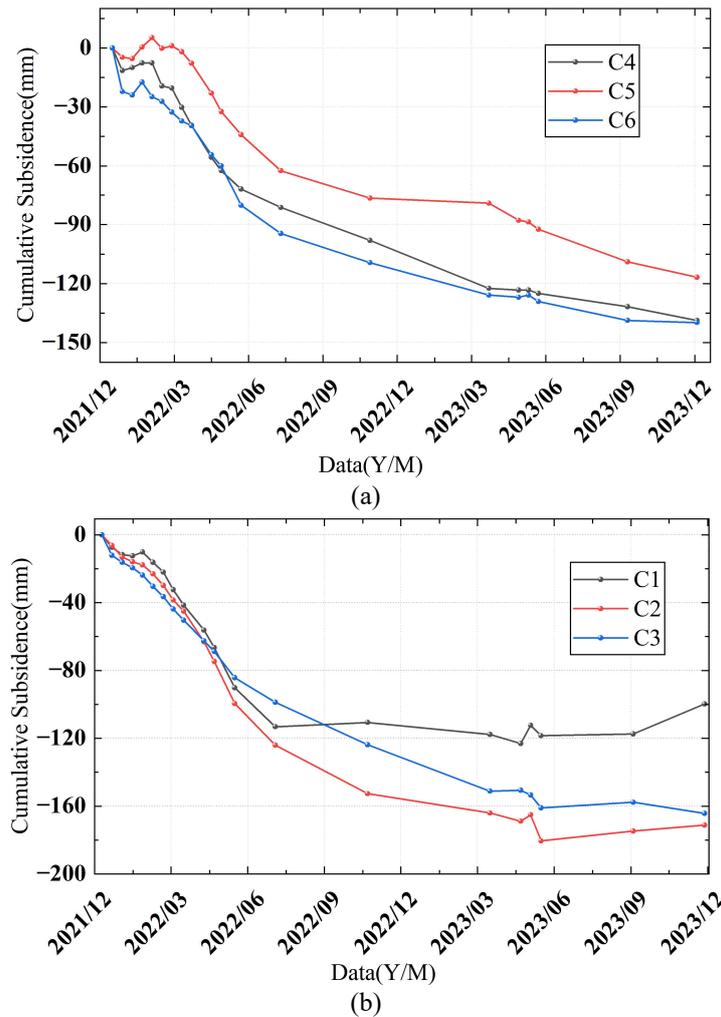
Figure 5 Cumulative Deformation Curve

Table 2 Deformation Information at Monitoring Points

Monitoring Points	Mean Deformation Rate (mm/a)	Cumulative Displacement (mm)	Maximum Displacement (mm)
C1	-51.9	-99.6	-123.1
C2	-89.3	-171.1	-180.4
C3	-85.7	-164.3	-165.8
C4	-72.3	-138.6	-138.6
C5	-60.8	-116.6	-116.6
C6	-72.9	-139.7	-140.1

5 CONCLUSIONS

This study employed time-series InSAR techniques to analyze 20 Sentinel-1 SAR images (2022-2023) covering the Xinjing Mining Area, deriving a mean deformation velocity field and revealing its spatiotemporal evolution patterns. The reliability of deformation monitoring results was cross-validated using optical remote sensing imagery. Key conclusions are as follows：

(1) Areas with deformation rates of -10 to -30 mm/a account for 60.6% of the total subsidence zone, predominantly distributed along the periphery of the main deformation region. Rapid deformation zones (rates < -30 mm/a) cluster centrally, consistent with the "downward excavation and progressive convergence" mining pattern of open-pit operations, confirming data validity.

(2) From January 2022 to December 2023, surface deformation in the Xinjing Mining Area progressed through three distinct phases:The first phase, spanning January to June 2022, was characterized by accelerated deformation, during which intense ground displacement occurred and the majority of cumulative subsidence in the mining area was generated. Subsequently, from July 2022 to March 2023, deformation entered a moderate phase marked by reduced rates of both subsidence velocity and cumulative displacement accumulation, significantly differing from the preceding period. The final stabilization phase, lasting from April to December 2023, exhibited minimal deformation magnitudes, including localized uplift anomalies in cumulative displacement. These stabilization trends are attributed to the successful implementation of post-landslide mitigation measures, including operational suspension and slope

reinforcement engineering, which effectively suppressed further deformation progression.

(3) InSAR technology effectively monitors mining-induced surface deformation, providing critical data support for hazard prevention and serving as a vital bridge between geohazard mechanism research and engineering practices. Its application significantly enhances mine safety management and disaster mitigation capabilities.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] T Muhammad, G Ni, Z Chen, et al. Addressing resource curse: How mineral resources influence industrial structure dynamics of the BRI 57 oil-exporting countries. Resources Policy, 2024, 99: 105420.

[2] Z Luo, L Zhou, J Qin, et al. Assessing surface deformation in the Chengdu Plain: A comprehensive time-series InSAR study of urban development and natural environmental factors. Advances in Space Research, 2024, 73(3): 1780–1798.

[3] J Geng, X Zhang, J Lin, et al. 3S technology and its application in the assessment and monitoring of mining geological disasters. Digital Technology and Application, 2013, (04): 93–95.

[4] H Kawabata, S Yoshioka. Spatiotemporal changes in the strain field in the Niigata-Kobe Tectonic Zone: long-term analysis using GNSS time series data. Progress in Earth and Planetary Science, 2025, 12(1): 22–22.

[5] M D Lucio, C Riccardo, C Valentina, et al. Integrated approaches for field mapping by traditional investigation and satellite PSInSAR data: Results from the Montemartano Landslide (Central Italy). Remote Sensing, 2023, 15(5): 1221–1221.

[6] K Zhang, Q Li, H Dai, et al. Research on the integrated monitoring technology of "air and sky" for surface movement in mining area. Coal Science and Technology, 2020, 48(02): 207–213.

[7] J Huang, B Wang, X Cai, et al. Coastal reclamation embankment deformation: Dynamic monitoring and future trend prediction using multi-temporal InSAR technology in Funing Bay, China. Remote Sensing, 2024, 16(22): 4320–4320.

[8] D Zhou, X Zuo. Surface deformation monitoring and prediction in mining area based on SBAS-InSAR and PSO-BP neural network algorithm. Journal of Yunnan University (Natural Science Edition), 2021, 43(05): 895–905.

[9] E Bayramov, N Sydyk, S Nurakynov, et al. Quantitative assessment of urban surface deformation risks from tectonic and seismic activities using multitemporal microwave satellite remote sensing: A case study of Almaty city and its surroundings in Kazakhstan. Frontiers in Built Environment, 2024, 10: 1502403.

[10] Q Fan, C Tang, Y Chen, et al. Application of GPS and InSAR technology in landslide monitoring. Surveying and Mapping Science, 2006(05): 60–62+5.

[11] K E, A N, F H, et al. Monitoring of land surface displacement based on SBAS-InSAR time-series and GIS techniques: A case study over the Shiraz metropolis, Iran. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2023, X-4/W1: 371–378.

[12] R Wang, S Xiong, H Nie, et al. Research on remote sensing geological survey technology and application. Journal of Geology, 2011, 85(11): 1699–1743.

[13] M Manunta, D C Luca, I Zinno, et al. The Parallel SBAS approach for Sentinel-1 interferometric wide swath deformation time-series generation. Algorithm description and products quality assessment. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(9): 6259–6281.

[14] H Yin, J Zhu, C Li, et al. Research on deformation monitoring of mining area based on SBAS. Journal of Surveying and Mapping, 2011, 40(01): 52–58.

[15] Q Yang, P Xu, C Cao, et al. Research on land subsidence prediction in mining areas based on SBAS-InSAR and multi-model comparison. Journal of Research in Science and Engineering, 2024, 6(12): 63–71.

[16] A Mugabushaka, Z Li, X Zhang, et al. Mapping surface deformation in Rwanda and neighboring areas using SBAS-InSAR. Remote Sensing, 2024, 16(23): 4456–4456.

[17] W Shen. Mindful of people's well-being, dedication to escort a better life. China Mining News, 2024(001).

[18] F Ma, R Li, X Liu, et al. Surface settlement monitoring of Xinjing coal mine based on time series InSAR technology. Journal of Shanxi Datong University (Natural Science Edition), 2023, 39(04): 123–127.

AN ECOSYSTEM-SERVICE EVALUATION METHOD OF A RIVER BASIN WITH RAPID CHANGES IN LAND USE WITHIN A SHORT TERM

HeSheng Zhang, MengHao Li, ZiYao Xing, YuLing Zhang[*], Xin Zan, XiaoYu Jia
Hebei Key Laboratory of New Energy Environmental Safety and Resource Utilization, Department of Environmental Science and Engineering, North China Electric Power University, Baoding 071003, Hebei, China.
Corresponding Author: YuLing Zhang, Email: zhangyuling_hit@163.com

Abstract: In 2015, Zhangjiakou City was selected as the co-host city for the 2022 Winter Olympics. In order to improve ecosystem service functions, land use types in the region were changed greatly within the short term. Therefore, existing evaluation methods of the ecosystem service value (ESV) are not applicable to such an underdeveloped region. With the year of 2015 as the key time node, a dynamic process evaluation method of ESV was established for Zhangjiakou City based on classic equivalent factors in this study. The five economic coefficients of net primary productivity (NPP) coefficient, precipitation adjustment coefficient, soil conservation adjustment coefficient, habitat adjustment coefficient and accessibility adjustment coefficient were introduced into the classic method to correct the commonly Chinese equivalent factors to obtain the actual ESV. Firstly, the results obtained from the dynamic process evaluation method could clearly distinguish the changing trend of ecosystem service value (ESV) after the change of land use type. Secondly, the ESV increment calculated with the dynamic process evaluation method was more sensitive to land use change. The change in the forest area proportion was the most significant and reached 6.8%, and corresponding ESV increment was 14.89%. Thirdly, from the perspectives of 11 sub-categories of ecosystem services, the ESV of hydrological regulation was the largest and reached 30%. The ESVs of food production, raw material production, maintenance of nutrient cycle, water supply and aesthetic value supply were relatively small. From the sub-categories of enhancing ecosystem services, the ESV of each sub-category increased after changing land use type. Fourthly, after the introduction of five economic adjustment coefficients, the ESVs of the other four categories of ecosystem services, except the ESV of biodiversity, increased significantly after the change in the land use types. The results indicated that the dynamic evaluation method could better identify the sensitive factors affecting the value of regional ecosystem services.

Keywords: Dynamic process evaluation; Rapid change in land use types; Ecosystem service value; Zhangjiakou City

1 INTRODUCTION

Ecosystem service refers to the products and services that ecosystem provides for human beings through ecological structures, functions and processes [1]. In recent years, due to rapid urban expansion and rapid economic development, human beings have utilized ecosystem services excessively, thus leading to ecosystem degradation [2]. The evaluation of ecosystem service value (ESV) is the foundation of ecosystem protection and management [3]. However, a widely accepted evaluation system of ESV has not yet been developed at home and abroad. The differences in evaluation methods also lead to the significant differences in evaluation results, thus limiting the objective knowledge of ecosystem service functions and their values.

Two types of ESV evaluation methods are currently available, namely, the functional value evaluation method and the equivalent factor evaluation method. In the functional value evaluation method, some economic methods are used to convert the mass of different materials into a unified monetary value [4,5]. This method can accurately measure the size of certain service functions in a region. However, the evaluation of different service functions often requires different ecological equations and parameters and the calculation process is more complicated [6]. Therefore, the method is mostly applied in the regions with developed economy and comprehensive basic data. In addition, when this evaluation method is used, it is necessary to choose the service function to be evaluated and set corresponding parameters, thus leading to the great uncertainty in evaluation results and increasing the difficulty in the comparison among the evaluation results of various ecosystems [7]. The equivalent factor evaluation method was first proposed by Costanza et al. [4]. According to the proposed method, different ecosystem service functions are divided into different categories and the equivalent values are obtained based on the meta-analysis results. Then the regional ESV is obtained with the area of each ecosystem. Compared with the functional value evaluation method, the equivalent factor evaluation method can evaluate the ESV in a large area more effectively and is widely used. Mendoza-González et al. introduced the economic coefficient to evaluate the ESV along the Gulf Coast [8]. Kreuter et al. investigated the impact of urban expansion on ESV in Bexar County from 1976 to 1991 by combining land exchange analysis with the ecosystem service value factor provided by Costanza et al. [9]. Yi et al. evaluated the ecosystem services of the San Antonio River Basin (SARB) in the United States based on land use change by modifying the global equivalent factors. Richardson et al. believed that only the immediate ESV value transfer factor could provide meaningful value information to be transformed [10,11]. Yi et al. pointed out that before the adoption of Costanza's value factor, the economic coefficient

should be corrected [10]. The evaluation results of the equivalent factor method were not reliable if the equivalent factor could not accurately reflect the ecological background of study areas [12]. After Costanza et al. first expressed ecosystem services in monetary units, Silvestri et al. [13] evaluated ESVs at different scales, such as global scale [4], regional scale [5], watershed [14], coastal zones [15], and individual ecosystems, such as wetlands [16], forestland and highway [17]. However, when the unit value for the global ecosystem proposed by Costanza et al. was applied in other regions, individual differences were ignored, thus resulting in large errors. In order to solve the current situation of ESV evaluation in China, Xie et al. conducted a questionnaire survey on more than 700 ecologists based on the evaluation model proposed by Costanza and others and obtained the "China Ecosystem Service Value Equivalent Factor Table" [5]. The values of mutual contributions between ecosystem service functions and the economic values of farmland food production services are considered in the table and it is more comprehensive and more targeted than previous evaluation methods. The method developed by Xie et al. solved the conversion problem of global equivalents per unit area in China and had been applied in ESV evaluation of various cases from a regional scale to a national scale [18,19]. However, although the equivalent factor method of Xie et al. solved the regional difference between the world and China, the method evaluated the ESV of the static ecological environment and belonged a static evaluation method. This method could not achieve an accurate evaluation of the ESV of the changing ecological environment. In order to accurately reflect the temporal and spatial changes of the ESV, most scholars currently only provided corrections for the values of food production per unit farmland area. The simple correction was obviously too monotonous and not consistent with the reality. Based on the equivalent factors proposed by Xie et al., five economic indicators were introduced to identify the key factors affecting each service function, and then a spatial heterogeneity evaluation model was established which could reflect the differences in service functions. The model provides a more comprehensive and objective evaluation methodology for rapidly changed land utilization patterns.

This study is based on the ecological status of Zhangjiakou City, which is located in the northwest of Beijing. Zhangjiakou City was originally an economically underdeveloped region. The opportunity of 2022 Winter Olympics promoted its rapid economic development and rapid expansion and the ecosystem services and functions were severely challenged. In order to enhance the ecological services for the 2022 Winter Olympics, local land utilization pattern was rapidly transformed within the short term. Therefore, it is necessary to establish a process ESV evaluation method considering land use change. It can calculate the value of ecosystem service value after land use change, analyze the important factors affecting ESV increment, and guide the conversion of land use type toward ESV increment direction. It provides a new idea for the restoration of environmental ecological systems and the sustainable green development in the region.

2 STUDY AREA AND DATA SOURCES

2.1 Study Area

Zhangjiakou City is located at the junction of Beijing, Hebei, Shanxi, and Inner Mongolia and in the northwestern part of Hebei Province. It is the meeting point of the Beijing-Tianjin-Hebei Economic Circle. Zhangjiakou City has 6 districts (Qiaodong District, Qiaoxi District, Xuanhua District, Xiahua District, Wanquan District, and Chongli District) and 10 counties (Zhangbei County, Kangbao County, Wuyuan County, Shangyi County, Yu County, Yangyuan County, Huai'an County, Huailai County, Zhuolu County, and Chicheng County). Its total area is about 36800 km2 (113°50′~116°30′ E, 39°30′~42°10′ N). In 2015, Zhangjiakou City was officially confirmed as one of the venues of the 24th Winter Olympic Games. Then, regional land use types and their spatial structures have also undergone significant changes, which in turn affected the ESV in the region.

2.2 Data Sources

The land use distribution data from 2013 to 2017 were provided by the Land and Resources Bureau of Zhangjiakou City. The land types in the study area included cultivated land, forest land, grassland, lakes/rivers and unused land. Due to the small area of wetlands in Zhangjiakou City, wetlands were not considered in the study. The data of the types and prices of foods were provided by the Zhangjiakou Municipal Development and Reform Commission and the meteorological data were from the China Meteorological Data Network. Other data were from China Statistical Yearbook and Zhangjiakou City Economic Statistics Yearbook from 2013 to 2017.

3 METHODS

In this study, the ESV change caused by the land use change in Zhangjiakou City was calculated with the unit area value equivalent factor method based on dynamic correction. Firstly, we used the Millennium Ecosystem Evaluation (MA) method as a baseline for the classification of ecosystem services and divided ecosystem services into 11 ESVs. Then, based on previous studies [20-22], five eco-economic indicators were selected to spatially correct the 11 ESVs and the annual ESV changes in the two stages were compared in order to explore the impact of venue announcement of the Winter Olympic Games and the applicability of the process evaluation method in Zhangjiakou City. The plant net primary productivity index (NPP) is generally positively correlated with food production, raw material production, gas regulation, climate regulation, the maintenance of circulating nutrients and environment purification [6]. The precipitation adjustment coefficient could measure the actual level of water supply and hydrological regulation in the

study area. The soil conservation spatial-temporal adjustment coefficient was used to regulate the soil conservation function in the study area. The habitat regulation coefficient could modify the biodiversity value according to the resistance encountered by the organism during migration. The accessibility adjustment coefficient could correct the cultural aesthetic service value according to the recreation opportunity spectrum theory. Finally, with the dynamically corrected equivalent factors, the spatial distribution of the ESV of Zhangjiakou City was calculated from the perspectives of land use and 11 ESV subcategories.

3.1 Division of Study Stages

In 2015, Zhangjiakou City was determined as one of the venues for the 2022 Winter Olympics. The land use type change in Zhangjiakou City is shown in Fig. 1. The land use situation was significantly changed from 2015 to 2017. Therefore, we divided the study period from 2013 to 2017 into two stages (one stage from 2013 to 2015 and the other stage from 2015 to 2017), compared the annual ESV changes in the two stages, and further explored the impact of the preparation for the Winter Olympic and the applicability of the process evaluation method in Zhangjiakou City.

3.2 Adjustment Coefficient

The correction of the equivalent factor in this study is based on the equivalent factor table proposed by Xie et al. in 2015. In this study, based on the previous study [12], five key ecological economic indicators were selected to spatially correct 11 ecological service functions as follows:

$$F_{ij} = F_j \times E_{ij} \quad (j=1,2\ldots5) \tag{1}$$

where Fij refers to the equivalent factor of the j-th type of service function of a certain ecosystem in the i-th region; Fj refers to the national average equivalent factor of the ecosystem (where the first category of service functions include food production, raw material production, gas regulation, climate regulation, maintenance of nutrient cycling and environment purification; Category 2 includes water supply and hydrological regulation; Category 3 refers to biodiversity; Category 4 refers to soil conservation; Category 5 refers to aesthetic landscape); Eij refers to the spatial regulation factor of the ecosystem in the i-th region for the j-th service function (Ei1 to Ei5 correspond to five types of service functions, respectively referred to as NPP regulatory factors, rainfall adjustment factors, habitat quality adjustment factors, soil erosion degree adjustment factors and accessibility adjustment factor).

3.2.1 NPP adjustment coefficient
Supply service refers to products obtained from ecosystems, such as foods, fuels, fibers, freshwater, and genetic resources. These service functions are generally positively correlated with biomass. The net primary productivity index (NPP) reflects the efficiency of carbon fixation and light energy conversion. NPP directly reflects the production capacity of vegetation community in natural environment and is the main factor regulating the terrestrial ecological process. The NPP regulation factor is calculated as follows:

$$E_{i1} = N_{iv} / \overline{N} \tag{2}$$

where Niv refers to the NPP (t•hm-2) of the v-th month in the i-th region of this ecosystem; $\overline{N}$ represents the national annual average NPP (t•hm-2) of this ecosystem.

3.2.2 Precipitation adjustment coefficient
The water supply and hydrological regulation are related to the change in precipitation. The precipitation regulation factors are as follows:

$$E_{i2} = H_{iv} / \overline{H} \tag{3}$$

where Hiv is the average unit area precipitation (mm•hm-2) in the v-th month of the i-th region; $\overline{H}$ is the average unit area precipitation (mm•hm-2) in the whole country. According to the precipitation data of each meteorological station in Zhangjiakou from 2013 to 2017, the monthly average precipitation of each region is calculated by Kriging interpolation in ArcGIS10.2 and the precipitation regulating factor is obtained with Eq. (3).

3.2.3 Soil conservation adjustment coefficient
Based on the most widely used soil loss equation for estimating soil erosion, the adjustment factors of spatio-temporal variations of soil conservation equivalent factors in different regions are calculated as follows:

$$E_{i3} = E_{iv} / \overline{E} \tag{4}$$

where Eiv is the simulation amount of soil conservation in the v-th month of the i-th region; $\overline{E}$ is the average simulation quantity of soil conservation per unit area in China. In this study, based on the simulation quantity of soil conservation calculated with the general soil loss equation, the spatiotemporal adjustment factors of soil conservation were obtained.

3.2.4 Habit quality adjustment coefficient
Creatures encounter different resistances when they migrate between different land types. The greater the resistance is, the greater the migration difficulty is. The resistance leads to poor habitat quality and affects regional biodiversity levels [23]. Forest land is the most suitable habitat for terrestrial organisms. Therefore, in this study, forest lands were

extracted from land use data as the ecological source, whereas other land types were set as cost grids. Cost parameters were set according to the report by Li. et al [24]. The habitat quality adjustment factor is calculated as follows:

$$E_{i4} = 1 / \left(\frac{R_i}{\overline{R}_i} \right)$$

(5)

where R_i is the average resistance of the i-th region and $\overline{R}_i$ is the national average resistance.

3.2.5 Accessibility adjustment coefficient

According to the theory of Recreation Opportunity Spectrum (ROS) [21], cultural service function is determined by two key factors, namely, entertainment potential index (RPI) and recreation accessibility index. The equivalent factors of cultural services reflect the RPI of different land ecosystems and the accessibility determines the actual supply level of cultural services in these ecosystems. The accessibility adjustment factor is calculated as follows:

$$E_{i5} = RD_i / \overline{RD}_i$$

(6)

where RD_i is the average road density (km/hm2) of the i-th region; $\overline{RD}_i$ is the national average road density.

3.3 ESV Per Unit Area in Zhangjiakou

In the ESV equivalent factor table revised by Xie et al. in 2015 [6], the economic value of annual natural grain output of the farmland with an area of 1 hm2 is defined as 1 and the ESV equivalent factors of other ecosystems are the contribution of corresponding ecosystems to the food production and services of the farmland. The weighted average of the net profit of corn, naked oats, potatoes and other food crops was treated as the standard equivalent factor value. Due to the fluctuation of planting area and grain price caused by natural and market factors, the average value of net grain profit from 2013 to 2017 was taken as the economic value of standard equivalent factor in this study. The equivalent factors in Zhangjiakou City are shown in Table 1. The economic value of the standard equivalent factor in this study is calculated as follows:

$$E_a = 1/7 \sum_{i=1}^{n} \frac{m_j p_j q_j}{M} \quad (i=1,2,3..., n)$$

(7)

$$P_{ij} = F_{ij} \times E_a$$

(8)

where Ea indicates food production value per unit area of farmland (RMB/hm2); i is the crop type, pi is the average price of the i–th crop (RMB / t); qi is the unit yield of the i–th crop (t/hm2); mi is the area of the i–th crop (hm2); m is the total area of n crops (hm2); Pij is the unit area value (RMB/(hm2·a)) of the j-th ecosystem services in the i-th region; F_{ij} is the unit area value equivalent of the j-th ecosystem in the i-th region.

3.4 Dynamic Evaluation Model of Zhangjiakou's ESV

ESV is calculated as follows:

$$ESV = \sum S_j \times P_j$$

(9)

where ESV is the total value of regional ecosystem services (RMB); S_j is the area of the j-th ecosystem (hm2).

4 RESULTS

4.1 ESVs of Different Land-Use Types

According to the dynamic evaluation method, the ESVs of various land use types in Zhangjiakou City from 2013 to 2017 were evaluated with Eq. (9). The ESV of Zhangjiakou City firstly decreased from 2013 to 2015 and then increased from 2015 to 2017 (Table 2). In 2017, the total ESV of Zhangjiakou City reached ¥ 242.22 billion, which was ¥ 5.06 billion higher than that of 2015, indicating that the process evaluation method could detect the total ESV change in this area. Fig. 2 shows the ESVs of various land use types. The ESV differences among different types of lands are significant (Fig. 2). In the study period, forests, lakes and grasslands contributed the majority to the entire ESV, about 90%. Compared to forests, lakes and grasslands, the other two types of land use (farmland and unused land) provided relatively small ESVs. Among them, the ESV ratio of cultivated land decreased firstly and then increased, and the ESV of unused land decreased year by year. At the same time, the process evaluation method could identify the sub-category changes of ESV in the region. The area increments of cultivated lands, forest, lake/river and grasslands in the second stage were increased compared with those in the first stage (Fig. 3). The area increment of forest land was 6.26% and corresponding ESV increment was 14.89%. The area increment of cultivated land was 5.60% and corresponding ESV increment was 14.51%. The area of unused land was decreased by 0.7% and corresponding ESV increment was -8.37%. The above results indicated that the ESV calculated with the process evaluation method was sensitive to the change in land use type.

4.2 ESVs of Different Sub-Categories

The ESVs of different sub-categories of ecosystem services are shown in Table 3. ESV proportions of hydrological regulation and climate regulation were the largest and respectively reached 30% and 20%. The ESV proportions of gas regulation, waste treatment, biodiversity maintenance, soil conservation were more than 5% and the ESV proportions for the other five sub-categories of food production, the production of raw materials, maintenance of nutrient cycling, water supply and providing aesthetic values were all less than 5%. The hydrological regulation service, which was mainly provided by forestland and lakes/rivers, accounted for the largest proportion of 32.55% in 2017 (Fig. 4). The services of climate regulation, gas regulation, waste treatment, biodiversity maintenance and soil conservation, which were mainly ascribed to forestland, accounted for 55% of the total value. However, the ESVs of the services of food production, raw material production, maintenance of nutrient cycling, water supply and providing aesthetic values shared the smaller proportion. The results of the sub-category of the process growth of ecosystem services are shown in Fig. 5. The ESV growth rate of each sub-category in the period from 2015 to 2017 was higher than that in the period from 2013 to 2015 (Fig. 5), especially the two ecological functions (food production; soil conservation). The ESV growth rates of these two functions from 2013 to 2015 were only - 0.14% and - 0.42%, whereas the ESV growth rates of the two functions from 2015 to 2017 were respectively 4.51% and 5.27%. The above results showed that various ecological transformation policies for the Winter Olympics increased the total ESV and the growth rate of multiple sub-categories. The process evaluation method could accurately identify the changes.

4.3 Influences of Economic Adjustment Coefficients on Different Types of ESVs

The functions of food production, raw material production, waste treatment, climate regulation, gas regulation and nutrient cycling were changed significantly after the correction (Fig. 6) because the introduction of NPP significantly improved the equivalent factors of the service value of the first category of ecosystems. The values of ecosystem services, such as water supply and hydrological regulation, were respectively increased by ¥ $1.1×10^7$ RMB and ¥ $3.6×10^8$ after the correction, indicating that the values of ecosystem services in the second category was significantly affected by precipitation. Due to the serious soil erosion in Zhangjiakou, before the ecological policy of returning farmland to forest was adopted, the habitat was damaged seriously and the migration process of animals was damaged. No significant change was observed in the study period and the value of biodiversity was reduced by ¥ $1.7×10^8$. The growth rate of soil conservation value was 5.23% after the correction (Fig. 6) because the change of soil conservation value in Zhangjiakou City could be accurately identified after adjusting the service value of such ecosystem with soil erosion factors. The providing aesthetic value was significantly changed after the correction because the actual supply level of landscape aesthetics in Zhangjiakou City was significantly improved after the introduction of accessibility factors. In conclusion, the five economic adjustment coefficients used in this paper could identify the sensitive factors that affect ESV in this area. The method could provide the more targeted results.

5 DISCUSSION

The ecosystem service value of Zhangjiakou City was obtained with the process evaluation method before and after Zhangjiakou City was determined as the venue of the 2022 Winter Olympic Games. According to the detailed composition of ecosystem service functions, the advantages and influencing factors of the process evaluation method are summarized below.
Firstly, in the process evaluation method, NPP was selected to correct the equivalent factor of the first category of service functions because the net primary productivity of vegetation could measure the impact of the regional land use change process on vegetation with a unified scale. Ecosystem services such as climate regulation, gas regulation and nutrient cycling were closely related to forest covering. The ecosystem service functions such as climate regulation, gas regulation and nutrient cycling were increased significantly in this area, indicating that the process evaluation method could accurately identify the spatio-temporal regulation factors, and quantitatively reflect the ESV changes of the above ecosystem service functions during the land use transformation period.
Secondly, the functions of water supply and hydrological regulation were strongly affected by precipitation. In the study, precipitation regulation factors were used to modify the service value of the second category of ecosystems. The precipitation of Zhangjiakou City from 2013 to 2017 was slightly higher than that of the whole country, and the ESV increment of water supply and hydrological regulation were respectively 1.68% and 0.95%, indicating that the process evaluation method could identify the influencing factors of ecosystem service functions more accurately than the uncorrected equivalent factor evaluation method.
Thirdly, soil conservation was affected by natural soil erosion, precipitation, vegetation coverage and other factors. Soil erosion coefficient was affected by the above factors and could be used to measure soil conservation capacity. The process dynamic evaluation method modified by soil erosion factors indicated that the soil conservation value of Zhangjiakou City was obviously affected by land use change. Soil conservation value accounted for about 10% of the total ESV and increased year by year, indicating that the soil conservation value in Zhangjiakou City could be dynamically identified by the equivalent factor modified with soil erosion factor. In addition, the soil conservation value was increased by 5.23% compared with the evaluation result obtained with the national equivalent factor, indicating that the process evaluation method could accurately identify the influencing factors of soil conservation function.

Fourthly, the values of ecosystem services, such as food production, raw material production and environment purification, were closely related to the economic level and the technological development level. The proportion of the primary industry in Zhangjiakou was higher and the proportions of secondary industry and tertiary industry were lower. The ecological preparation for the Winter Olympic Games greatly promoted the optimization of economic structure and rapid economic development. The values of food production and raw material production also increased significantly with the growth of economic level (Fig. 4). Zhangjiakou City, due to the 2022 Olympic Games, promoted the establishment of diversified clean energy supply systems such as wind power, solar photovoltaic, clean coal power, biomass energy and geothermal energy, accelerated the establishment of a clean and efficient Olympic energy supply system, and significantly improved the environment purification ability.

Fifthly, aesthetic value reflects the actual potential of landscape aesthetics and the accessibility factor reflects the supply level of landscape aesthetics by measuring the convenience of a certain location to a designated location. The more convenient the transportation is, the stronger the supply functions of culture and entertainments are. The aesthetic value of Zhangjiakou was significantly affected by accessibility factors. The aesthetic value provided by the region was increased by 1.21% from 2015 to 2017 (Fig. 3) and the increment was ¥ 1100×104, indicating that the process evaluation method could adjust the aesthetic value of the region with accessibility factors.

6 CONCLUSION

In underdeveloped areas experiencing dramatic changes of land use types within a short period, a suitable evaluation method of ecological values is not available. In the study, based on the consideration of NPP adjustment coefficient, rainfall adjustment coefficient, soil conservation adjustment coefficient, habitat adjustment coefficient and accessibility regulatory factor adjustment coefficient, a dynamic evaluation method of ecosystem service functions is proposed. Taking Zhangjiakou City as an example, the dynamic process evaluation method was used to calculate and evaluate the changes of ESV. During the entire study period, the ESV of Zhangjiakou City declined in the first phase (2013-2015) and then increased rapidly in the second phase (2015-2017). This change between the two phases clearly indicated that the supporting policy played an active role in protecting ecological zones. The ESV increment of soil conservation was 5.23% after the correction. The ESV increment of cultural aesthetic value was 3.42% and the ESV increments of water supply and hydrological adjustment were respectively 1.68% and 0.95%. The above results showed that the process evaluation method adopted in this study provided more details than the uncorrected evaluation method. From the perspective of the impact of land use on the ESV, the ESV increment was significantly sensitive to land use change. The change proportion of the forest area was the highest and reached 6.8% and corresponding ESV increment was 14.89%. The change proportion of cultivated land area was 5.60% and corresponding ESV increment was 14.51%. In addition, the five economic factor adjustment coefficients selected by the process evaluation method were suitable for analyzing the economically underdeveloped regions with rapid changes in land use types within a short period. The ESV assessment method is applicable to the regions with rapid changes in the land use pattern within a short period. In summary, the evaluation results were consistent with the rapid development of Zhangjiakou City in recent years, indicating that the evaluation results were reasonable and in line with the actual situation. The study provides an effective method for calculating the short-term rapid change of land use in the future. Therefore, Zhangjiakou City should seize the opportunity of the 2022 Winter Olympics to avoid post-Olympic effects and actively promote the reform of ecosystem management and regional green sustainable development. The dynamic process evaluation method also helps to understand or deeply study the dynamic influencing factors of ecosystems, and supports regional planning for the sustainable development of ecosystems.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

REFERENCES

[1] Costanza R, D'Arge R, De Groot R, et al. The value of the world's ecosystem services and natural capital. Nature. 1997, 387, 253-260.
[2] Wu J. Landscape sustainability science: ecosystem services and human well-being in changing landscapes. Landsc Ecol. 2013, 28, 999-1023.
[3] Lei J C, Wang S, Wang J M, et a;. Effects of land use change on ecosystem services value of Xunwu County. Acta Ecologica Sinica. 2019, 39, 3089-3099.
[4] Costanza R, De Groot R, Sutton P, et al. Changes in the global value of ecosystem services. Global Environmental Change. 2014, 26, 152-158.
[5] Xie G D, Lu C X, Leng Y F, et al. Ecological assets valuation of the Tibetan Plateau. Journal of Natural Resources, 2003, 18, 189-196.

[6]   Xie G D, Zhang C X, Zhang L M, et al. Improvement of the evaluation method for ecosystem service value based on per unit area. Journal of Natural Resources. 2015, 30, 1243-1254.

[7]   Yu Z Y, Bi H. Status quo of research on ecosystem services value in China and suggestions to future research. Energy Procedia. 2011, 5, 1044-1048.

[8]   Mendoza-González G, Martínez M L, Lithgow D, et al. Land use change and its effects on the value of ecosystem services along the coast of the Gulf of Mexico. Ecological Economics. 2012, 82, 23-32.

[9]   Kreuter U P, Harris H G, Matlock M D, et al. Change in ecosystem service values in the San Antonio area, Texas. Ecol. Econ. 2001, 39, 333-346.

[10] Yi H, Güneralp B, Anthony M F, et al. Impacts of Land Change on Ecosystem Services in the San Antonio River Basin, Texas, from 1984 to 2010. Ecological Economics. 2017, 135, 125-135.

[11] Richardson L, Loomis J, Kroeger T, et al. The role of benefit transfer in ecosystem service valuation. Ecological Economics. 2015, 115, 51-58.

[12] Xue M G, Xing L, Wang X Y. Spatial Correction and Evaluation of Ecosystem Services in China. China Land Science. 2018, 32, 81-88.

[13] Silvestri S, Zaibet L, Said M Y, et al. Valuing ecosystem services for conservation and development purposes: a case study from Kenya. Environmental Science Policy. 2013, 31, 23-33.

[14] Gren I M, Groth K H, Sylvén M. Economic values of Danube floodplains. Environ. Manage. 1995, 45, 333-345.

[15] Rao N S, Ghermandi A, Portela R, et al. Global values of coastal ecosystem services: a spatial economic analysis of shoreline protection values. Ecosystem Services. 2015, 11, 95-105.

[16] Sharma, B., Rasul, G., Chettri, N. The economic value of wetland ecosystem services: evidence from the Koshi Tappu Wildlife Reserve, Nepal. Ecosystem Service. 2015, 12, 84-93.

[17] Hejazi R, Shamsudin M N, Rahim K A, et al. Measuring the economic values of natural resources along a freeway: a contingent valuation method. Environ. Plann. Manage. 2012, 57, 629-641.

[18] Wang W, Guo H, Chuai X, et al. The impact of land use change on the temporospatial variations of ecosystems services value in China and an optimized land use solution. Environ. Sci. Policy. 2014, 44, 62-72.

[19] Zhao Y H, Zhang L L, Wang X F, et al. Assessment and spatiotemporal difference of ecosystem services value in Shaanxi Province. Chin. J. Appl. Ecol. 2011, 22, 2662-2672.

[20] He J, Yan Z, Wan Y. Trade-offs in ecosystem services based on a comprehensive regionalization method: a case study from an urbanization area in China. Environmental Earth Sciences. 2018, 77, 179.

[21] Sun X Crittenden J C, Li F, et al. Urban expansion simulation and the spatio-temporal changes of ecosystem services, a case study in Atlanta Metropolitan area, USA. Science of the Total Environment. 2018, 622, 974-987.

[22] Paracchini M L, Zulian G, Kopperoinen L, et al. Mapping cultural ecosystem services: a framework to assess the potential for outdoor recreation across the EU. Ecological Indicators. 2014, 45, 371-385.

[23] Adriaensen F, Chardon J P, De Blust G, et al. The application of 'least-cost' modelling as a functional landscape model. Landscape and urban planning. 2003, 64, 233-247.

[24] Li B, Chen D, Wu S, et al. Spatiotemporal assessment of urbanization impacts on ecosystem services: Case study of Nanjing City, China. Ecological Indicators. 2016, 71, 416-427.

CO-PYROLYSIS OF BIOMASS-BASED FURAN AND METHANOL TO LIGHT AROMATICS OVER MODIFIED HIERARCHICAL ZEOLITES

Qing Xu[1], JunMing Liang[1,2,3], DongYan Zhang[2,3], ShengPeng Xia[2,3], AnQing Zheng[2,3*], Kun Zhao[2,3], ZengLi Zhao[2,3]

[1]*School of Mechanical Engineering, Guangdong Ocean University, Zhanjiang 524088, Guangdong, China.*
[2]*Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, Guangzhou 510640, Guangdong, China.*
[3]*Guangdong Provincial Key Laboratory of High-Quality Recycling of End-of-Life New Energy Devices, Guangzhou 510640, Guangdong, China.*
Corresponding Author: AnQing Zheng, Email: zhengaq@ms.giec.ac.cn

Abstract: Light aromatics, including benzene, toluene, and xylene, are essential chemical feedstocks, traditionally produced via catalytic reforming of petroleum fractions. Due to the depletion of fossil resources, lignocellulose has attracted considerable interest as a renewable carbon source. Furans as key platform molecules can be catalyzed into aromatics over zeolites; however, their application is often limited by low selectivity and catalyst deactivation from coke. In this work, a co-pyrolysis approach using 2-methylfuran and methanol was developed to enhance aromatic production. Methanol served both as a heat-releasing agent during the aromatization process and as a hydrogen donor due to its high effective hydrogen-to-carbon ratio, promoting a synergistic effect with furan compounds, enhancing aromatization efficiency. Hierarchical HZSM-5 modified with Ga, Zn, Mo, and W were synthesized and characterized using various characterization techniques to correlate pore structure and acidity with product distribution. It is found that W-modified HZSM-5 can further improve the carbon yield of aromatics (70.62%), while other metal modifications resulted in decreased performance due to reduced acidity and pore blockage, accompanied by increased coke formation. This strategy provides new insight into the efficient catalytic conversion of biomass-derived platform molecules into high-value aromatics.
Keywords: 2-Methylfuran; Methanol; Zeolite; Co-pyrolysis; Light aromatic

1 INTRODUCTION

In the context of achieving carbon peaking and carbon neutrality goals, the depletion of fossil fuels and escalating environmental concerns have garnered increasing global attention. As a result, the search for clean and renewable energy sources has become a major research focus[1]. Biomass, with its capacity to be converted into high-value chemicals (e.g., sugars, fragrances, and aromatics) and liquid biofuels (e.g., biodiesel and ethanol), is widely regarded as one of the most promising clean energy sources and a key resource for realizing dual-carbon objectives.
Catalytic pyrolysis of biomass enables enhanced feedstock conversion, selective product formation, and targeted yield improvements through the action of catalysts. Catalysts can promote specific structural transformations during biomass pyrolysis, such as C-O bond cleavage or the removal of functional groups (e.g., hydroxyl, alkyl), thereby directing the reaction toward the formation of desired products and enhancing their yiedl[2-8].
Light aromatics are in high and growing demand worldwide, with the global consumption of benzene expected to exceed 107 million tonnes by 2035. Extensive research has demonstrated that co-pyrolysis of furan and olefins (such as ethylene and propylene) facilitates aromatization via the Diels–Alder (DA) reaction[9]. While olefins are highly reactive under ambient conditions, alcohols (e.g., methanol and ethanol) offer comparable functionality through methanol-to-olefins (MTO) pathways[10]. Moreover, as liquids at room temperature, alcohols are more stable and easier to store. Ivo and his co-workers have shown that replacing ethylene with ethanol significantly lowers the activation barrier, enhancing both the conversion of feedstocks and selectivity of aromatic[11]. Among alcohols, methanol—an essential building block in the chemical industry—achieved an annual production of 91.82 million tonnes in China by 2024, with a growth rate of 10.4%, substantially surpassing that of ethanol. Therefore, the co-pyrolysis of methanol with biomass-derived furans presents a viable alternative approach.
Zeolites are widely used in industrial applications due to their excellent acid-catalytic activity, thermal stability, and hydrothermal resistance. Catalyst deactivation is typically caused by low diffusion coefficients that hinder mass transfer of reactants and products within the pores, leading to pore blockage. Introducing hierarchical pore structures into zeolites helps mitigate this by retaining the shape-selectivity of micropores while improving mass transport through mesopores. Transition metals such as Fe, Cu, and Ni are frequently employed to further modify catalytic behavior, as metal incorporation can significantly alter the acidity and nature of acid sites. Incorporating metals like Mo, Zn, Ga, and W into hierarchical zeolites leverages the structural advantages of hierarchical pore structures to facilitate the aromatization of biomass-derived furans with methanol via co-pyrolysis[12-14].
This study proposes a simple and effective strategy for the efficient production of light aromatics via the co-pyrolysis of biomass-derived furans. The exothermic nature of methanol aromatization helps lower the energy barrier associated with furan conversion, while methanol's high effective hydrogen-to-carbon ratio (H/C$_{eff}$) provides an external hydrogen

source, resulting in synergistic enhancement. Coupling this with hierarchical zeolites modified by various transition metals allows for catalyst design. The influence of different metals on product distribution during co-pyrolysis is systematically investigated. A range of characterization techniques is employed to assess changes in surface area, pore structure, acid site type and distribution, and overall acidity. This work aims to elucidate the regulatory mechanisms by which metal-modified hierarchical pore structures influence pyrolysis product selectivity.

2 EXPERIMENTAL

2.1 Catalyst Synthesis

Preparation of hierarchical zeolite: Commercial HZSM-5 (Si/Al = 25) was treated with 0.3 M NaOH solution. 10 g of HZSM-5 was added to 100 ml of 0.3 M NaOH aqueous solution and stirred in a water bath at 80 °C for 1 h. After treatment, the mixture was immediately filtered using a vacuum filtration apparatus and thoroughly washed with deionized water until the filtrate reached a neutral pH. The resulting solid was dried in an oven at 105 °C for 12 h. To remove Na+ from the alkaline treatment, the solid was then ion-exchanged with 100 ml of 1 M NH4Cl solution at 80 °C for 4 h. The exchanged material was filtered, extensively washed with deionized water, and dried overnight. The dried sample was ground and calcined in a muffle furnace at 550 °C for 2 h. The resulting hierarchical zeolite was denoted as 0.3 ZSM.

Metal modification: Metal loading was carried out via excess impregnation. Appropriate amounts of metal nitrates were dissolved in 100 ml of deionized water to prepare precursor solutions with the desired concentration. Then, 10 g of 0.3 ZSM was added to the solution and stirred in a water bath at 60 °C for 8 h. The mixture was subsequently dried overnight in an oven and calcined at 550 °C for 4 h in a muffle furnace. Each metal-modified hierarchical zeolite was synthesized with a fixed metal loading of 2 wt% and denoted as M/ZSM, where M refers to the specific metal species. The resulting catalysts were pelletized, crushed, and sieved to obtain particles in the 40-60 mesh size range for use in subsequent fixed-bed experiments.

2.2 Catalyst Tests

2-methylfuran (2-MF) and methanol (MeOH) co-pyrolysis was conducted in a fixed-bed reactor. The reaction was carried out inside a quartz tube (inner diameter: 18 mm), which was fixed and heated using a tubular furnace. Before reactions, 1 g zeolite was loaded into the quartz tube. N2 was used as the purge gas at a flow rate of 120 ml/min, while Ar served as the carrier gas during the reaction at a flow rate of 60 ml/min. A syringe pump was employed to feed the liquid reactants (mixture of 2-MF:MeOH = 1:2). All reactions were performed under atmospheric pressure. Schematic diagram of reaction device is shown in Figure 1.

The detailed procedure is as follows: before the reaction, the reactor was purged with high-purity N2 at 200 ml/min to eliminate residual air from the fixed-bed system. Meanwhile, the tubular furnace was heated to the desired temperature by temperature control system. Once the target temperature was reached, Ar was introduced as the carrier gas to establish the reaction atmosphere. The mixture of 2-MF and MeOH was then introduced at a constant flow rate via the syringe pump.

The reaction products were divided into gas, liquid and solid. The liquid products were initially captured in a condenser immersed in ethanol maintained at –60 °C, and subsequently passed through a series of gas-washing bottles for further absorption and collection. All gas-washing bottles were placed in an ice-water bath to ensure a low-temperature environment. Gas productions were collected using gas sampling bags. After the reaction, the spent catalyst was recovered as the solid product. The liquid products collected from the condenser and washing bottles, then diluted to 250 ml, and stored for subsequent analysis and characterization.

2.3 Catalyst Characterization

X-ray diffraction (XRD) patterns were recorded using an X'Pert Pro MPD diffractometer to determine the crystallinity and elemental composition of the catalysts.

Ammonia-Temperature-Programmed Desorption (NH₃-TPD) was performed using to evaluate the total acidity and acid site distribution of the zeolite catalysts. NH₃ desorption was monitored by a thermal conductivity detector (TCD). The desorption was conducted by heating the sample from 90 °C to 750 °C at a rate of 10 °C/min.

N₂ adsorption-desorption analysis was conducted using an IQ-2 instrument to determine the specific surface area, total pore volume, average pore diameter, and pore size distribution. Prior to measurement, the samples were degassed. And followed by N₂ adsorption-desorption analysis under liquid N₂ conditions.

Pyridine-adsorbed Fourier-transform infrared spectroscopy (Py-FTIR) was carried out using a Nicolet IS50 spectrometer to identify and quantify Brønsted and Lewis acid sites in the zeolite. Pyridine adsorption was performed for semi-quantitative analysis, and the Brønsted-to-Lewis acid ratio (B/L) was calculated. Combined with the total acidity from NH₃-TPD results, the individual concentrations of Brønsted and Lewis acid sites were determined.
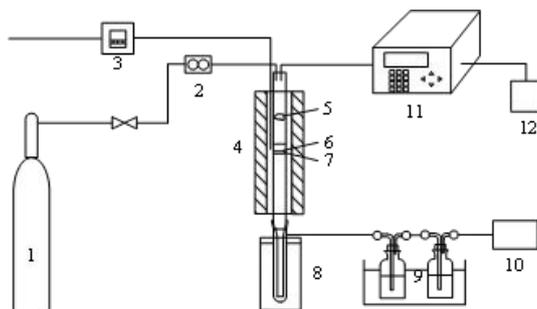
Figure 1 Schematic Diagram of Reaction Device

## 3 RESULTS AND ANALYSIS

3.1 Effect of Diverse Metal-Loaded Hierarchical Zeolites on Product Distribution

The specific reaction conditions were as follows: feedstock: a mixed solution of 2-MF and MeOH (molar ratio=1:2), reaction conditions: 550 °C, reaction pressure: 1 atm, WHSV: 4 h-1, carrier gas: Ar, carrier gas flow rate: 60 ml/min, reaction time: 15 min.

Figure 2 shows the mass spectra of liquid products obtained from the co-pyrolysis of 2-MF and MeOH over various metal-loaded hierarchical zeolites. As Figure 2, the distribution of products varies depending on the type of metal loaded. Nevertheless, under the pivotal role of the hierarchical pore structure all catalysts predominantly yield light aromatics, with small amounts of polycyclic aromatics also detected. The detailed retention times and identified compounds are summarized in Table 1.

As shown in Figure 3, the W-modified zeolite (W/ZSM) notably enhances the carbon yield of light aromatics (70.62%). In contrast, other metal-modified zeolites' aromatic yield decreased in different degrees. This performance deterioration can be attributed to two main factors, as supported by subsequent catalyst characterization: (i) the introduction of metal species reduces the number of acidic active sites on the catalyst surface, thereby diminishing overall acidity and catalytic efficiency; and (ii) metal oxide particles partially block the hierarchical pores, impairing mass transfer within the catalyst.
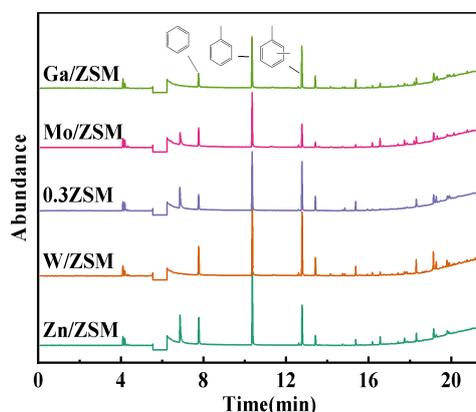
Figure 2 GC/MS Spectra of Catalytic Pyrolysis Reactions

Table 1 Main Productions of Catalytic Pyrolysis

Time	Substance
6.41	2-Methylfuran
6.88	Methoxyethane
7.77	Benzene
10.36	Toluene
12.77,12.62,13.42	Xylene
15.37	Mesitylene
16.18	Indane
16.56	indene
18.31	Naphthalene
19.15	2-Methylnaphthalene

Additionally, coke further support these findings. Compared with the unmodified hierarchical zeolite, the addition of Mo, W, or Zn significantly increases coke, with Mo/ZSM showing the highest coke of 15.61%. This excessive coke accumulation is primarily attributed tot pore blockage, which affects the mass transfer efficiency in the reaction.

Although W incorporation reduces the number of acidic active sites and partially disrupts the hierarchical pore structure, but the enhanced carbon yield of light aromatics indicates that W promotes aromatization. Moreover, both W and Ga modifications were found to suppress the formation of by-products such as methoxyethane, which is often promoted by the hierarchical pore structure. Additionally, the presence of Mo, W, and Zn appears to effectively activate C-H bonds and facilitate their cleavage, allowing residual carbon-containing fragments to recombine more freely which thereby increasing the yields of coke and polycyclic aromatics [15, 16]. This decomposition process also generates substantial amounts of low molecular weight gaseous products.

Table 2 summarizes the major gas products obtained from the co-pyrolysis of 2-MF and MeOH over different catalysts. The data clearly show that the introduction of Zn, Ga, and Mo significantly enhances hydrogen production, which is consistent with previous studies demonstrating that these metals promote the direct aromatization of methanol [17, 18]. Although direct methanol aromatization can increase aromatic yields to some extent, it bypasses the methanol-to-olefins (MTO) pathway, resulting in reduced olefin formation. This leads to limits the Diels-Alder (DA) reactions between 2-MF and olefins, ultimately reducing the overall yield of aromatics. In contrast, the W-modified zeolite facilitates the release of a larger amount of CO, thereby promoting the synergistic conversion of 2-MF and MeOH into light aromatics more effectively.
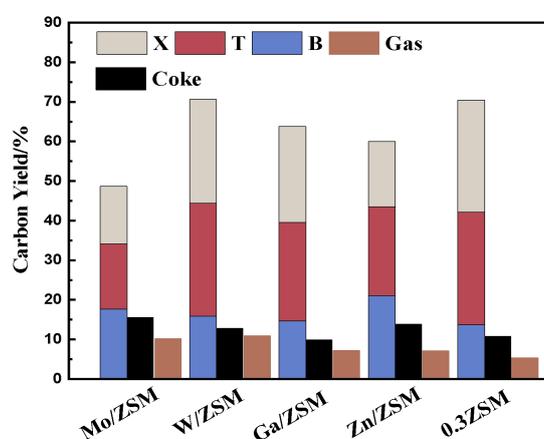
Figure 3 Carbon Yield of Different Concentration of Diverse Metal-Loaded Hierarchical Zeolites

Table 2 Gas Productions of Diverse Metal-Loaded Hierarchical Zeolites

Gas	Mo/ZSM	W/ZSM	Ga/ZSM	Zn/ZSM	0.3ZSM
CH_4	31.92%	26.24%	15.76%	11.36%	20.47%
C_2H_6	0.46%	0.77%	0.77%	0.28%	1.91%
C_2H_4	5.62%	14.14%	10.87%	4.86%	17.75%
C_3H_8	0.00%	0.28%	0.45%	0.06%	1.08%
C_3H_6	0.42%	1.35%	0.94%	0.39%	2.13%
CO_2	7.39%	5.98%	6.64%	7.87%	8.82%
CO	30.65%	41.39%	30.05%	22.28%	37.10%
H_2	23.54%	9.85%	34.53%	52.91%	12.56%

3.2 Characterization of Catalyst

The catalysts were characterized using XRD, BET, NH_3-TPD, and pyridine infrared (Py-IR) analyses.

X-ray diffraction (XRD) was employed to examine the crystallinity and elemental composition of the metal-modified hierarchical zeolites. As shown in Figure 4, all metal-loaded samples were based on the 0.3 ZSM. Characteristic diffraction peaks were observed at $2\theta = 8°$, 8.8°, 23.1°, 23.9°, and 24.4°, which are consistent with the MFI-type framework of HZSM-5. These results indicate that the incorporation of various transition metals and the presence of hierarchical pore structures did not alter the fundamental crystalline structure of the zeolite. Furthermore, the absence of distinct diffraction peaks corresponding to metal oxides suggests that the metals are either highly dispersed on the surface or incorporated into the zeolite pore channels. Combined with the results of BET, it can further prove that the

metals incorporated into the channels.

N_2 adsorption-desorption isotherms were used to characterized the specific surface area, pore volume and pore size distribution of the catalysts (Figure 5). The corresponding values are summarized in Table 3. It is noteworthy that among the four metal-modified zeolites, only the Ga-loaded sample exhibited an increase in surface area and pore volume. In contrast, the other three metal-modified samples showed decreases in both parameters. Given that micropores dominate the pore structure of zeolites and contribute significantly to surface area and pore volume, the observed reductions suggest that the introduction of transition metals during impregnation leads to partial blockage of micropores. The increase observed for Ga/ZSM is likely due to Ga-induced partial recrystallization of the zeolite framework, which alters the pore structure and mitigates micropores blockage.

Figure 5a displays the N_2 adsorption-desorption isotherms for the various catalysts, all of which exhibit type IV curves with H4-type hysteresis loops, indicating the formation of mesopores and capillary condensation of micropores [19]. As shown in Figure 5b, Ga/ZSM exhibits a marked increase in mesopore volume within the 2-10 nm range, it consistent with the results of BET which further supports the role of Ga in promoting framework recrystallization and mitigating micropores blockage. In contrast, Zn-loaded zeolite significantly reduced both the mesopores (2-10 nm) and macropores (30-50 nm), suggesting severe pore blockage due to agglomeration of Zn oxides within the channels. This conclusion is corroborated by the significantly reduced surface area and pore volume listed in Table 3 for Zn/ZSM. While Mo and W also led to decreases in surface area, they resulted in increased average pore diameter, indicating that these metals mainly block micropores while leaving mesopores and macropores relatively intact, thus minimizing mass transfer limitations.
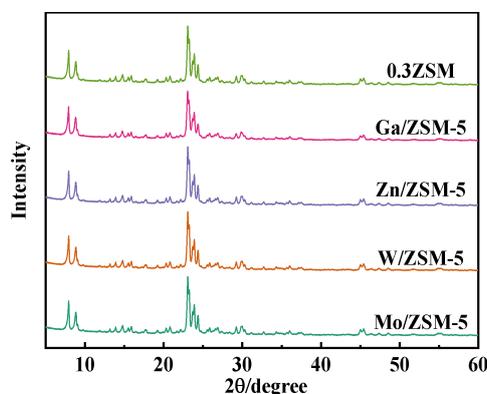
Figure 4 X-ray Diffraction Spectra of Diverse Metal-Loaded Hierarchical Zeolites
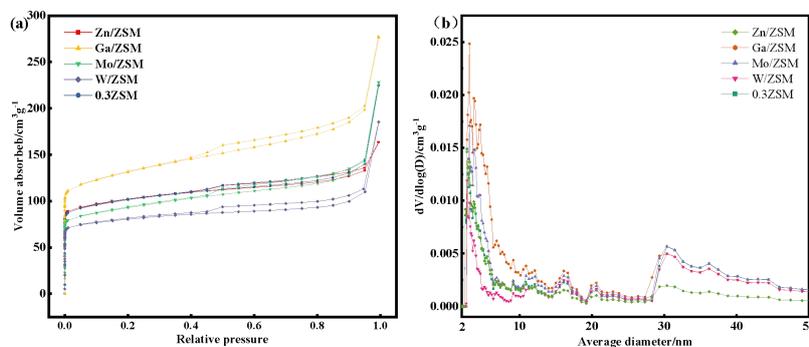
Figure 5 N_2 Adsorption and Desorption Isotherms (a) and Pore Size Analysis (b) of Diverse Metal-Loaded Hierarchical Zeolites

Table 3 Physical Properties of Diverse Metal-Loaded Hierarchical Zeolites

Catalysts	Ga/ZSM	Mo/ZSM	W/ZSM	Zn/ZSM	0.3ZSM
S_{BET}[1] (m^2/g)	481.742	343.4	308.202	384.781	379.705
V_{total}[2] (cm^3/g)	0.4288	0.3541	0.2875	0.2536	0.3489
Average Pore size[1] (nm)	3.56003	4.12443	3.73116	2.63647	3.67558

1 BET method, 2 Volume adsorbed at P/P$_0$=0.99
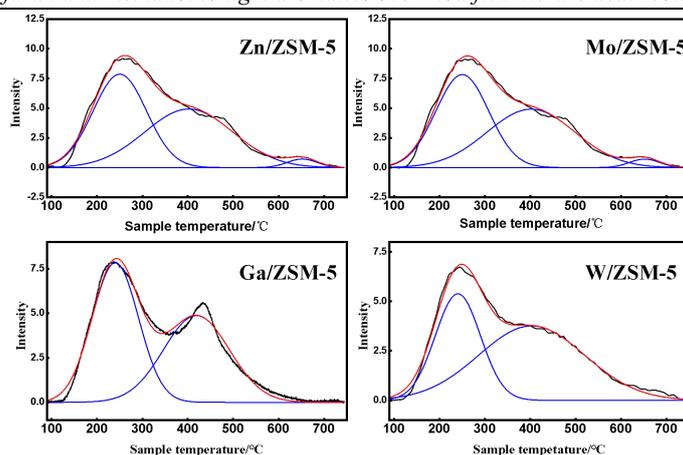
Figure 6 NH3-TPD Diagram of Diverse Metal-Loaded Hierarchical Zeolites

Table 4 NH3-TPD Results of Diverse Metal-Loaded Hierarchical Zeolites

| Catalysts	First peak (°C)	Second peak (°C)	Third peak (°C)	Total acid (umol/g)
Ga/ZSM	240.21	418.97	-	1425.8
W/ZSM	240.62	403.22	-	1328.8
Zn/ZSM	250.27	401.67	651.77	1678.4
Mo/ZSM	250.16	402.10	651.80	1749.8

NH$_3$-temperature programmed desorption (NH$_3$-TPD) was employed to characterize the strength and distribution of acid sites in the catalysts (See Table 4). The acidity profiles of various metal-modified hierarchical zeolites are shown in Figure 6. Two characteristic desorption peaks were observed: one in the low-temperature region around 250 °C, typically attributed to NH$_3$ desorption from weak acid sites, and another in the high-temperature region near 450 °C, generally corresponding to NH$_3$ desorption from strong acid sites [15]. Table 5 presents the fitted curves of NH$_3$-TPD and the corresponding total acidity values. The data indicate that the introduction of metal oxides weakens the strength of medium-to-strong acid sites, as evidenced by the diminished intensity of the corresponding peaks. Notably, a very weak desorption peak appeared near 650 °C for Mo/ZSM and Zn/ZSM, which is typically associated with NH3 desorption from strong acid sites. This phenomenon may result from a synergistic interaction between Mo or Zn with the zeolite, leading to NH$_3$ desorption at elevated temperatures.

Overall, these results demonstrate that the incorporation of transition metals significantly alters both the type and number of acid sites on the catalyst. For Ga- and W-loaded samples, the total acidity was found to be lower than that of the unmodified hierarchical zeolite, suggesting that Ga and W species may have entered the pore channels of the zeolite [15].

Figure 7 shows the FTIR spectra of the zeolites measured at 200 °C. The characteristic absorption peaks at 1545, 1454, and 1490 cm-1 respectively correspond to pyridine adsorbed on Brønsted acid sites, Lewis acid sites, and the combined B+L acid sites [20].
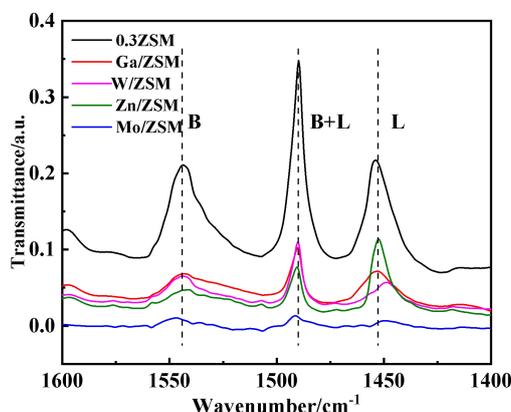
Figure 7 Py-IR Spectra of Diverse Metal-Loaded Hierarchical Zeolites

Table 5 Py-IR Acidity Distribution of Diverse Metal-Loaded Hierarchical Zeolites

Catalysts	Lewis (umol/g)	Brønsted (umol/g)	B/L	Total acid (umol/g)
0.3ZSM	635.7	896.3	1.41	1532.0
Mo/ZSM	438.0	1311.8	2.99	1749.8
Zn/ZSM	1271.2	407.2	0.33	1678.4
W/ZSM	437.5	891.3	2.04	1328.8
Ga/ZSM	684.5	741.3	1.08	1425.8

4 CONCLUSIONS

Based on the hierarchical zeolite, metal-modified catalysts were designed and applied to the co-pyrolysis of bio-based furans and methanol, resulting in a significant enhancement in aromatic yield. The results showed that tungsten (W) loading on hierarchical HZSM-5 further improved the catalytic performance, leveraging the advantages of the hierarchical pore structure, and achieved a light aromatics yield of 70.62%. In contrast, the other three metals led to a decline in overall acidity due to interactions with the acid sites of the zeolite. Moreover, partial pore blockage caused by certain metals restricted mass transfer of the reactants, thereby diminishing the catalytic activity and reducing aromatic yields.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] WANG S R, DAI G X, YANG H P, et al. Lignocellulosic biomass pyrolysis mechanism: A state-of-the-art review. Progress in Energy and Combustion Science, 2017, 62: 33-86. DOI: 10.1016/j.pecs.2017.05.004.

[2] BRIDGWATER A V. Review of fast pyrolysis of biomass and product upgrading. Biomass & Bioenergy, 2012, 38: 68-94. DOI: 10.1016/j.biombioe.2011.01.048.

[3] BUTLER E, DEVLIN G, MEIER D, et al. Fluidised bed pyrolysis of lignocellulosic biomasses and comparison of bio-oil and micropyrolyser pyrolysate by GC/MS-FID. Journal of Analytical and Applied Pyrolysis, 2013, 103: 96-101. DOI: 10.1016/j.jaap.2012.10.017.

[4] CHENG Y T, JAE J, SHI J, et al. Production of Renewable Aromatic Compounds by Catalytic Fast Pyrolysis of Lignocellulosic Biomass with Bifunctional Ga/ZSM-5 Catalysts. Angewandte Chemie-International Edition, 2012, 51(6): 1387-90. DOI: 10.1002/anie.201107390.

[5] ICHARD F, STEFAN C. Catalytic pyrolysis of biomass for biofuel production. Fuel Processing Technology, 2010, 91(1): 25-32.

[6] DHYANI V, BHASKAR T. A comprehensive review on the pyrolysis of lignocellulosic biomass. Renewable Energy, 2018, 129: 695-716. DOI: 10.1016/j.renene.2017.04.035.

[7] MISSON M, HARON R, FADHZIR M, et al. Pretreatment of empty palm fruit bunch for production of chemicals via catalytic pyrolysis. Bioresource Technology, 2009, 100(11): 2867-73. DOI: 10.1016/j.biortech.2008.12.060.

[8] SUN L Z, WANG Z B, CHEN L, et al. Catalytic Fast Pyrolysis of Biomass into Aromatic Hydrocarbons over Mo-Modified ZSM-5 Catalysts. Catalysts, 2020, 10(9). DOI: 10.3390/catal10091051.

[9] CHENG Y-T, HUBER G W. Production of targeted aromatics by using Diels–Alder classes of reactions with furans and olefins over ZSM-5. Green Chemistry, 2012, 14(11): 3114. DOI: 10.1039/c2gc35767d.

[10] RAHIMI N, KARIMZADEH R. Catalytic cracking of hydrocarbons over modified ZSM-5 zeolites to produce light olefins: A review. Applied Catalysis a-General, 2011, 398(1-2): 1-17. DOI: 10.1016/j.apcata.2011.03.009.

[11] TEIXEIRA I F, LO B T, KOSTETSKYY P, et al. From Biomass-Derived Furans to Aromatics with Ethanol over Zeolite. Angew Chem Int Ed Engl, 2016, 55(42): 13061-6. DOI: 10.1002/anie.201604108.

[12] LIU Y, LI D, WANG T, et al. Efficient Conversion of Methane to Aromatics by Coupling Methylation Reaction. ACS Catalysis, 2016, 6(8): 5366-70. DOI: 10.1021/acscatal.6b01362.

[13] SU X, ZAN W, BAI X, et al. Synthesis of microscale and nanoscale ZSM-5 zeolites: effect of particle size and acidity of Zn modified ZSM-5 zeolites on aromatization performance. Catalysis Science & Technology, 2017, 7(9): 1943-52. DOI: 10.1039/c7cy00435d.

[14] CHENG Y T, JAE J, SHI J, et al. Production of renewable aromatic compounds by catalytic fast pyrolysis of lignocellulosic biomass with bifunctional Ga/ZSM-5 catalysts. Angew Chem Int Ed Engl, 2012, 51(6): 1387-90. DOI: 10.1002/anie.201107390.

[15] ZHENG Y, WANG F, YANG X, et al. Study on aromatics production via the catalytic pyrolysis vapor upgrading of biomass using metal-loaded modified H-ZSM-5. Journal of Analytical and Applied Pyrolysis, 2017, 126: 169-79. DOI: 10.1016/j.jaap.2017.06.011.

[16] YEH Y-H, GORTE R J. Study of Zn and Ga Exchange in H-[Fe]ZSM-5 and H-[B]ZSM-5 Zeolites. Industrial & Engineering Chemistry Research, 2016, 55(50): 12795-805. DOI: 10.1021/acs.iecr.6b03659.

[17] SHEN X, KANG J, NIU W, et al. Impact of hierarchical pore structure on the catalytic performances of MFI zeolites modified by ZnO for the conversion of methanol to aromatics. Catalysis Science & Technology, 2017, 7(16): 3598-612. DOI: 10.1039/c7cy01041a.

[18] CHU S, GUO X, LI J, et al. Synthesis of Ga2O3/HZSM-5@cubic ordered mesoporous SiO2 with template Pluronic F127 to improve its catalytic performance in the aromatization of methanol. Journal of Porous Materials, 2016, 24(4): 1069-78. DOI: 10.1007/s10934-016-0347-0.

[19] SING K S W, WILLIAMS R T. Physisorption Hysteresis Loops and the Characterization of Nanoporous Materials. Adsorption Science & Technology, 2004, 22(10): 773-82. DOI: 10.1260/0263617053499032.

[20] ZHANG X, ZHONG J, WANG J W, et al. Catalytic performance and characterization of Ni-doped HZSM-5 catalysts for selective trimerization of n-butene. Fuel Processing Technology, 2009, 90(7-8): 863-70. DOI: 10.1016/j.fuproc.2009.04.011.

RECONSTRUCTION OF SEVERAL LANDSLIDE DAMMED LAKES IN THE UPPER REACHES OF JINSHA RIVER AND ITS INFLUENCE ON RIVER PROFILE EASTERN TIBETAN PLATEAU

LiQin Zhou[1,2], Lu Cong[3*]

[1]*State Key Laboratory of Mountain Hazards and Engineering Resilience, Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, Sichuan, China.*
[2]*University of Chinese Academy of Sciences, Beijing 100049, China.*
[3]*Chengdu Academy of Educational Sciences (Chengdu Research Office of Basic Education), Chengdu 610036, Sichuan, China.*
Corresponding Author: Lu Cong, Email: cong_1u@yeah.net

Abstract: The damming phenomenon caused by landslides is a significant disturbance factor in the evolution of mountainous rivers, influencing river morphology across multiple timescales. The temporal characteristics of this phenomenon are often analyzed by examining the relationship between dam locations and the river longitudinal profile within a specific area. In the Batang section of the upper Jinsha River, located within the Jinsha river suture zone, intense tectonic activity makes the region highly susceptible to geological hazards such as landslides. Through data collection, remote sensing interpretation, and field investigations, we identified nine landslide dams in this area. Systematic studies of these dams were conducted, including preliminary measurements of dam heights and volumes. By reconstructing dam elevations, we calculated the areas, volumes, and peak discharge flows of the associated dammed lakes. A detailed analysis of the Wangdalong landslide dam revealed that its maximum lake surface area reached 106.70 million m², with a maximum volume of 14.57 billion m³. The peak discharge flow was calculated at 31.98×10^4 m³/s, and the dammed lake existed for a duration of approximately 1.09-1.76 ka. Additionally, a comparison of river geomorphological parameters indicated that the impact of a single dammed lake on the river's longitudinal profile is limited when its duration is relatively short. However, clusters of dammed barriers significantly affect the river steepness index, increasing the likelihood of forming knickpoints. These findings provide valuable insights into the role of landslide dams in shaping the geomorphological evolution of mountainous river systems.
Keywords: Landslide dammed lake; River profile; Luminescence dating; Jinsha River

1 INTRODUCTION

Landslide-induced river blockages represent a prevalent geohazard in mountainous terrain with steep topography and confined valleys. Even moderate slope failures can trigger cascading disasters through valley obstruction, subsequent lake formation, and catastrophic dam breaches, posing substantial risks to downstream communities. Particularly along the tectonically active Jinsha River on the Tibetan Plateau margin, recurrent damming events are facilitated by intense fluvial incision and neotectonic deformation [1]. The 2018 Baige landslides exemplify modern catastrophic damming incidents, generating two sequential barrier lakes and destructive outburst floods [2]. The resulting upstream inundation and downstream outburst floods caused extensive damage to infrastructure, including residential areas, bridges, and roadbeds along National Highway 318, underscoring the significant hazards associated with such events.

Landslide-dammed lakes represent a prevalent geological phenomenon in mountainous regions, arising from the interplay of topography, climatic conditions, and tectonic activities [3-5]. The vestiges of such events—including fluvial erosion traces and lacustrine sediments on riverbanks—serve as critical evidentiary markers for reconstructing and identifying historical river-blocking incidents [6]. By analyzing the elevation of lacustrine deposits, paleo-dammed lake water levels can be estimated, while the integration of geomorphic datasets enables the reconstruction of lake areal extent and storage capacity [7,8]. On geological timescales, landslide-induced river blockages and subsequent outburst flood events exert significant influences on fluvial geomorphology through modifications to sediment transport fluxes [9]. Recent research has demonstrated that such dams profoundly alter river longitudinal profiles, thereby complicating the interpretation of tectonic activities and reshaping the understanding of regional geomorphological evolution [1,10].

The upper Jinsha River in the Hengduan Mountain Range exhibits intense tectonic activity and steep topography, predisposing it to recurrent river-blocking landslides. While previous studies have reconstructed individual paleo-damming events and analyzed lacustrine sediments [11-14], key gaps remain in understanding their geomorphic impacts [7], particularly on longitudinal river profiles and quantitative landscape evolution. This study focuses on the Batang reach, a hotspot for paleo-landslide dams, employing field surveys, OSL dating, and digital topographic analysis to characterize dam structures, reconstruct reservoir capacities, and establish event chronologies. By evaluating the influence of these dams on river morphology, the research advances insights into landslide-fluvial interactions in tectonically active settings.

2   REGIONAL SETTING

The study area is situated in the northern Hengduan Mountains section of the upper Jinsha River on the southeastern margin of the Tibetan Plateau, within the Jinshajiang suture zone (Figure 1). Influenced by the rapid uplift of the Tibetan Plateau and intense fluvial incision, steep canyon landforms dominate the region [6]. Tectonic activity is pronounced, with major active faults such as the Batang Fault and Xiongsong-Suwalong Fault. Since 1722, seven earthquakes of magnitude ≥6 have been recorded, reflecting frequent geohazards [15]. The lithology is complex (Figure 2), primarily comprising Permian schists, basic volcanic rocks, crystalline limestones, Triassic siltstones, intermediate-basic volcanic rocks, hercynian medium-coarse diorites, and indosinian biotite granites.
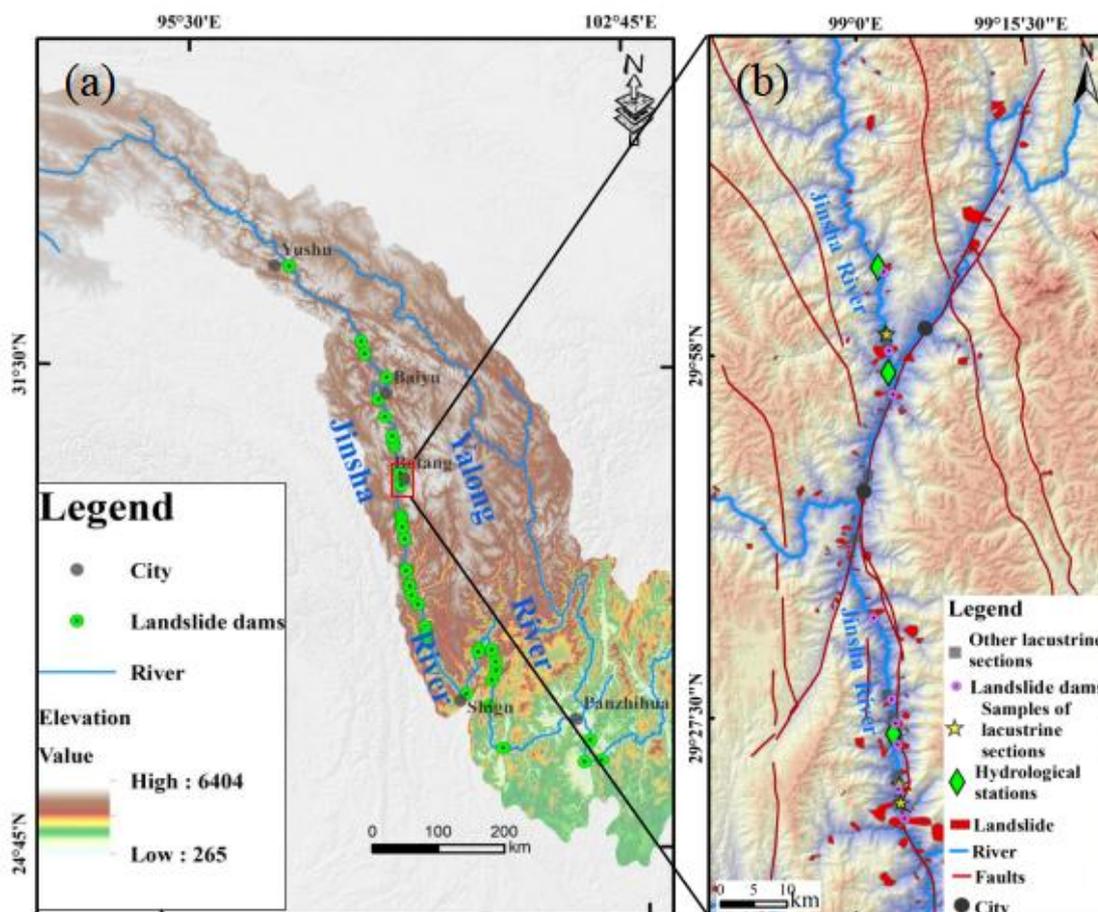
Figure 1 Geological Map of Batang Section of the Upper Jinsha River

The area experiences a dry-hot valley climate, with an average annual precipitation of 480.4 mm (94% concentrated between May and October). Sparse vegetation and intense physical weathering prevail. Quaternary deposits are dominated by fluvial and slope sediments, including low-liquid-limit clays, gravelly soils, and sand-cobble accumulations along riverbanks. Lacustrine sediments, primarily distributed in riverbeds and localized bank sections, serve as critical evidence for identifying paleo-damming events. Fine-grained lacustrine deposits (e.g., clays and silts) formed during stable lake phases contrast markedly with coarse-grained fluvial deposits (e.g., gravels and well-rounded pebbles).
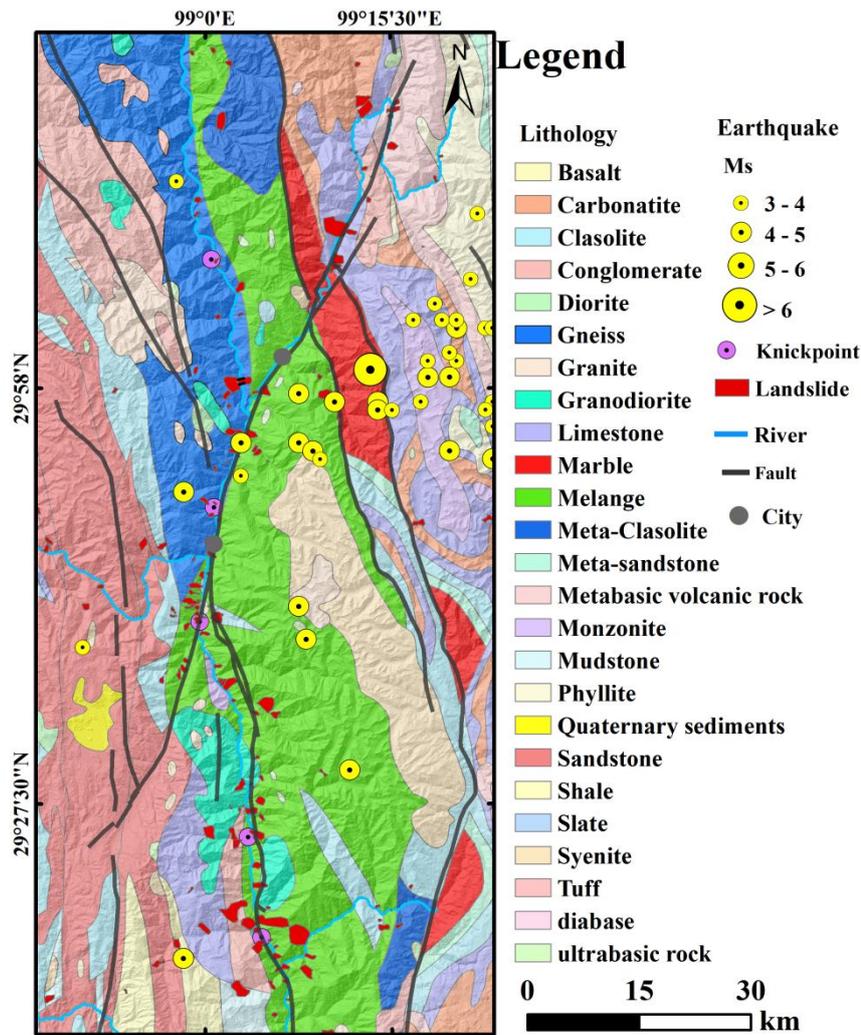
**Figure 2** Geological Map of Batang Section of the Upper Jinsha River

3   MATERIALS AND METHODS

3.1 Geomorphic Field Work and Sampling

First, through remote sensing interpretation and data collection, we obtained the location of some landslides in the Batang section of the main stream of the Jinsha River. Then the field survey along the Jinsha River to verify the accuracy of the remote sensing interpretation results. Through the investigation of the residual dam body accumulation and the lacustrine sediment, the exact position of the landslide dam is finally obtained. At the same time, we also sampled some of the lacustrine deposits for dating testing. We mapped the position, altitude, relative elevation to river levels using a laser TruPulse 200 range finder and handheld Trimble Juno 3B GPS (global positioning system). We obtained the height of the dam by the difference between the elevation of the dam outburst and the channel (Figure 3). Using field results, we delineated the extent of the landslide relict dam on DEM snd Google Earth, estimates the area of the dam, and then uses the landslide volume and area relationship derived [16]:

$$V = \alpha A^{\gamma} \tag{1}$$

Where V is the volume (m^3), A is the area (m^2), and α and γ are power-law scaling parameters, the α=0.23 and γ=1.41 is used to estimate the volume of the landslide dam [7,16].
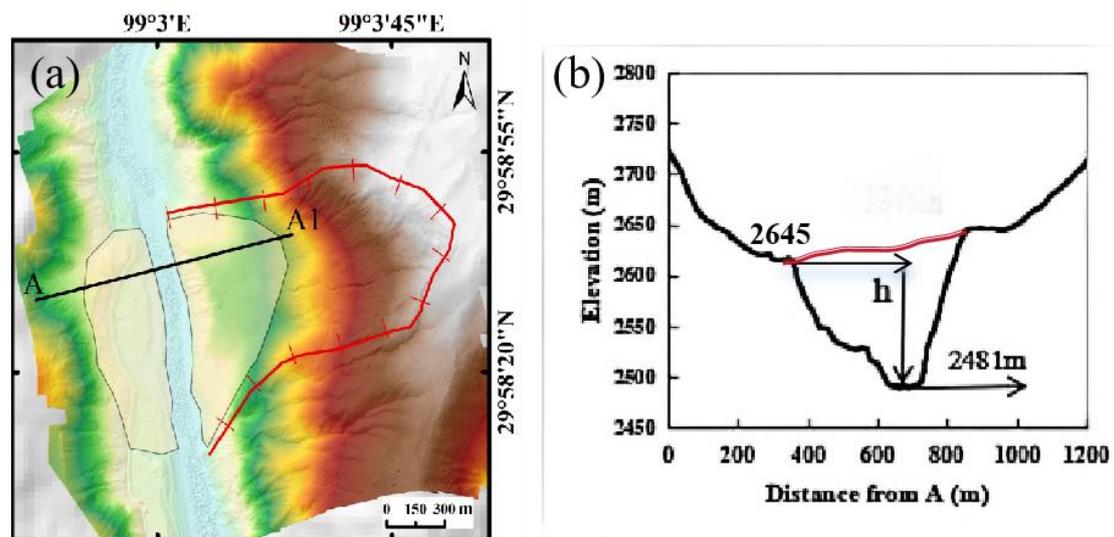
Figure 3 Range and Height of Temi Landslide Dam. (a) Range of Relict Dam of Temi Landslide (b) Line Graphs of Cross-Section AA1, h is Dam Height

3.2 Optically-stimulated luminescence dating

Ten samples were collected from four lacustrine sections for optically stimulated luminescence (OSL) dating to estimate the chronological limits of the studied dammed lake events. The chemical extraction of pure quartz grains from sediment and OSL measurements, which performed under subdued red light, based on the pretreatment steps and measurement procedure reported in Lai and Ou (2013) [17]. All the processes were completed in the OSL laboratory of the Qinghai Institute of Salt Lakes, Chinese Academy of Sciences, which is equipped with a RisøTL/OSL-DA-20 reader. In this study, we applied a combination of the single-aliquot regenerative-dose (SAR) protocol [18,19] and the standard growth curve (SGC) method [20] to determine the equivalent dose (De) values of samples. The optical measurements were performed after heating at a temperature of 260 °C for 10 s for natural and regenerative doses, and after preheating at a temperature 220 °C for 10 s for test doses. For 38-63 μm grains, the α efficiency value of 0.035 ± 0.003 was adopted during dose rate calculation [21]. The U, Th contents were measured using ICP-MS, and the K content was measured using ICP-OES in the Qinghai Institute of Salt Lakes, Chinese Academy of Sciences. And we used water content of $15\pm5\%$ to calculate dose rate calculations for lacustrine samples based on data on moisture content and modern annual precipitation [7,19]. The cosmic ray contribution was determined based on the burial depth and the geomagnetic coordinate of the research sites [22].

3.3 Digital Terrain Analysis

The 30 m SRTM DEM data used in this paper are all obtained from the US Geological Survey United States Geological Survey (USGS) data sharing platform (https://earthexplorer.usgs.gov/). In this paper, the TopoToolbox was used to extract the river geomorphic parameters [23], and the TopoToolbox to extract the DEM containing projection information as the input data (https://topotoolbox.wordpress.com/). Considering the geographical location of the Jinsha river, the projection coordinate system used in this study is WGS_1984_UTM_zone_47 N. To learn more about the effect of the landslide dam on the river channel, the channel width was measured using Google Earth images, the channel width selected is the modern wetted-width calculated every kilometer over 160 km.

To obtain the effect of the landslide dam on the river longitudinal profile in the Jinsha River, we used the Stream-Power River Incision Model, it can be quantified by relationship between upstream catchment area and channel gradient [24]:

$$S = k_s A^{-\theta} \tag{2}$$

where S is channel slope, k_s is the steepness index，A is the upstream catchment area. In this study we fixed concavity index of $\theta = 0.45$ to extract the normalized steepness index (k_{sn}) to facilitate the comparison of the differences between of more reaches [25]. In here, the value of k_{sn} was calculated by the algorithms in TopoToolbox with K=5 and τ=0.5, which is software for the analysis of DEMs [26]. In addtion, We also applied the KnickpointFinder function in TopoToolbox, which reproducibly extracts knickpoint locations from smooth river profiles, over the whole river network [23,26].

4 RESULTS

4.1 Characteristics of Landslide Dam

By remote sensing interpretation and field investigation, we found nine landslide dams in the Batang section of Jinsha River, From upstream to downstream is the lawa landslide dam, Temi landslide dam, Niuying landslide dam, Xiaguiwa landslide dam, Xuelongnang landslide dam, Suoduoxi landslide dam, Suwalong landslide dam, Biji landslide dam,

Wangdalong landslide dam (Figure 4). At the same time, we also found a total of 18 lacustrine sediments sections upstream of the landslide dam, but not all of the upstream dam retained lacustrine deposits. These lacustrine deposits are mainly distributed in upstream of Temi landslide dam, Xuelongnang landslide dam, Suoduoxi landslide dam, Suwalong landslide dam, Biji landslide dam, Wangdalong landslide. At the same time, we determined the area and height of the dam according to according to the position and elevation of the residual dam body, and then estimated the volume of the 9 landslide dams using the classical formula of area and volume [16]. The parameters of the landslide dam obtained are shown in Table 1, among them, the volume and height of the dam of Wangdalong landslide dam are the largest.

Figure 4 Nine Damming Landslides in the Upper Reaches of Jinsha River. (a) Lawa Landslide; (b) Temi Landslide; (c) Niuying Landslide; (d) Xiawagui Landslide; (e) Xuelongnang Landslide; (f) Suoduoxi Landslide; (g) Suwalong Landslide; (h) Biji Landslide; (i) Wangdalong Landslide. Image Credit: Google Earth

Table1 Scale Statistics of Landslide Dams

Name	Longitude (N)	Latitude (E)	Dam area ($10^4 m^2$)	Dam height (m)	Dam volume ($10^6 m^3$)
Lawa landslide	30° 5'4.44"	99° 2'35.16"	39.18	46	17.70
Temi landslide	29°58'25.50"	99° 3'3.01"	94.70	164	61.43
Niuying landslide	29°54'46.66"	99° 3'23.11"	92.83	46	59.72
Xiaguiwa landslide	29°35'58.15"	99° 1'38.05"	61.45	134	33.38
Xuelongnang landslide	29°29'5.58"	99° 3'18.86"	91.64	87	58.65
Suoduoxi landslide	29°27'5.33"	99° 3'40.81"	90.53	90	57.65
Suwalong landslide	29°25'27.65"	99° 3'57.76"	65.96	78	36.89
Biji landslide	29°21'33.24"	99° 3'56.00"	51.62	100	26.11
Wangdalong landslide	29°19'5.90"	99° 4'30.05"	224.75	316	207.78

4.2 OSL Dating Results

We selected 10 OSL samples from two lacustrine deposits sections (BT-1, BT-2) in upstream of Temi landslide dam

(Figure 5a, 5b), one lacustrine deposits section (BT-3) in upstream of Biji landslide dam (Figure 5c) and one lacustrine deposits section (BT-4) in upstream of WangDalong landslide dam (Figure 5d), the OSL dates for lacustrine sediments are summarized in Table 2. The BT-1 section height is approximately 11 m thick, the elevation is 2521 m, consists of silty clay varves. The BT-2 section height is approximately 1 m thick, the elevation is 2535 m, consists of climbing ripples of silt and clay vares, it belongs to the shallow lacustrine deposit. The BT-3 section height is about 5 m thick, the elevation is 2445 m, consists of Planar and parallel laminations lacustrine clay varves. The BT-4 section height is about 22.3 m thick, the elevation is 2391m, the upper 3.4-m-thick part contains fluvial gravel layer. The middle part, which is 6.7-mthick, consists of wavy lamination and typical lacustrine silt and clay varves. The lower part, which is about 13-m-thick, consists of contains climbing ripples of fine sand and silt, and the lamination is slightly wavy. A preliminary analysis of the results of the dating of the lake sediments in Table 2, we found that the dating results of these lacustrine section tended to be consistent and fit to the formation order, except for JSJ 23-26 sample. Further research found that there were more roots near the sample, which may have affected his chronological results. The bottom age of the four lake facies was mainly about 1.7 ka, while the top age of the lake facies was different.
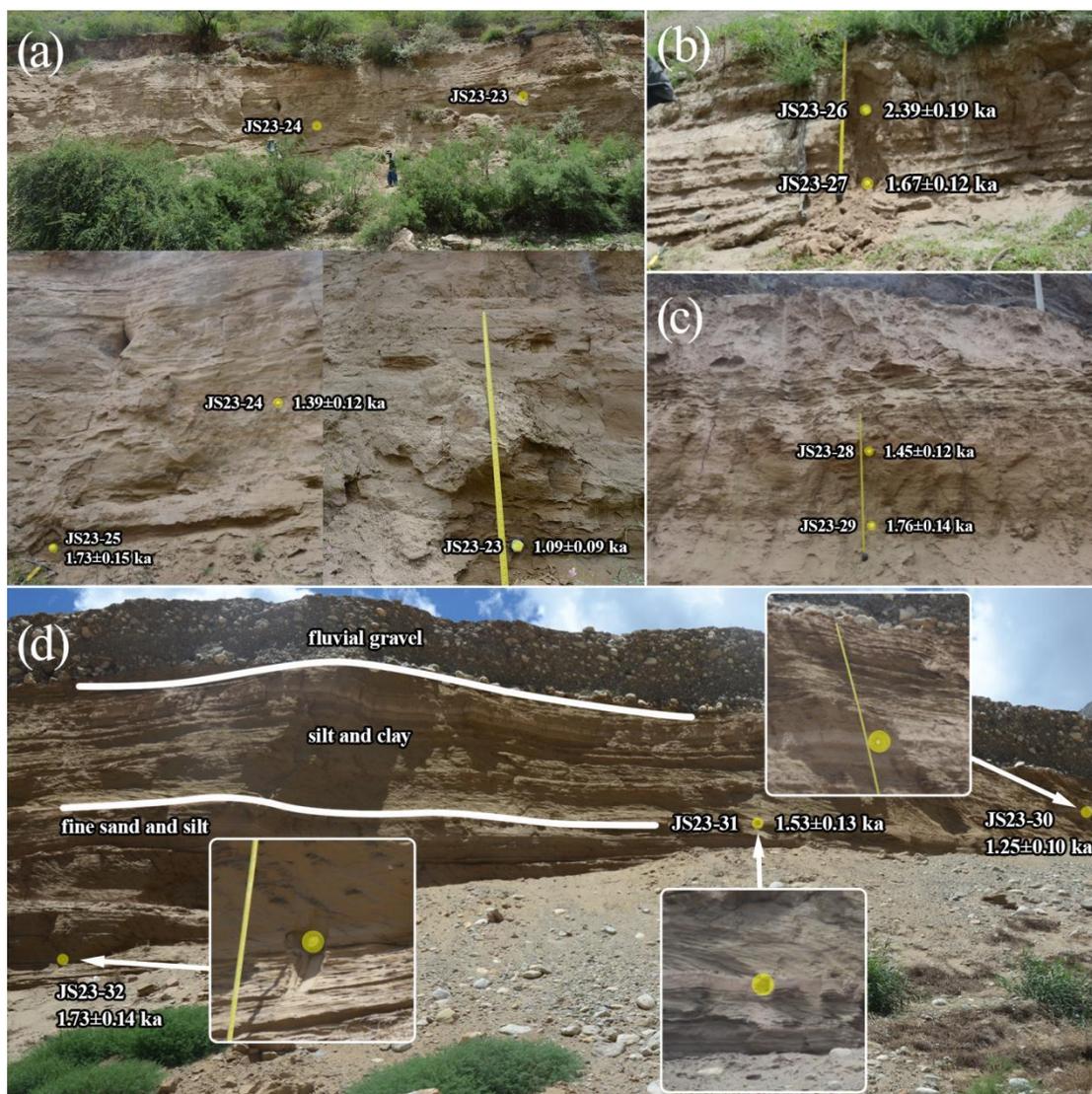
Figure 5 OSL Sample Sections of Lacustrine Sediments

Table 2 Palaeodose, Dose Rate and Ages Obtained from Quartz Extracted from Deposits

Section	Sample ID	Depth (m)	K (%)	Th (ppm)	U (ppm)	Water content (%)	Dose rate(Gy/ka)	No. of aliquotes	Final De (Gy)	OSL Age(ka)
BT-1	JS23-23	7.5	2.244±0.11	9.72±0.49	2.94±0.15	15±5	3.35±0.26	6[a]+12[b]	3.65±0.11	1.09±0.09
	JS23-	9.2	1.572±0.08	8.83±0.44	2.55±0.13	15±5	2.62±0.20	6[a]+12[b]	3.65±0.14	1.39±0.12

	24								
BT-2	JS23-25	10.6	1.342±0.07	7.33±0.37	2.14±0.11	15±5	2.21±0.17	6[a]+12[b]	3.82±0.15	1.73±0.15
	JS23-26	0.6	1.539±0.08	7.65±0.38	2.32±0.12	15±5	2.64±0.19	6[a]+12[b]	6.32±0.21	2.39±0.19
	JS23-27	1	1.572±0.08	7.54±0.38	2.30±0.11	15±5	2.68±0.19	6[a]+12[b]	4.48±0.08	1.67±0.12
BT-3	JS23-28	3.8	1.587±0.08	8.77±0.44	2.56±0.13	15±5	2.71±0.20	6[a]+12[b]	3.93±0.15	1.45±0.12
	JS23-29	4.7	1.339±0.07	6.68±0.33	1.94±0.10	15±5	2.19±0.16	6[a]+12[b]	3.85±0.09	1.76±0.14
	JS23-30	5.4	1.637±0.08	8.29±0.41	2.55±0.13	15±5	2.68±0.20	6[a]+12[b]	3.35±0.09	1.25±0.10
BT-4	JS23-31	9.2	1.572±0.08	7.87±0.39	2.52±0.13	15±5	2.54±0.19	8[a]+12[b]	3.90±0.15	1.53±0.13
	JS23-32	13.1	1.225±0.06	5.73±0.29	1.60±0.08	15±5	1.85±0.14	6[a]+12[b]	3.21±0.10	1.73±0.14

[a] Aliquot numbers using the SAR protocol; [b] aliquot numbers using the SGC protocol. OD is the overdispersion (%) of De values for each sample.

5 DISCUSSION

5.1 Extent of the Landslide-Dammed Lake

We found a total of 18 lacustrine deposits sections in the study area, and these deposits were all at elevations between 2377 and 2537 m, the elevation of most lacustrine deposits was much lower than the elevation of the nearest landslide dam downstream. Due to the presence of nine landslide dams in the area, we cannot decide which dam these lacustrine deposits were formed. We deduced the original or maximum lake level from the morphology of the relict dam, and the lake area, lake volume and backwater distance are calculated by using 30 m SRTM DEM data. The estimated the peak discharge (Q_p) of the break flood was determined by the formula [27]:

$$Q_P = 0.024 (V)^{0.701} \tag{3}$$

V is lake volume, the outburst flood peak discharge of the nine landslides dammed lake are calculated as below in table 3.

We superimposed the extent of the Wangdalong, Biji, Temi landslide dammed lake on the map (Figure 6), If the lake area is calculated according to lake level 2662 m from the morphology of the relict dam, Its backwater distance is enough to flood the upstream eight paleo landslide dam and all lacustrine deposits (Figure 6a). Therefore, it can be considered that the lacustrine deposits was formed by the Wangdalong landslide dam, the dammed lake duration time is between 1.09-1.76 ka, the lake area is 106.70 million m^2, the maximum lake volume is 14573.16 million m^3, the peak discharge is $31.98 \times 10^4 \, m^3/s$.

Table 3 Landslide Dammed Lake Volume and Peak Discharge Statistics

| Number	Name	Lake elevation (m)	Dam height (m)	Backwater distance （km）	Lake area ($10^6 m^2$)	Lake volume ($10^6 m^3$)	Peak discharge ($10^4 m^3/s$)
1	Lawa dammed lake	2585	46	6	1.17	39.31	0.51
2	Temi dammed lake	2645	164	20	6.92	613.8	3.47
3	Niuying dammed lake	2525	46	15	5.05	158.07	1.34
4	Xiaguiwa dammed lake	2535	134	56	26.52	1359.05	6.06
5	Xuelongnang dammed lake	2475	87	18	6.23	319.47	2.20
6	Suoduoxi dammed lake	2469	90	22	8.25	404.24	2.59
7	Suwalong dammed lake	2440	78	14	4.95	183.11	1.49
8	Biji dammed lake	2455	100	34	11.71	628.7	3.53
9	Wangdalong dammed lake	2662	316	107	106.70	14573.16	31.98
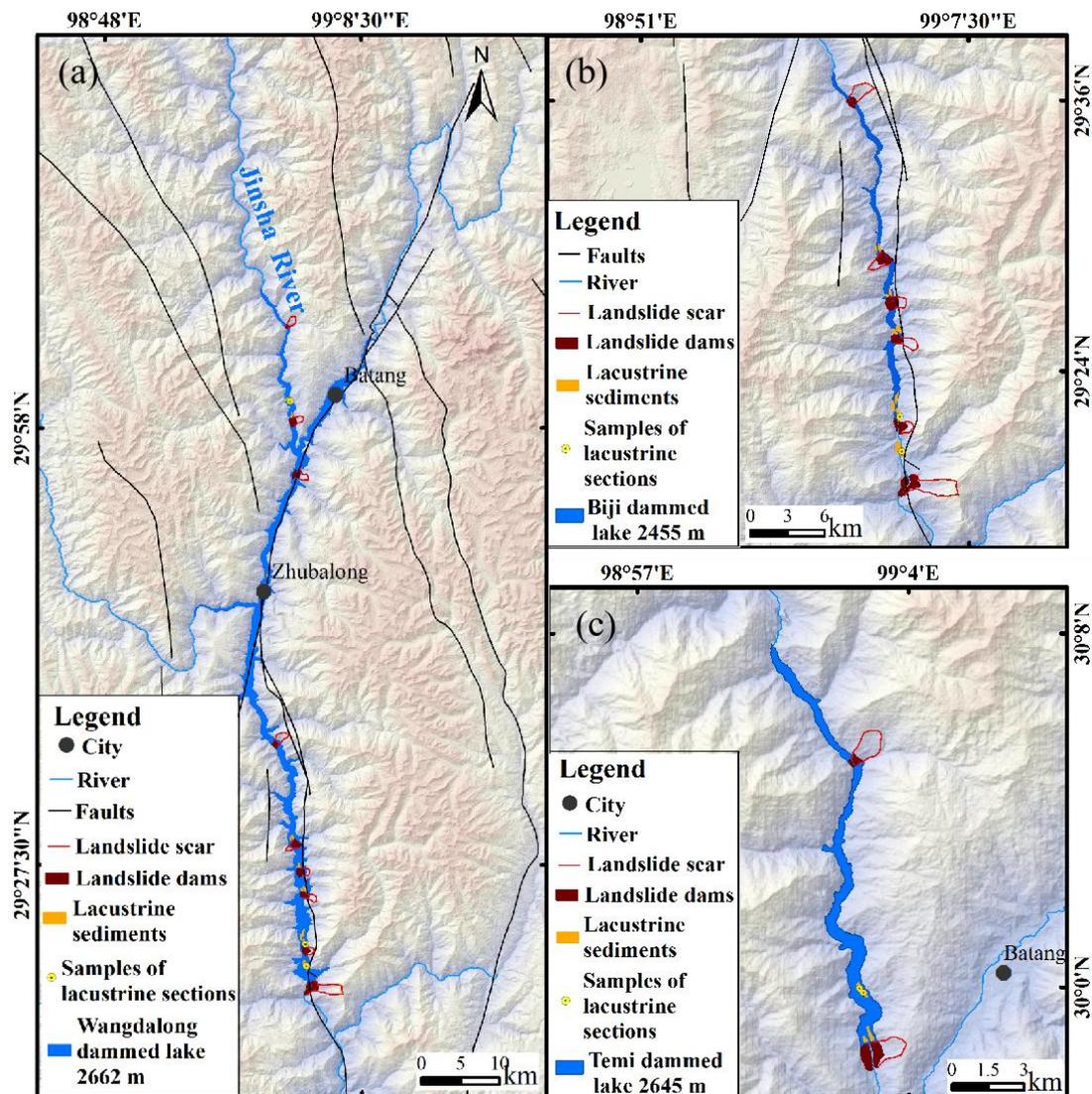
Figure 6 Paleo-Extent of the Wangdalong Dammed Lake. (a)Paleo -Extent of the Wangdalong Dammed Lake Levels of 2662 m a.s.l. (b) Palaeo-Extent of the Biji Dammed Lake Levels of 2455 m a.s.l. (c) Paleo Extent of the Temi dammed Lake Levels of 2645 m a.s.l

Based on optically stimulated luminescence (OSL) dating results of the Temi, Biji, and Wangdalong landslide-dammed lakes, this study reveals that the basal ages of these deposits cluster between 1.7-1.8 ka, demonstrating high chronological consistency. This suggests that these lakes likely formed contemporaneously. Previous studies by Chen et al. (2013, 2018) [11,12] using OSL and 14C dating proposed that at least one major landslide event occurred around 1.9 ka, damming the Jinsha River, with the Xuelongnang landslide dated to ~2.1 ka and the Temi landslide to ~1.8 ka. Wang et al. (2014) [14] reported that the Wangdalong landslide formed at ~1.9 ka, while the Suwalong landslide occurred later (~1.36 ka). Integrating these findings, this study posits that the Temi, Biji, and Wangdalong landslides likely occurred between 1.7–1.9 ka. Accounting for dating uncertainties, the formation periods of these dams broadly overlap. The spatiotemporal synchronicity of clustered landslide events strongly implies triggering by a single paleo-earthquake. A similar phenomenon was documented by Brooks [28] in the Quyon Valley (Quebec, Canada), where multiple landslides co-occurred between 980–1060 a and were seismically induced.The Jinsha River Fault Zone (particularly its Batang segment) has historically exhibited intense seismicity, with seven strong earthquakes (6≤Ms≤7.5) recorded since 1722. Numerous paleo-landslide dams along the left bank of the Jinsha River, aligned with the Xiongsong-Suwalong active fault, further support a tectonic origin [15]. The study area's two regional active faults—the Jinsha River Fault and Batang Fault—have driven recurrent strong earthquakes. Such tectonic activity progressively weakens rock mass integrity, promoting slope instability. Climatically triggered landslides typically exhibit temporal randomness and spatial dispersion [29], whereas strong earthquakes can induce widespread slope failures within short intervals. Consequently, this study infers that the paleo-landslide dams formed during 1.7–1.9 ka BP were likely seismically triggered under intense tectonic activity.

5.2 Effects of Damming on River Longitudinal Profile

We superimposed the nine landslide dams in the area on the longitudinal section of the Jinsha River according to the

dam volume (Figure 7). We found that the longitudinal profile is steeper where the dams more concentrated, It is especially prominent in the reach of Xuelongnang-Wangdalong. The steepness index (k_{sn}) of the this section river has high values, and form two river knickpoint (kp4 and kp5). However, not all dams have a significant impact on the river profile, the other four landslides (lawa, Temi, Niuying, Xiaguiwa) does not correspond to the river steepness index high value, this may be related to the landslide of the short duration, a single landslide river less impact on river landscape, and then when they are densely distributed together, the agglomeration effect is not ignore. On the Indus River and the Yarlung River, the longitudinal profile of the river is obviously convex where the dam is gathered, and forms the river knickpoint [10]. Further comparing the slope and width of the upstream and downstream of the dam, we find that the slope of the downstream of the dam becomes higher and the width is relatively smaller, Similar findings have been found in the geomorphic effect of baimakou landslide in Jinsha Rive [7]. Another scholars found that the landslide dam aligns spatially with the elevated sections of the river longitudinal profile and areas exhibiting high k_{sn} values, and coincides with specific river knickpoints in Min, Dadu, Indus and Parlung River [30-32].
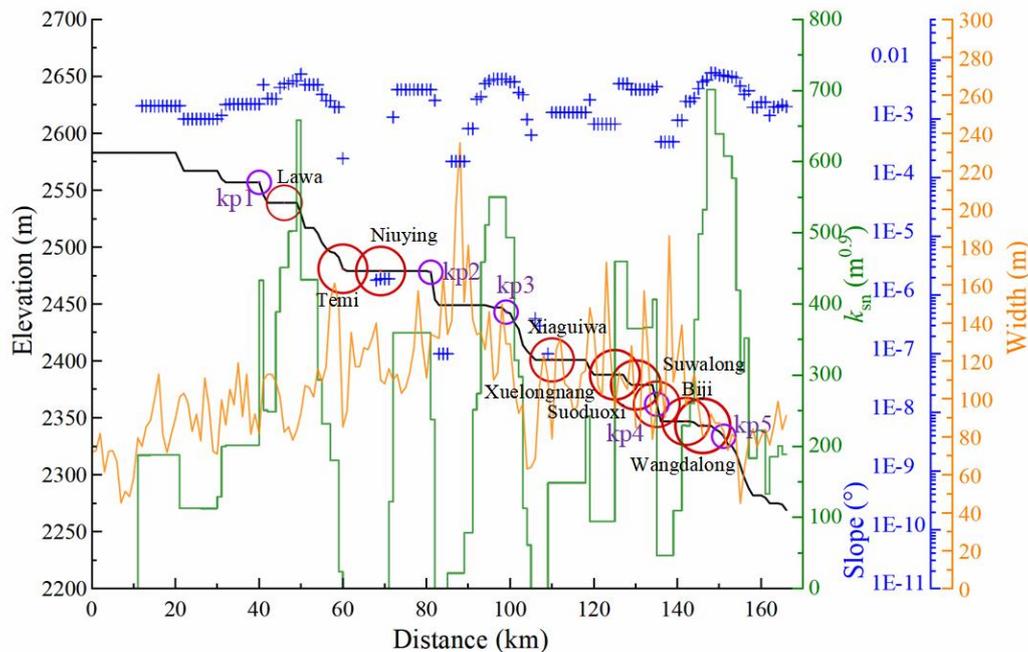
Figuer 7 Relationship between Dams, Knickpoints, k_{sn} , Longitudinal Profile, Channel Slope and width of the Jinsha River. River Bed Elevation (Black Line), Knickpoint (Purple Circle), Dam (Red Circle) Log Slope (Blue Cross), Channel Width (Orange Line), k_{sn} (Green Line)

6   CONCLUSION

Through data collection, remote sensing interpretation and field investigation, we found a total of nine landslide dams in the Batang section of the Jinsha River, among which the largest one was Wangdalong landslide dam，the dam height is 316 m, the dam volume is 207.78.×10^4 m³. Further reconstruction of the area landslide damed lake and OSL dating test, We determined that Wangdalong landslide dam maximum lake surface area reached 106.70 million m², with a maximum volume of 14.57 billion m³. The peak discharge flow was calculated at 31.98×10^4 m³/s, and the dammed lake existed for a duration of approximately 1.09-1.76 ka. The formation age of this dammed lake is estimated to be between 1.7 ka and 1.9 ka, while the outburst time is between 1.09 ka and 1.25 ka. The dammed lake lasted for approximately 400 to 600 years, and its formation may have been triggered by an earthquake. The densely distributed landslide damming events significantly impact river profiles, making it easier for river knickpoints to form.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Ouimet W B, Whipple K X, Royden L H, et al. The influence of large landslides on river incision in a transient landscape: Eastern margin of the Tibetan Plateau (Sichuan, China). Geological Society of America Bulletin, 2007, 119(11–12): 1462–1476.

[2] Ouyang C, An H, Zhou S, et al. Insights from the failure and dynamic characteristics of two sequential landslides at Baige village along the Jinsha River, China. Landslides, 2019, 16: 1397–1414.

[3]   Ding Y, Zhang X, He Z, et al. Sedimentary environment of a dammed lake buried in the modern riverbed of the Yalong River during the Last Glacial Maximum and its implication for fluvial geomorphic evolution. Geomorphology, 2021, 378: 107588.

[4]   Scheingross J S, Limaye A B, McCoy S W, et al. The shaping of erosional landscapes by internal dynamics. Nature Reviews Earth & Environment, 2020, 1: 661–676.

[5]   Wang H, Wang B B, Cui P, et al. Disaster effects of climate change in high-Mountain Asia: state of art and scientific challenges. Advances in Climate Change Research, 2024, 15(3): 367–389.

[6]   Li Y, Chen J, Zhou F, et al. Identification of ancient river-blocking events and analysis of the mechanisms for the formation of landslide dams in the Suwalong section of the upper Jinsha River, SE Tibetan Plateau. Geomorphology, 2020, 368: 107351.

[7]   Liu W, Hu K, Carling P A, et al. The establishment and influence of Baimakou paleo-dam in an upstream reach of the Yangtze River, southeastern margin of the Tibetan Plateau. Geomorphology, 2018, 321(15): 167–173.

[8]   Wang H, Wang P, Hu G, et al. An Early Holocene river blockage event on the western boundary of the Namche Barwa Syntaxis, southeastern Tibetan Plateau. Geomorphology, 2021, 395: 107990.

[9]   Korup O, Densmore A L, Schlunegger F. The role of landslides in mountain range evolution. Geomorphology, 2010a, 120(1–2): 77–90.

[10] Korup O, Montgomery D R, Hewitt K. Glacier and landslide feedbacks to topographic relief in the Himalayan syntaxes. Proceedings of the National Academy of Sciences of the United States of America, 2010b, 107(12): 5317–5322.

[11] Chen J, Dai F, Lv T, et al. Holocene landslide-dammed lake deposits in the Upper Jinsha River, SE Tibetan Plateau and their ages. Quaternary International, 2013, 298: 107–113.

[12] Chen J, Zhou W, Cui Z, et al. Formation process of a large paleolandslide-dammed lake at Xuelongnang in the upper Jinsha River, SE Tibetan Plateau: constraints from OSL and 14C dating. Landslides, 2018, 15(12): 2399–2412.

[13] Higgitt D L, Zhang X, Liu W, et al. Giant palaeo-landslide dammed the Yangtze river. Geoscience Letters, 2014, 1: 1–7.

[14] Wang P, Chen J, Dai F, et al. Chronology of relict lake deposits around the Suwalong paleolandslide in the upper Jinsha River, SE Tibetan Plateau: Implications to Holocene tectonic perturbations. Geomorphology, 2014, 217: 193–203.

[15] Xia M, Ren G M, Tian F. Mechanism of an ancient river-damming landslide at Batang Hydropower Station, Jinsha River Basin, China. Landslides, 2023, 20(10): 2213–2226.

[16] Larsen I J, Montgomery D R, Korup O. Landslide erosion controlled by hillslope material. Nature Geoscience, 2010, 3(4): 247–251.

[17] Lai Z, Ou X. Basic procedures of optically stimulated luminescence (OSL) dating. Progress in Geography, 2013, 32: 683–693.

[18] Murray A S, Wintle A. Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. Radiation Measurements, 2000, 32(1): 57–73.

[19] Liu W, Lai Z, Hu K, et al. Age and extent of a giant glacial-dammed lake at Yarlung Tsangpo gorge in the Tibetan Plateau. Geomorphology, 2015, 246: 370–376.

[20] Lai Z. Testing the use of an OSL standardised growth curve (SGC) for De determination on quartz from the Chinese Loess Plateau. Radiation Measurements, 2006, 41(1): 9–16.

[21] Lai Z, Brückner H. Effects of feldspar contamination on equivalent dose and the shape of growth curve for OSL of silt-sized quartz extracted from Chinese loess. Geochronometria, 2008, 30: 49–53.

[22] Prescott J R, Hutton J T. Cosmic ray contributions to dose rates for luminescence and ESR dating: large depths and long-term time variations. Radiation Measurements, 1994, 23(2–3): 497–500.

[23] Schwanghart W, Scherler D. TopoToolbox 2–MATLAB-based software for topographic analysis and modeling in Earth surface sciences. Earth Surface Dynamics, 2014, 2(1): 1–7.

[24] Whipple K X. Bedrock rivers and the geomorphology of active orogens. Annual Review of Earth and Planetary Sciences, 2004, 32(1): 151–185.

[25] Korup O. Rock-slope failure and the river long profile. Geology, 2006, 34(1): 45–48.

[26] Schwanghart W, Scherler D. Bumps in river profiles: uncertainty assessment and smoothing using quantile regression techniques. Earth Surface Dynamics, 2017, 5(4): 821–839.

[27] Liu W, Carling P A, Hu K, et al. Outburst floods in China: a review. Earth-Science Reviews, 2019, 197: 102895.

[28] Brooks G R. A massive sensitive clay landslide, Quyon Valley, southwestern Quebec, Canada, and evidence for a paleoearthquake triggering mechanism. Quaternary Research, 2013, 80(3): 425–434.

[29] Korup O, Görüm T, Hayakawa Y. Without power? Landslide inventories in the face of climate change. Earth Surface Processes and Landforms, 2012, 37(1): 92–99.

[30] Wang H, Cui P, Liu D, et al. Evolution of a landslide-dammed lake on the southeastern Tibetan Plateau and its influence on river longitudinal profiles. Geomorphology, 2019, 343: 15–32.

[31] Zhou L, Liu W, Chen X, et al. Relationship between dams, knickpoints and the longitudinal profile of the Upper Indus River. Frontiers in Earth Science, 2021, 9: 660996.

[32] Zhou Y, Liu W, Yanites B J, et al. Transient geomorphic response after landslide-induced river damming in the eastern margin of the Tibetan Plateau. Catena, 2024, 246: 108386.

CAN OPEN WEB DATA ASSESS URBAN FLOOD RISK? EVIDENCE FROM ZHENGZHOU

Da Mao[1,2*], Yan Xie[1], ZhiYu Yuan[1]

[1]*School of Horticulture and Landscape Architecture, Henan Institute of Science and Technology, Xinxiang 453003, Henan, China.*
[2]*Xinxiang Urban-Rural and Landscape Digital Technology Engineering Research Center, Xinxiang 453003, Henan, China.*
Corresponding Author: Da Mao, Email: maoda@foxmail.com

Abstract: Under extreme rainfall scenarios, urban stormwater pipe networks are prone to saturation and failure, rendering flood risk assessment under such conditions of critical importance. This study established a flood risk assessment framework using GIS technology and open web data to evaluate 270 catchment subunits in Zhengzhou's main urban area. Results reveal an overall stepped spatial pattern of flood risks, characterized by lower risks in the southwest and higher risks in the northeast, with contiguous high-risk clusters identified in the city center. Spatial autocorrelation analysis confirms strong spatial dependence of flood risks: High-High clusters concentrate in the urban core and northeastern districts, while Low-Low clusters are distributed across the higher-elevation southwestern region. The study further conducts bivariate correlation analysis between zonal flood risks and influencing factors, culminating in the proposal of targeted mitigation strategies.
Keywords: Waterlogging; Flood; Risk; Open data; Evaluate

1 INTRODUCTION

Urban waterlogging refers to the phenomenon of flood caused by excessive rainfall intensity or prolonged duration, which exceeds the drainage capacity of pipe networks. In recent years, extreme heavy rainfall events have occurred with increasing frequency in various regions[1-2]. Once urban flooding disasters triggered by torrential rains occur, disaster-bearing bodies such as pedestrians, roads, and buildings exposed to floods are highly vulnerable to threats, leading to casualties, waterlogging-induced traffic disruptions, water and power outages, economic losses, and even potential collapse of urban emergency command systems and basic support systems such as communication, electricity, and medical services . Urban flood not only affects the work and life of urban residents but also easily causes severe losses of life and property.How to assess the risk of urban flood disasters and ensure the safety of urban drainage and flood prevention has become an urgent issue for urban development[3].

Risk assessment is an effective means to alleviate urban flood disasters and has attracted the attention of many studies[4-5]. Previous studies generally suggest that the formation of flooding disasters is primarily caused by two factors: natural and anthropogenic. Natural factors mainly include urban topography and climatic characteristics, while anthropogenic factors include: the increase in impervious surfaces during urbanization, which reduces natural buffer zones, slows down rainwater infiltration, and leads to surface water accumulation during storms[6].

The high spatial heterogeneity of the built environment and the complexity of disaster-causing processes pose considerable challenges to urban pluvial flooding research, manifesting as insufficient model representativeness, low computational efficiency, and scarcity of basic and validation data[7]. Under extreme conditions, intensified surface water accumulation renders underground stormwater pipe networks nearly ineffective. Meanwhile, for large cities, the traditional method of establishing catchment areas involves enormous workloads, and data on underground pipelines are difficult to obtain. Many scholars from various countries have conducted research and attempts on the assessment of urban flood risks[8-11]. However, due to different ways of obtaining data, many indicators lack universality[12]. In this context, more and more research has begun to use open data based on the Internet[13-15]. This study proposes using open web data for flood risk assessment under extreme conditions to derive evaluation conclusions more efficiently and concisely.

2 DATA AND METHODS

2.1 Study Area

Zhengzhou, the capital of Henan Province, serves as a key central city in China's central region. On July 20, 2021, under the impact of extreme weather, Zhengzhou was struck by severe rainstorms and flooding. The hourly rainfall reached a staggering 201.9 mm that day, with a daily cumulative rainfall of 624.1 mm. Urban waterlogging has imposed significant impacts on Zhengzhou's strategic status and its economic and social fabric. The selection of Zhengzhou's main urban area as the study area is highly representative.

This study selected a continuous area within the central urban district specified in the public *Zhengzhou Sponge City Special Plan (2017-2030)* for assessment. Geographic information data were established through ArcGIS digitization

and georeferencing. The assessment area, projected using WGS1984-UTM-Zone50N, encompasses 838.56 km² divided into 270 subunits (Figure 1). The subunits range from 0.77 km² to 13.91 km² in area. These subunits are assigned distinct annual runoff control rate targets in the Zhengzhou Sponge City Special Plan (2017-2030), which well align with practical requirements for urban water management. The numbers in the right figure are the partition serial numbers given by the author.
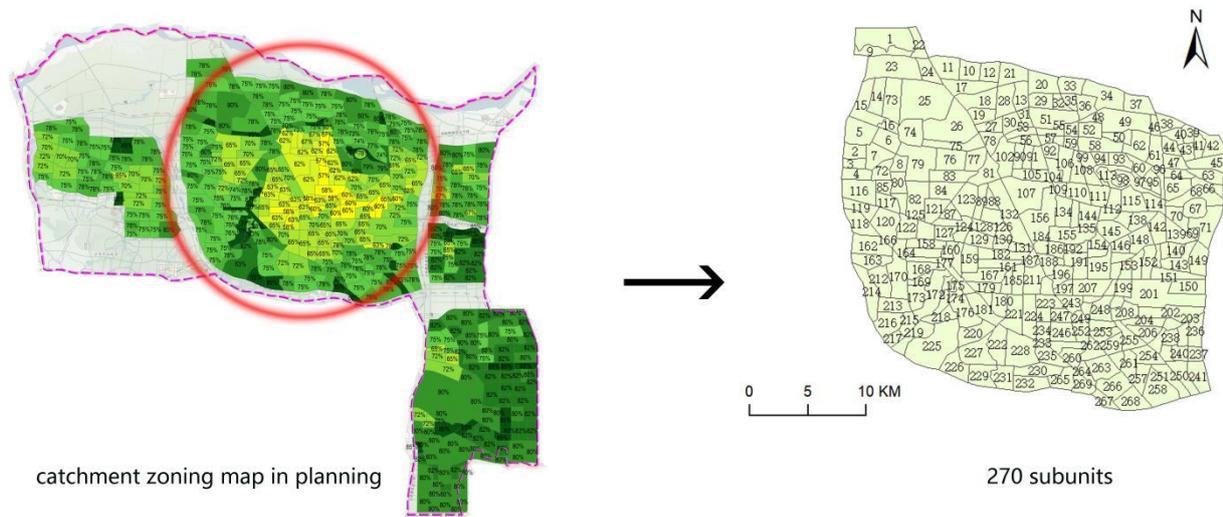

Figure 1 Study Area and Subunits Diagram

2.2 Open Web Data

2.2.1 DEM data

DEM (Digital Elevation Model) is a digital simulation of terrain through limited topographic elevation data, which realizes the digital expression of terrain surface morphology. It is currently widely applied in geographic assessment and analysis. The DEM data used in this study was downloaded from the Geographic Information Data Cloud, with a precision of 30 meters. By overlaying the evaluation units with DEM rasters in ArcGIS, three indices—minimum elevation, average elevation, and average slope—of each evaluation unit were extracted using tools such as Spatial Analyst (Figure 2).
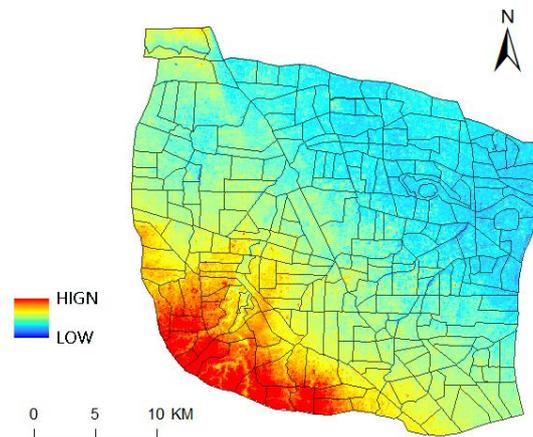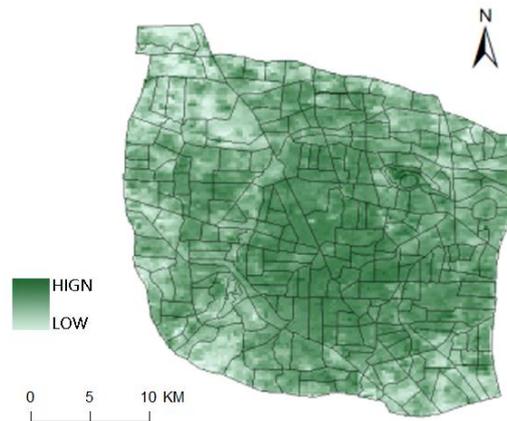

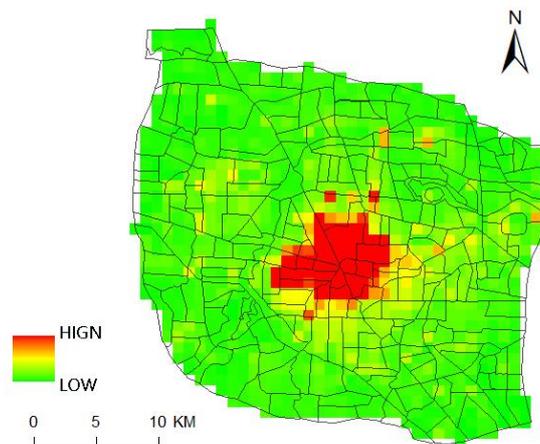**Figure 2** DEM Data and Overlay Map of Subunits

Figure 3 NDVI Data and Overlay Map of Subunits

2.2.2 NDVI data

NDVI (Normalized Difference Vegetation Index) is a digital characterization of surface vegetation coverage achieved through normalized operations of the difference and sum of reflectance values from remote sensing bands. Its core value lies in transforming vegetation physiological status into quantifiable numerical signals, which has been widely applied in ecological monitoring, agricultural yield estimation, and climate change research. The NDVI data used in this study was derived from the Geospatial Data Cloud in August 2023, with a spatial resolution of 250 meters. Using the ArcGIS platform, the vector boundaries of the study area were spatially overlaid with NDVI raster data (Figure 3), and the average NDVI values within each unit were extracted.

2.2.3 Population density data

The population density data in this study was primarily obtained from the LandScan global population density spatial distribution dataset, with a spatial resolution of 1 kilometer and a temporal coverage of 2020. Although population distribution is dynamic, this dataset effectively characterizes the spatial distribution and density of human populations. The study achieved a more refined simulation of population spatial distribution by allocating the total population of statistical units to raster grids and performing overlay extraction (Figure 4).

Figure 4 Population Density Data and Overlay Map of Subunits

2.2.4 Building density data

Building density is calculated by dividing the total building footprint area by the unit area. As a key indicator for measuring urban spatial development intensity, building scale, and land-use efficiency, it also serves as a critical metric for assessing urban impervious surfaces and flood risk. The building area data used in this study was derived from the vector building datasets of 77 major Chinese cities published online in 2022. After projecting these datasets, we overlaid them with the evaluation units (Figure 5) to calculate the building density of each unit.
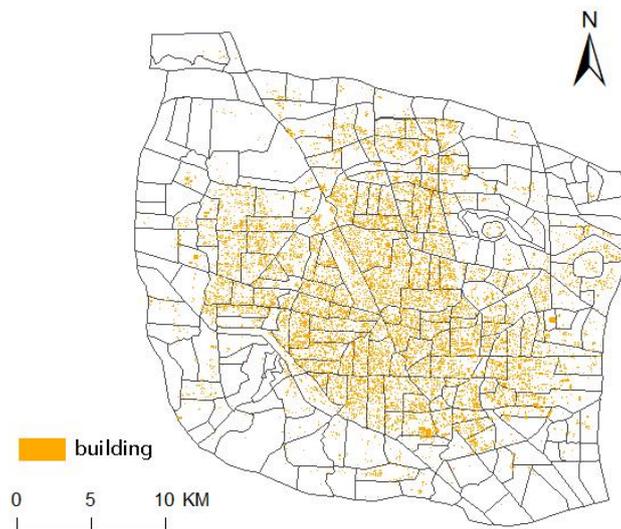
Figure 5 Building Density Data and Overlay Map of Subunits

2.3 Methods

2.3.1 Evaluation System

The unit-based flood risk assessment for Zhengzhou's urban core under extreme conditions is a concept with relatively clear objectives but vague details, necessitating the introduction of multiple evaluation indicators. The indicators were selected based on their maximum correlation with surface flood risk after urban stormwater pipe networks become saturated under extreme conditions. Referencing relevant research, an evaluation system was established (Table 1).

Table 1 Evaluation System of unit-based Flood Risk

Code	Indicator Name	Calculation Method	Effect
A1	Minimum Elevation	Extract the minimum value from the intersection of the unit area and DEM raster.	Negative
A2	Average Elevation	Extract the average value from the intersection of the unit area and DEM raster.	Negative
A3	Average Slope	Extract the average value from the intersection of the unit area and slope raster.	Negative
A4	Building Density	Calculate by dividing the building area by the area of the unit.	Positive
A5	Average NDVI	Extract the average value from the intersection of the unit area and NDVI raster.	Negative
A6	Population Density	Extract the average value from the intersection of the unit area and population density raster.	Positive

The scientific hypothesis of this assessment is that under extreme conditions, urban stormwater pipe networks are saturated and unable to function effectively in the short term, leading to urban waterlogging risks. The minimum elevation and average elevation within units are negative indicators, meaning that lower elevation values correspond to higher flood risks. The average slope is also a negative indicator because gentle slopes under extreme conditions slow surface runoff velocity, prolonging rainwater retention time in catchment units. If drainage facilities (e.g., pipe networks, drainage ditches) have insufficient design capacity for flood control, rainwater easily accumulates in low-lying areas, forming standing water.

Building density is a positive indicator: as building roofs are impervious surfaces, higher building density amplifies risks. The average NDVI is a negative indicator—lower average NDVI values signify higher waterlogging risks. Although green spaces lose much of their infiltration capacity under extreme conditions, vegetation itself still retains and dissipates rainwater. Finally, population density is a positive indicator: as vulnerable recipients of waterlogging disasters, higher population density correlates directly with greater risks.

2.3.2 Weight Determination

The flood risk assessment for Zhengzhou's urban core area under extreme conditions is a concept with relatively clear objectives but vague extensions, thus multivariate statistical analysis methods can be adopted for evaluation. The established evaluation system exhibits strong complexity, fuzziness, and dynamics in data information. To avoid weight bias caused by subjective analysis methods, the entropy weight model was selected for evaluation to objectively reflect data characteristics and patterns. Specifically, in this evaluation, smaller differences among units within the same indicator imply lower indicator weights. The specific calculation process is as follows:

First, the range method was adopted to standardize the statistical values of each indicator. Considering the requirement for logarithmic calculations in subsequent formulas, which necessitates avoiding zero values, the linear interpolation method was used to standardize the data within the range of 0.1 to 1. The formula is as follows:

$$X'_{ij} = \frac{X_{ij} - \min(X_j)}{\max(X_j) - \min(X_j)} \times 0.9 + 0.1 \quad \text{(Positive)} \tag{1}$$

$$X'_{ij} = \frac{\max(X_j) - X_{ij}}{\max(X_j) - \min(X_j)} \times 0.9 + 0.1 \quad \text{(Negative)} \tag{2}$$

Then, the standardized indicator data are normalized to obtain a matrix:

$$Z_{ij} = (X'_{ij} / \sum_{i=1}^{m} X'_{ij})_{m \times n} \tag{3}$$

Calculate the entropy value of the *j*-th indicator:

$$E_j = -\frac{1}{\ln m} \sum_{i=1}^{m} Z_{ij} \ln Z_{ij}; \tag{4}$$

Thus, the weight of the *j*-th indicator can be derived as:

$$W_j = (1 - E_j) / \sum_{j=1}^{n} (1 - E_j) \tag{5}$$

2.3.3 Model Calculation

Multiply the standardized indicator values of each subunit by the indicator weights determined via the entropy weight method to derive the flood risk assessment model for each catchment subunit in Zhengzhou's urban core under extreme conditions. This model is expressed as the sum of the products of individual risk indicators and their respective weights within each subunit. The formula is as follows:

$$R_i = \sum_{j=1}^{n} W_j X_{ij} \tag{6}$$

3 RESULTS AND ANALYSIS

3.1 Results

Results from the entropy weight method show that the differences in indicator weights are not significant (Table 2), indicating minimal variation in the within-indicator data across the 270 subunits. Specifically, Minimum Elevation and Average Elevation exhibit the highest weights, highlighting the most pronounced topographic differences.

Table 2 Indicator Weights Derived from Entropy Weight Method

Minimum Elevation	Average Elevation	Average Slope	Building Density	Average NDVI	Population Density
0.1669	0.1669	0.1668	0.1664	0.1667	0.1663

By substituting the indicator weights and standardized indicator values into the calculation model, the risk values for individual indicators and total flood risk values of the 270 subunits were derived. The average risk value across subunits is 0.4994, with the maximum and minimum values being 0.7066 and 0.1909, respectively—showcasing remarkable differences.

Jenks natural breaks classification divides data into different categories by minimizing the sum of squared within-class variances, enabling the identification of classification intervals to most appropriately group similar values. By inputting the total risk values of each subunit into ArcGIS and applying the Jenks natural breaks classification method for legend categorization, a flood risk assessment distribution map for each subunit was generated (Figure 6). The legend values represent flood risk levels, where higher numbers indicate greater risks.

The flood risks in Zhengzhou's urban core under extreme conditions exhibit an overall stepped spatial distribution, lower in the southwest and higher in the northeast. However, contiguous high-risk areas appear in the urban central region, mainly concentrated west of Zhongzhou Avenue and north of Zhengbian Road-East Street, as well as in the Longhu and Longzi Lake areas. Contiguous high-risk zones between Huayuan Road and Lianhuo Expressway in the northern region are also notably present.
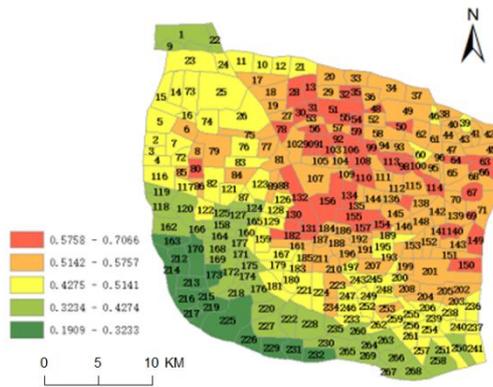
Figure 6 Flood Risk Assessment Distribution Map of Subunits

3.2 Spatial Autocorrelation Analysis

It is necessary to employ spatial autocorrelation to analyze the spatial characteristics of flood risks in Zhengzhou's urban core, using *Moran's I* index to determine whether spatial clustering exists within the region. The *Moran's I* statistic was computed using GeoDa 1.20 to quantify spatial clustering patterns of flood risk.
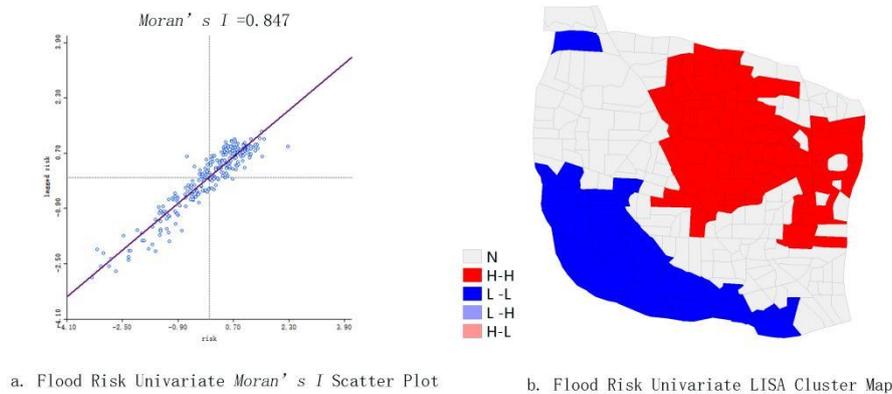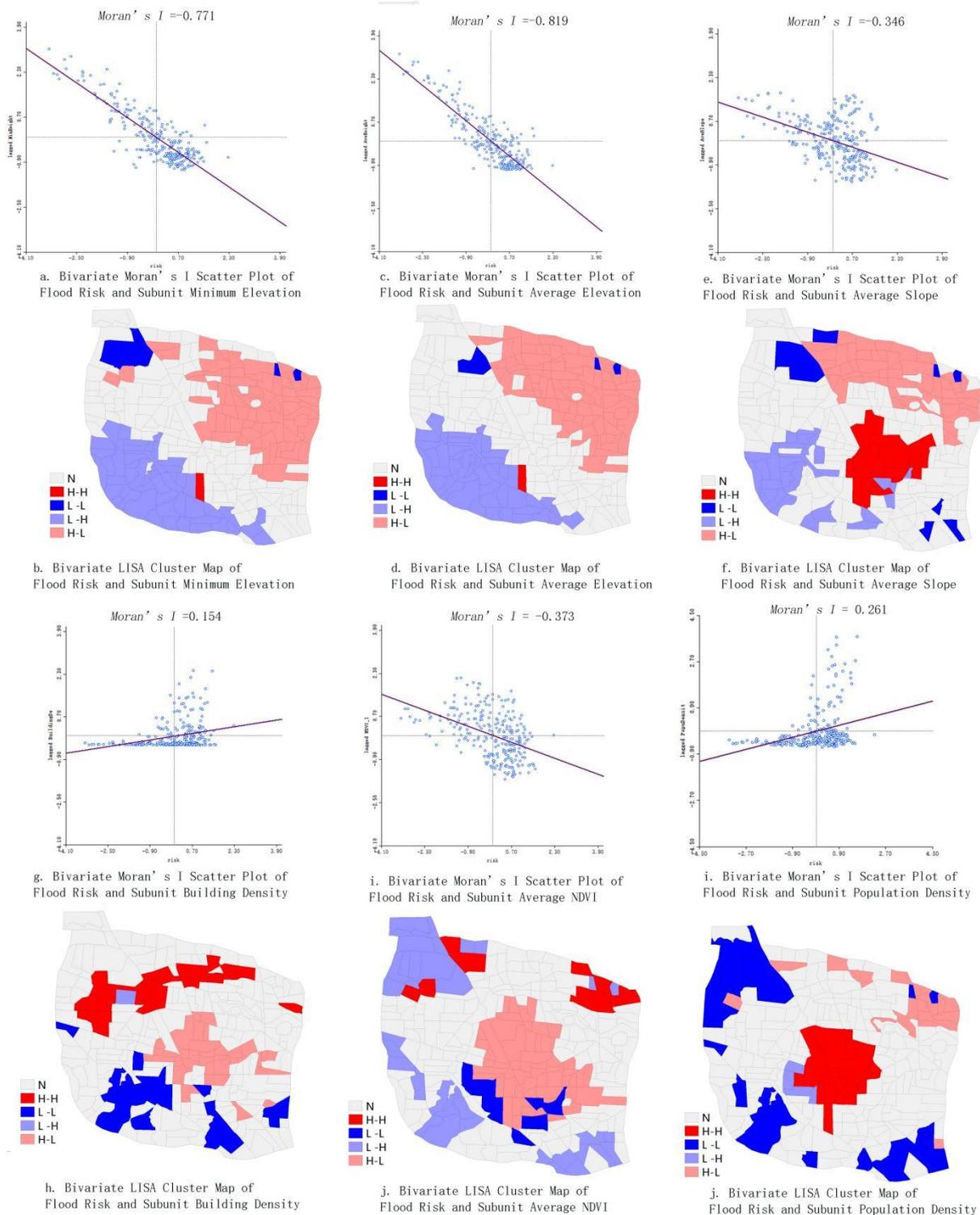

a. Flood Risk Univariate *Moran's I* Scatter Plot

b. Flood Risk Univariate LISA Cluster Map

Figure 7 Flood Risk *Moran' I* Scatter Plot and LISA Cluster Map

Moran's I value reflects the spatial clustering of an indicator within a region: a value greater than 0 indicates positive correlation, meaning high-indicator subunits are surrounded by high-indicator neighbors, and low-indicator subunits are surrounded by low-indicator neighbors. The univariate *Moran's I* value for flood risk across subunits was first calculated as 0.847, indicating significant spatial autocorrelation (Figure 7, a). Meanwhile, a local LISA cluster map generated by Geo-Da software (Figure 7, b) shows that High-High clusters are primarily concentrated in the city center and northeastern areas, while Low-Low clusters are mainly distributed in the higher-terrain southwestern regions.

The GeoDa software was also used to calculate local *Moran's I* values between bivariate variables and generate LISA cluster maps. This study sequentially calculated the raw values of flood risk against the minimum elevation, average elevation, average slope, building density, NDVI values, and population density of each subunit, generating corresponding bivariate scatter plots and LISA cluster maps(Figure 8) in turn.

Figure 8 Bivariate *Moran's I* Scatter Plots and LISA Cluster Maps of Flood Risk and Various Parameters

Bivariate spatial autocorrelation analysis shows that subunit average elevation has the highest correlation with flood risk, with a *Moran's I* value of -0.819, indicating a negative correlation. In terms of cluster distribution, High-Low and Low-High clusters are contiguous in the northeastern and southwestern urban areas. Both average slope and population density also exhibit negative correlations with flood risk, but unlike elevation indicators, High-High clusters are remarkably concentrated in the city center. The building density indicator has the smallest correlation with flood risk but shows a positive correlation (*Moran's I* = 0.154), still demonstrating a significant relationship. In spatial clustering, High-High clusters are aggregated in the northwestern region, High-Low clusters in the city center, and Low-Low clusters in the southern region. NDVI also has a significant negative correlation with flood risk (*Moran's I* = -0.373), with High-Low clusters mainly in the city center and Low-High clusters on the urban periphery, except for High-High clusters appearing only in the northern region.

4 CONCLUSION

This study presents a comprehensive flood risk assessment for Zhengzhou's urban core under extreme conditions, incorporating open web data with entropy weight modeling, spatial autocorrelation analysis, and GIS-based spatial distribution mapping. Key findings include:

(1) Spatial Distribution Pattern: Flood risks exhibit a stepped gradient from southwest to northeast, with high-risk clusters concentrated in central urban areas (west of Zhongzhou Avenue, north of Zhengbian Road-East Street) and northeastern districts (Longhu, Longzi Lake areas). Contiguous high-risk zones between Huayuan Road and Lianhuo Expressway further validate topographic influence.

(2) Indicator Significance: Topographic indicators (minimum/average elevation) dominate entropy weights, highlighting terrain as the primary risk determinant. Negative correlations between elevation/slope and flood risk confirm that lower-lying areas face higher vulnerability.

(3) Spatial Autocorrelation: Univariate *Moran's I* (0.847) reveals strong global spatial clustering, while bivariate analysis identifies elevation as the most influential factor *(Moran's I = -0.819)*. Building density shows weak positive correlation (*Moran's I* = 0.154), indicating urban construction's indirect impact.

(4) Policy Implications: High-risk clusters in central business districts and northeastern developments necessitate targeted drainage upgrades, while NDVI-negative correlation underscores green infrastructure's flood mitigation potential.

5 DISSCUSION

The study bridges methodological gaps by integrating entropy weight objectivity with spatial autocorrelation's locational insight. The -0.819 *Moran's* I between elevation and flood risk aligns with hydrological theory, but the unexpected positive correlation in building density (0.154) warrants further exploration — potentially reflecting impervious surface expansion overriding topographic effects. High-risk zones coincide with both low-lying areas and urban heat islands, possibly exacerbated by climate change-induced extreme precipitation.

Limitations include:

(1) Data Timeliness: NDVI and population density data represent snapshots, failing to capture recent urban expansion;

(2) Climate Scenario Simplification: Extreme conditions assume historical rainfall peaks, ignoring projected climate variability;

(3) Microtopography Omission: Subunit-scale elevation data may obscure localized depressions critical for flood accumulation.

Future research should incorporate real-time rainfall-runoff modeling, fine-scale LiDAR topography, and climate change projections to refine risk zonation. The framework also offers transferability to other rapidly urbanizing regions with complex topographies.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

REFERENCES

[1] Koc K, Isik Z. Assessment of Urban Flood Risk Factors Using Monte Carlo Analytical Hierarchy Process. Natural Hazards Review, 2021, 22(4): 14.

[2] Idowu D , Zhou WY. Global Megacities and Frequent Floods: Correlation between Urban Expansion Patterns and Urban Flood Hazards. Sustainability, 2023, 15(3). DOI: 10.3390/su15032514.

[3] Liu J. Research and Countermeasure on Key Issues in the Construction of Shenzhen Waterlogging Disaster Assessment System Oriented to Risk Management. Urban Geotechnical Investigation & Surveying, 2021(01): 22-26.

[4] Bin LL, Xu K, Pan H, et al. Urban flood risk assessment characterizing the relationship among hazard, exposure, and vulnerability. Environmental Science and Pollution Research, 2023, 30(36): 86463-86477.

[5] O'Donnell EC, Thorne CR. Drivers of future urban flood risk. Philosophical Transactions of The Royal Society a-Mathematical Physical and Engineering Sciences, 2020, 378(2168). DOI:10.1098/RSTA.2019.0216

[6] Jia B, Li Y. Study on Roughness Height and Resistance Characteristics of SubmergedVegetation Buffer Zone in Open Channel Flow Condition. Water Resources and Power, 2021, 39(6): 60-63,115.

[7] Huang H, Wang X, Liu L. A review on urban pluvial floods: Characteristics.mechanisms, data, and research methods. Progress in Geography, 2020, 40(6): 1048-1059.

[8] Caradot N, Granger D, Chapgier J, et al. Urban flood risk assessment using sewer flooding databases. Urban flood risk assessment using sewer flooding databases, 2011, 64(4): 832-840.

[9] Bakhtiari V, Piadeh F, Behzadian K, et al. A critical review for the application of cutting-edge digital visualisation technologies for effective urban flood risk management. Sustainable Cities and Society, 2023, 99: 14.

[10] Seemuangngam A, Lin HL. The impact of urbanization on urban flood risk of Nakhon Ratchasima, Thailand. Applied Geography, 2024, 162. DOI: 10.1016/j.apgeog.2023.103152.

[11] Peixoto JPJ, Costa DG, Portugal P. Flood-Resilient Smart Cities: A Data-Driven Risk Assessment Approach Based on Geographical Risks and Emergency Response Infrastructure. Smart Cities, 2024, 7(1): 662-679.

[12] Zhu SA, Jiang YX, Zhang J, et al. Evaluating the effect of urban flooding on spatial accessibility to emergency shelters based on social sensing data. Transactions in GIS, 2024, 28(1): 23-39.

[13] Niemi TJ, Warsta L, Taka M, et al. Applicability of open rainfall data to event-scale urban rainfall-runoff modelling. Journal of Hydrology, 2017, 547: 143-155.

[14] Lin T, Liu XF, Song JC, et al. Evaluating the effect of urban flooding on spatial accessibility to emergency shelters based on social sensing data. Habitat International, 2018, 71: 88-96.

[15] Ren C , Cai M, Li XW, et al. Developing a rapid method for 3-dimensional urban morphology extraction using open-source data. Sustainable Cities and Society, 2020, 53. DOI: 10.1016/j.scs.2019.101962.

INTELLIGENT PREDICTION OF WATER QUALITY PARAMETERS BASED ON IMPROVED PARTICLE SWARM OPTIMIZATION COUPLED WITH RANDOM FOREST MODEL

JinYi Luo
School of Mathematical Science, Chengdu University of Technology, Chengdu 610059, Sichuan, China.
Corresponding Email: ljinyi77@gmail.com

Abstract: In recent years, the problem of water pollution has become increasingly serious, and traditional water quality monitoring methods are difficult to meet the demand for high precision. Constructing an efficient and reliable water quality prediction model is of great significance for governance decisions. To this end, this paper proposes a random forest (IPSO-RF) water quality prediction model optimized based on the improved particle swarm algorithm. Firstly, for the problems of traditional particle swarm algorithm (PSO), such as easy premature convergence and insufficient global search ability, an improved particle swarm algorithm (IPSO) with nonlinear iteration incorporating inertia weights is proposed, and its optimization performance is verified based on typical test functions. Secondly, the IPSO algorithm is combined with random forest (RF) to optimize the key hyperparameters of RF (e.g., the number of decision trees, the minimum number of samples for node splitting, etc.) in order to improve the prediction accuracy and generalization ability of the model. Simulation experiments were carried out based on the water quality monitoring data of a watershed, and the results showed that compared with RF and standard PSO-RF and other models, the IPSO-RF model showed lower MSE in the prediction of key indicators such as dissolved oxygen, phosphorus content of water body and ammonia nitrogen amount of water body, which verified its superiority in the prediction of water quality. This study not only provides new ideas for the application of intelligent optimization algorithms in the field of water environment, but also provides more accurate technical support for water quality monitoring and pollution prevention and control of environmental protection departments.
Keywords: Random forest; IPSO-RF; Inertia weight; PSO

1 INTRODUCTION

As China's economy and society continue to develop rapidly, the problem of water supply and demand has become increasingly prominent, and the problem of water pollution has become one of the important factors limiting sustainable development. According to the Bulletin of China's Ecological and Environmental Conditions released by the Ministry of Ecology and Environment, about 20% of the surface water monitoring sections in the country will still fail to meet the Class III water quality standard in 2022. Traditional water quality monitoring methods have limitations in terms of high cost and poor timeliness, making it difficult to meet modern environmental governance needs. In this context, intelligent optimization algorithms have brought new technological breakthroughs in the field of water quality monitoring. And the first large-scale application of the intelligent early warning system for the cyanobacterial outbreak in Lake Taihu in 2016 marked the entry of China's water quality monitoring into the intelligent development stage. In recent years, the application of machine learning algorithms in water quality prediction has become more and more widespread, which greatly promotes the environmental protection department, water enterprises and scientific research institutions to pay more attention to the innovation and application of intelligent monitoring technology, and provides new technical support for water environment management.

At present, scholars both at home and abroad have been conducting research on the construction of water quality monitoring models and evaluating various water quality indicators. The research methods involve machine learning methods and artificial intelligence. Taher Rajaee et al. used methods including artificial neural network (ANN), genetic programming (GP), fuzzy logic (FL), support vector machine (SVM), hybrid neural-fuzzy (NF), hybrid ANN-ARIMA, and hybrid genetic algorithm-neural network (GA-NN) to model and predict river water quality parameters such as dissolved oxygen (DO) and suspended sediment [1]. Mosleh Hmoud Al-Adhaileh et al. used feedforward neural network (FFNN) and K-nearest neighbor for water quality classification, and proposed advanced artificial intelligence methods that can assist water treatment and management. The ANFIS model accurately predicted WQI [2]. Yingyi Chen et al. conducted extensive investigations and analyses on water quality prediction based on ANN from three aspects: feedforward, feedback, and hybrid architectures [3]. Lu et al. used extreme gradient boosting (XGBoost) and random forest (RF) to predict six water quality indicators such as water temperature, dissolved oxygen, pH value, specific conductance, turbidity, and fluorescent dissolved organic matter [4].Compared with the traditional water quality monitoring methods, machine learning techniques show better performance and prediction accuracy in constructing water quality prediction models; among the many machine learning algorithms, the random forest (RF) model derived from decision trees has been widely used in water quality classification and regression prediction due to its excellent generalization ability and robustness. However, it should be pointed out that traditional machine learning methods are often limited by parameter setting issues, such as the selection of kernel function and penalty coefficient

setting of support vector regression (SVR), which can significantly affect the model performance and easily lead to problems such as overfitting or underfitting. The emergence of population intelligent optimization algorithms, on the other hand, provides an effective technical way to solve the parameter optimization problems of traditional machine learning models.

In this paper, based on the systematic analysis of related research at home and abroad, an improved particle swarm algorithm (IPSO) incorporating inertia weight nonlinear iteration is proposed and applied to the optimization study of water quality prediction model. The research content of the whole paper mainly includes the following aspects: firstly, introducing the basic principles of random forest (RF) model; secondly, elaborating the basic principles of the standard particle swarm algorithm, proposing the improvement strategy based on the inertia weight nonlinear iteration for the problems such as easy to fall into the local optimization, and choosing the two test functions of Sphere and Rastrigin to verify the performance of the IPSO algorithm, through the comparison experiment with the PSO algorithm, and through the optimization of water quality forecasting model. The comparison experiment with PSO algorithm confirms the superiority of the improved algorithm, constructs the IPSO-RF integrated prediction model, applies it to the task of water quality prediction, and finally puts forward the targeted optimization suggestions for water quality monitoring based on the experimental results.

2 DESCRIPTION OF APPLICATION METHODS

2.1 Particle Swarm Algorithm

Particle swarm optimization algorithm (PSO) belongs to an important branch of population intelligence optimization algorithms, and together with ant colony algorithms and artificial fish swarm algorithms, it constitutes a research hotspot in the field of intelligent computing [5]. The algorithm was first proposed by Kennedy and Eberhart in 1995, and its design was inspired by the observation study of bird group foraging behavior. In nature, an efficient strategy for birds to search for food is to congregate to the area of the individuals in the group that are closest to the food source.The PSO algorithm is an optimization method developed to simulate this characteristic of intelligent behavior of biological groups.Each particle in the algorithm characterizes a possible solution to the problem and evaluates the degree of its superiority or inferiority by means of a fitness function. During the search process, the particles continuously adjust their speed and direction according to their own historical optimal position and the optimal position of the group, thus realizing intelligent exploration and optimization in the solution space. In order to calculate the fitness value corresponding to the position of each particle according to the objective function, it is assumed that in a D-dimensional search space, a population consisting of n particles $X = (X_1, X_2, \ldots X_n)$, where the i-th particle is denoted as a D-dimensional vector $X_i = (X_{i1}, X_{i2}, \ldots, X_{iD})^T$, represents the position of the ith particle in the D-dimensional search space, which is a potential solution to the problem. The velocity of the ith particle is $V_i = (V_{i1}, V_{i2}, \ldots, V_{iD})^T$, Its individual extremes are $P_i = (P_{i1}, P_{i2}, \ldots, P_{iD})^T$, And the population extremes of the population are $P_g = (P_{g1}, P_{g2}, \ldots, P_{gD})^T$.

During each iteration, the particle updates its velocity and position through the individual poles and the population poles, i.e：

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k) \tag{1}$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \tag{2}$$

where ω is the inertia weight; $d = 1,2,3,\ldots D$; $i = 1,2,3,\ldots,n$; k is the current iteration number; V_{id} is the particle velocity; and c_1 and c_2 is a non-negative constant called acceleration factor; and r_1 and r_2 is a random number distributed in the interval $[0，1]$. In order to prevent blind search of particles, it is generally recommended to limit the position and velocity to a certain interval as $[-X_{max}，X_{max}]$、$[-X_{max}，X_{max}]$.

Basic steps:

Step 1. Parameter setting:

General order $c_1 = 1.49445$，$c_2 = 1.49445$

Step 2. Population initialization:

Randomly initialize particle position and particle velocity, and calculate particle fitness value according to fitness function.

Step 3. Finding initial extreme value.

Finding individual extreme value and population extreme value according to the initial particle fitness value

Step 4. Iterative optimization

Update the particle position and velocity according to Eqs (1) and (2), and update the individual extreme value and population extreme value according to the fitness value of the new particles.

Step 5. Results analysis

2.2 The Random Forest Model

Random Forest (RF) is a machine learning algorithm based on decision tree integration, which was first proposed by statistician Leo Breiman in 2001. As a typical integrated learning method, the algorithm significantly improves the stability and accuracy of the model by constructing multiple decision trees and synthesizing their predictions, effectively solving the problems of weak generalization ability and large fluctuations in prediction of a single classifier [6]. The core implementation process of the algorithm mainly includes the following key links:

1. Random sampling with put-back from the original training set containing N samples, i.e., Bootstrap sampling technique is used to generate multiple training subsets with the same capacity;

2. In the feature selection stage, for the samples possessing M-dimensional features, each time the node splitting is performed, not all the features are considered, but only the features composed from the m features selected randomly are taken into account. Instead of considering all the features, the optimal split features are determined from a subset of m randomly selected features;

3. Each decision tree adopts the full-growth strategy without pruning, in order to ensure the diversity of the base learner, it is necessary to make each decision tree adopt the full-growth strategy without pruning;

4. A large number of decision trees are constructed through the iterative process described above, and a complete Random Forest system is formed;

5. The prediction results of all decision trees are aggregated by the set strategy. Aggregate the prediction results of all decision trees.

Through the synergistic effect of "three randomnesses" (sample randomness, feature randomness, split randomness), this integration mechanism not only significantly enhances the model's generalization ability and overfitting resistance, but also improves the model's prediction accuracy, which makes it an effective tool for dealing with complex classification and regression problems. Similar to most machine learning algorithms, the performance of the random forest model is highly dependent on the setting of key parameters. In Python's scikit-learn implementation, the following three core parameters have a decisive impact on model performance: n_estimators, max_features, and max_depth. The settings of these three hyperparameters significantly affect the model's prediction accuracy and computational efficiency.

Insufficient n_estimators leads to an underfitted model. Excessive n_estimators will increase unnecessary computational overhead, and the choice of max_features and max_depth directly affects the degree of exploration of the feature space, and unreasonable settings will reduce the model generalization ability. In the subsequent study, we will focus on optimizing these two key parameters, and find the optimal parameter combinations through intelligent optimization algorithms to improve the performance of the Random Forest model in water quality prediction tasks.

3 DATA COLLECTION

The water quality monitoring data in this study were mainly obtained from the public database of the National Environmental Quality Monitoring Network for Surface Water (http://www.cnemc.cn), and the monthly monitoring data from key river basins in China were selected as the study samples. The dataset contains seven key water quality indicators, including pH, dissolved oxygen (O2), permanganate index (KMn), ammonia nitrogen (NH4N), total phosphorus (P), total nitrogen (N), and flow direction (Dir).

In the data preprocessing stage, the following quality control measures were used:

1. Missing value processing: for the missing data in the monitoring network, the standard deviation weighted interpolation method based on time series was used. Specifically, the inter-month standard deviation of the historical data of each monitoring station is first calculated, and then the standard deviation is used as the threshold to weight the missing values for interpolation to ensure data continuity.

2. Outlier detection and processing: Physical extreme value calibration for pH (normal range 6-9), dissolved oxygen (>2mg/L) and other indicators

3. Data standardization: Improvement of data normality for NH4N, P and other concentration indicators, and all characteristic variables were finally processed to the [0,1] interval using Min-Max standardization [7].

After pre-processing, it provides a reliable data base for subsequent modeling, and Table 1 shows some of the data:

Table 1 Treated Water Quality Data (Partial)

pH	O2	KMn	NH4N	P	N	Dir	WR
8.39	13.12	6.03	0.114	0.095	2.25	23.3	4
8.52	12.94	5.65	0.105	0.081	2.3	14.6	3
8.39	12.3	5.68	0.106	0.082	2.12	17.5	3
8.49	12.71	5.66	0.17	0.098	2.08	10.9	3
8.54	12.14	5.68	0.099	0.078	2.06	15.8	3
8.53	12.26	5.78	0.101	0.084	2.09	22.1	3
8.5	12.14	5.67	0.106	0.091	2.08	20.9	3
8.51	13.18	5.39	0.106	0.079	1.95	18.5	3
8.51	12.7	5.8	0.125	0.114	2.17	30.6	3
8.52	12.3	5.64	0.122	0.109	2.18	28.7	3
8.5	11.86	6.21	0.139	0.167	2.47	59	4
8.53	12.86	5.47	0.108	0.09	2.04	22.4	3
8.5	12.53	5.49	0.106	0.1	2.07	28.7	3

| 8.5 | 12.15 | 5.62 | 0.11 | 0.112 | 2.3 | 28.3 | 3 |

4 RESULTS

4.1 Introduction of Inertia Weights

The inertia weight reflects the degree of inheritance of the particle to the previous velocity [8].Shi.Y firstly incorporated the inertia weight into the PSO algorithm, which was analyzed to show that larger inertia weights are more favorable for the global search, while smaller inertia weights are more advantageous for the local search. In order to strike a balance between the global and local search capabilities of the algorithm more effectively, Shi.Y proposed linear decreasing inertia weight (LDIW), i.e:

$$\omega(k) = \omega_{start}(\omega_{start} - \omega_{end})(T_{max} - k)/T_{max} \tag{3}$$

where, ω_{start} is the initial inertia weight; ω_{end} is the inertia weight when iterating to the maximum number of iterations; k is the current number of iterations; and T_{max} is the maximum number of iterations.
Common inertia weight selection:

$$\omega(k) = \omega_{start} - (\omega_{start} - \omega_{end})\left(\frac{k}{T_{max}}\right)^2 \tag{4}$$

$$\omega(k) = \omega_{start} + (\omega_{start} - \omega_{end})\left[\frac{2k}{T_{max}} - \left(\frac{k}{T_{max}}\right)^2\right] \tag{5}$$

$$\omega(k) = \omega_{end}\left(\frac{\omega_{start}}{\omega_{end}}\right)^{1/(1+ck/T_{max})} \tag{6}$$

4.2 Test Functions

We chose two of the 23 Benchmark benchmark functions, the Sphere function and the Rastrigin function, to introduce the mechanism of variable inertia weights to improve the particle swarm algorithm to obtain the IPSO on the basis of the particle swarm algorithm with fixed inertia weights at the base.
Sphere test function:

$$f(x) = \sum_{i=1}^{n} x_i^2 \tag{7}$$

Rastrigin test function:

$$f(x) = 10n + \sum_{i=1}^{n}\left(x_i^2 - 10\cos\left(2\pi x_i\right) + 10\right) \tag{8}$$

4.3 Comparative Algorithm Performance Evaluation

In the comparison of the two algorithms PSO and IPSO, I chose the Sphere test function as well as the Rastrigin test function from the 23 Benchmark benchmark functions [9]. And used the above three inertia weights for the performance comparison of PSO and IPSO, respectively, so that the resultant mean and variance obtained by running them independently for 10 times are as follows:

Table 2 PSO and IPSO Performance Comparison Data

| Func		PSO	IPSO（Linear）	IPSO（Expotential）	IPSO（Quadratic）
Sphere	Mean	6.3262e-01	1.2184E-10	9.7546E-14	9.372E-09
	Var	1.8646e-01	1.3233E-19	5.196E-26	1.1378E-16
Rastrigin	Mean	309.3884	124.5944	129.7789	133.0352
	Var	426.1845	916.549	1050.3886	996.97
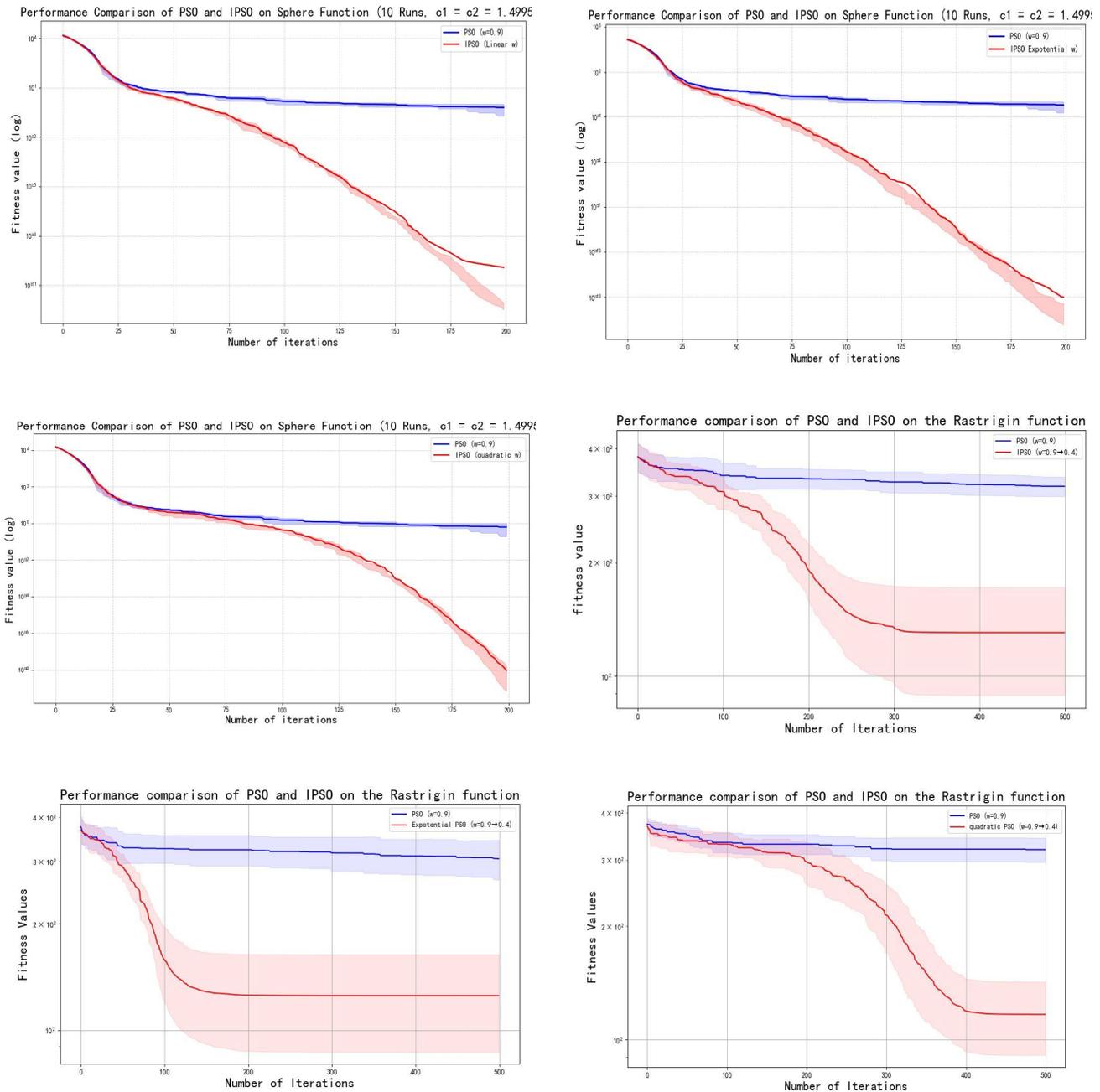
Figure 1 PSO and IPSO Performance Comparison Data

Based on the above six comparison graphs, it is clearly found that IPSO outperforms PSO regardless of the inertia weights and test functions used. The results of testing the performance of the IPSO algorithm using the Sphere test function as well as the Rastrigin test function are shown in Figure 1 as well as in Table 2, Figure1 presents the iterative curves of PSO and IPSO under Linear, Expotential, and Quadrati inertia weights, respectively, and Table 2 records the mean and the variance.

Based on the data and iterative curves in Table 2 and Figure1, it can be seen that the IPSO algorithm outperforms PSO in both the Sphere test function as well as the Rastrigin test function, with the mean values of IPSO (Linear) being 1.2184E-10, 124.5944, and the mean values of IPSO (Expotential) being 9.7546E-14, 129.7789, and the mean values of IPSO (Quadratic) are 9.372E-09, 133.0352, respectively, which are all better than the mean values of PSO, 6.3262e-01, 309.3884, and the minimum value of IPSO is improved by several orders of magnitude compared to PSO.

In summary, the performance of the IPSO algorithm using nonlinear iteration with inertia weights is improved compared to the PSO algorithm, proving that the improvements made on the traditional basis are effective.

4.4 Construction of Coupled Models
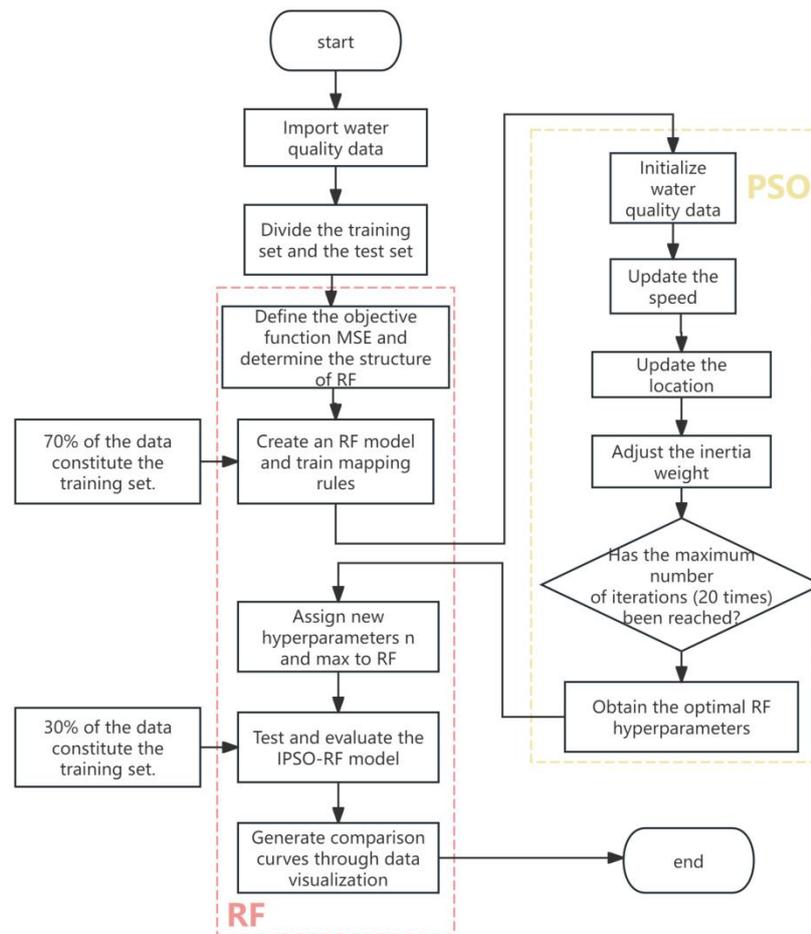
Figure 2 Flowchart of the Coupled Model

The logical structure of the IPSO-RF integrated algorithm constructed is shown in Figure 2. The specific steps are as follows:

Step 1: Import water quality data and normalize it;

Step 2: Divide the training set and test set. Use 70% of the data as the training set for subsequent training of the model, and 30% of the data as the test set for model evaluation;

Step 3: Determine the RF structure based on the input data, then create the basic RF model and train it;

Step 4: Implement the algorithm using the inertia-weight nonlinear iterative IPSO algorithm. According to the settings in the previous text, set relevant parameters and encode the three hyperparameters of RF, namely n_estimators, max_features, and max_depth;

Step 5: Use the mean square error of the sample data as the fitness function, and continuously iterate according to the IPSO algorithm process to update the optimal position and search for the optimal hyperparameter combination;

Step 6: Determine whether the maximum iteration times have been reached. If so, output the best hyperparameter combination to the random forest model and end the iteration. If not, return to Step 5;

Step 7: Assign the optimal n_estimators, max_features, and max_depth parameters of the RF model, construct a new IPSO-RF water quality prediction model, test whether the model accuracy meets the requirements, and output the classification results if it does. Otherwise, return to Step 3.

5 EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the water quality prediction model based on the IPSO-RF algorithm is constructed according to the above, and in order to verify that the model has a better prediction performance, it is validated against the PSO algorithm by comparing the mean square error (MSE) of the two algorithms [10].

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{9}$$

It is further demonstrated that the IPSO-RF model has better performance.

Among them, the parameter settings of the IPSO and PSO algorithms are consistent with those described above, and the population size and the number of iterations is uniformly set to 20 generations. The final model construction was carried out.

Among them, the final optimization results of the IPSO-RF integrated model for the parameters of random forest n and max are: n_estimators=16, max_depth=10, max_features=7; the specific visualization image of the mean square error is shown in Figure 3, and it can be seen that the MSE of PSO is converging to 0.021707, while the MSE of IPSO It is obvious that the mean square error of IPSO is smaller than that of PSO and its performance is more superior. Finally, it is proved that the IPSO-RF model constructed in this paper has certain theoretical and practical significance, and can make more accurate prediction of water quality composition.
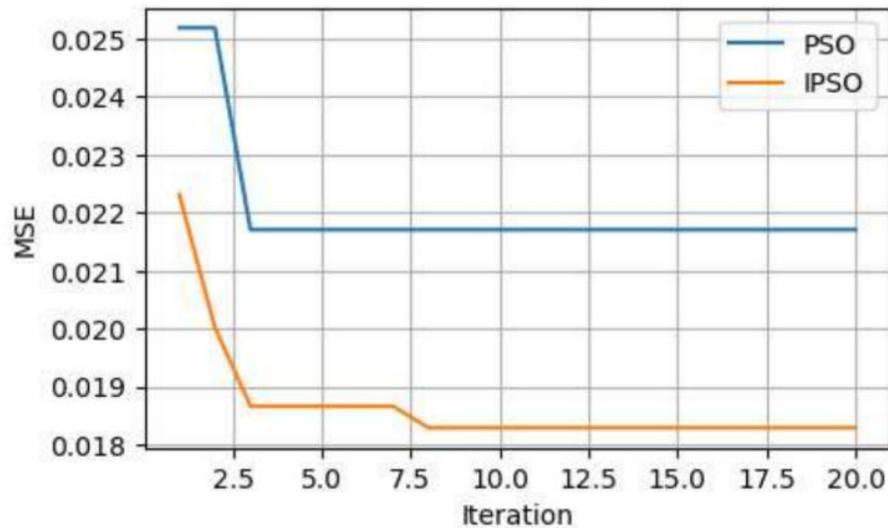
Figure 3 Water Quality Data IPSO and PSO Performance Comparison

6 CONCLUSIONS AND IMPLICATIONS

This study focuses on the construction of an intelligent water quality prediction model, demonstrating the advantages of machine learning models enhanced by intelligent optimization algorithms in the field of environmental monitoring. By improving the standard Particle Swarm Optimization (PSO) algorithm, we propose an IPSO algorithm that incorporates nonlinear iterative inertia weight adjustment. The optimization performance of IPSO and PSO is compared using two benchmark test functions, with experimental results confirming the superior convergence speed and accuracy of the IPSO algorithm. Furthermore, we integrate the IPSO algorithm with Random Forest (RF) to construct an IPSO-RF ensemble prediction model, which is applied to the prediction and analysis of key water quality indicators.

The findings of this study provide an effective technical solution for intelligent water quality monitoring, contributing to more accurate and stable environmental data analysis. The proposed model can assist in early warning systems for water pollution, support decision-making in water resource management, and promote sustainable environmental protection practices.

Innovation:

1. Algorithm Improvement: The introduction of nonlinear adaptive inertia weight in IPSO enhances optimization efficiency, addressing the limitations of standard PSO in local optima escape and convergence speed.

2. Model Integration: The novel combination of IPSO and RF improves prediction accuracy by optimizing hyperparameters and feature selection, offering a robust approach for water quality forecasting.

3. Practical Application: The IPSO-RF model demonstrates strong generalization capability in real-world water quality monitoring scenarios, providing a reliable tool for environmental big data analysis.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Rajaee T, Khani S, Ravansalar M. Artificial intelligence-based single and hybrid models for prediction of water quality in rivers: A review. Chemometrics and Intelligent Laboratory Systems, 2020, 200: 103978.
[2] Hmoud Al-Adhaileh M, Waselallah Alsaade F. Modelling and prediction of water quality by using artificial intelligence. Sustainability, 2021, 13(8): 4259.
[3] Chen Y, Song L, Liu Y, et al. A review of the artificial neural network models for water quality prediction. Applied Sciences, 2020, 10(17): 5776.
[4] Lu H, Ma X. Hybrid decision tree-based machine learning models for short-term water quality prediction. Chemosphere, 2020, 249: 126169.

[5]   Jain M, Saihjpal V, Singh N, et al. An overview of variants and advancements of PSO algorithm. Applied Sciences, 2022, 12(17): 8392.

[6]   Genuer R, Poggi J M, Genuer R, et al. Random forests. Springer International Publishing, 2020.

[7]   Deepa B, Ramesh K. Epileptic seizure detection using deep learning through min max scaler normalization. Int. J. Health Sci, 2022, 6: 10981-10996.

[8]   Choudhary S, Sugumaran S, Belazi A, et al. Linearly decreasing inertia weight PSO and improved weight factor-based clustering algorithm for wireless sensor networks. Journal of ambient intelligence and humanized computing, 2023: 1-19.

[9]   Chen T, Chen X, Chen W, et al. Learning to optimize: A primer and a benchmark. Journal of Machine Learning Research, 2022, 23(189): 1-59.

[10] Chicco D, Warrens M J, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. Peerj computer science, 2021, 7: e623.

REVEALING HETEROGENEITY OF N₂O CONCENTRATION IN SELECTIVE CATALYTIC REDUCTION DENITRIFICATION: A MULTI-POINT SAMPLING STUDY IN COAL-FIRED POWER PLANT

ZePeng Li[1], Yasser M. A. Mohamed[2*], YingHui Han[1*]

[1] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 101408, China.
[2] Photochemistry Department, National Research Centre, Dokki, Giza, P. O. 12622, Egypt.
*Yasser M. A. Mohamed and YingHui Han are both the Corresponding Authors.
Corresponding Authors: YingHui Han, Email: hanyinghui@ucas.ac.cn; Yasser M. A. Mohamed, Email: y.m.a.mohamed@outlook.com

Abstract: Selective catalytic reduction (SCR) plays a critical role in achieving ultra-low emissions from coal-fired power plants by enabling highly efficient removal of nitrogen oxides (NO_x). However, the denitrification process may concurrently generate nitrous oxide (N_2O), a potent greenhouse gas often overlooked in emission control strategies. Crucially, conventional continuous emission monitoring systems (CEMS) focus solely on NO_x reduction efficiency without tracking byproduct formation, creating a potential blind spot for N_2O emissions when reduction reactions are incomplete. To address this gap, this study conducted multi-point on-site sampling monitoring within SCR system of ultra-low emission units at a 693 MW coal-fired power plant and a separate 2× 300 MW plant. A pivotal question was investigated: Are N_2O emissions inevitably elevated by SCR operation? Our findings reveal significant spatial heterogeneity in N_2O distribution: while concentrations increased at certain locations (e.g., upstream catalyst layers), they unexpectedly decreased at others (e.g., optimized mid-reactor zones), demonstrating position-dependent synergistic reduction of NO_x and N_2O under specific operational conditions. This work highlights the necessity of multi-dimensional emission assessment beyond standard CEMS protocols and provides actionable insights for co-controlling air pollutants and greenhouse gases in industrial air pollution control systems.
**Keywords:** Nitrous oxide; Nitrogen oxide; Selective catalytic reduction; Ultra-low emission; Coal-fired power plant

1 INTRODUCTION

Air pollution is one of the most important challenging problems in China and Egypt, which have a major impact on increasing threat of development in all fields [1-2]. The Chinese government prioritizes ecological and environmental quality improvement, having implemented comprehensive "ultra-low emission standards" in the power generation sector to mitigate air pollution's adverse health impacts [3]. The main pollutants measured were sulfur Dioxide (SO_2), nitrogen oxides (NO_x), carbon monoxide (CO), ozone (O_3), and particulate matters with varied diameter (PM) [4]. It is known that various compounds exist in the nitrogen oxides depending on the bonding state of nitrogen and oxygen. Most of what is generated by combustion of fuels NO and NO_2, and is generally referred to as NO_x (Hereinafter referred to as NO_x) [5].

Figure 1 depicts a representative ultra-low emission control process for coal-fired power plants. Following emission from the boiler, flue gas undergoes sequential treatment: SCR for substantial NO_x abatement, followed by electrostatic precipitation (ESP) for particulate matter (PM) removal, and finally passes through a flue gas desulfurization (FGD) unit for sulfur dioxide (SO_2) elimination [6-7].

As the cornerstone NO_x control technology in ultra-low emission systems, SCR employs ammonia (NH_3) to catalytically reduce NO_x to nitrogen (N_2) in flue gas [8]. Nevertheless, SCR systems may simultaneously generate N_2O through competing reactions where NH_3 react with nitric oxide (NO) and oxygen (O_2) or with nitrogen dioxide (NO_2) [9-12]. This side-product formation warrants concern given N_2O's dual role as both a potent greenhouse gas and ozone-depleting substance. In addition, to address the challenge of climate change, increasing renewable energy integration has necessitated deep load-cycling operations in coal-fired power plants, inducing frequent boiler load fluctuations. Consequently, flue gas temperature variability alters the intrinsic activity of the SCR catalyst, potentially compromising denitrification efficiency while affecting N_2O formation pathways [13-15].
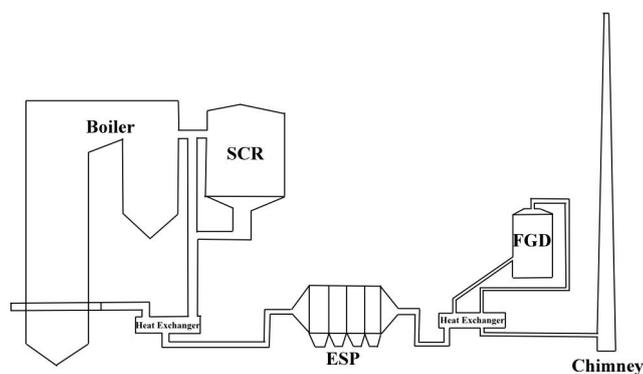
Figure 1 Flowchart of Ultra-Low Emission Technology in Coal-Fired Power Plants

Despite these operational complexities, systematic quantification of N_2O emissions from operational SCR systems remains scarce. Moreover, influential factors governing N_2O generation are inadequately characterized. This study bridges these knowledge gaps through field measurements quantifying N_2O concentrations at SCR inlets and outlets across multiple coal-fired plants. Furthermore, the correlative relationship between N_2O and NO_x concentrations in the flue gas were further examined to elucidate interactions mechanisms and evaluate co-control strategies for these interconnected pollutants.

2 EXPERIMENTAL METHODS

Figure 2 shows the field-deployed sampling setup used in the coal-fired power plant industrial environment. The vacuum pump operated at a flow rate of 10 L/min. The vacuum sampling box, with a total volume of 2 L, was connected to the vacuum inlet of the pump on its left side and to the sampler on its right side. A sampling bag was placed inside the box. The sampler is a steel tube approximately 2.5 meters in length, with a wall thickness of 2 mm and an outer diameter of 10 mm. One end of the steel tube was inserted into the sampling hole, while the other end was connected to the exterior of the vacuum sampling box.

Figure 2 In Situ Deployment of Plug-And Play Sampling Probes at Multi-Port SCR Monitoring Locations

The specific procedure for a single gas sample collection was as follows: First, with no sampling bag placed inside, the lid of the vacuum sampling box was securely closed, the pressure relief valve was tightened, and the vacuum pump was operated for approximately one minute to create a negative pressure environment. While keeping the vacuum pump running, the vacuum sampling box was then quickly opened, and the sampling bag was connected to the internal interface of the box. Subsequently, the valve of the sampling bag was opened, and the box was closed tightly again. Once the desired volume of gas had been collected, the vacuum pump was turned off, the pressure relief valve was loosened, and the vacuum sampling box was opened. Finally, the valve of the sampling bag was closed, and the bag was removed from the box.

3 RESULT AND DISCUSSION

Flue gas samples were collected from the SCR unit of a 693 MW coal-fired power plant located in China, featuring a supercritical once-through boiler with variable pressure operation and an integrated recirculating pump start-up system. Flue gas was collected into the gas bag from 10 sampling points located on both the A and B sides of the SCR reactor. N_2O concentrations were measured using gas chromatography. The collected flue gas samples were subjected to

detailed analysis and characterization using gas chromatography (GC). The test results are shown in **Figure 3**. In **Figure 3a**, the green line represents the test results from the A side, the blue line represents the test results from the B side, and the black dashed horizontal line indicates the approximate background concentration of atmospheric N_2O (340 ppb). Correlation between N_2O and NO_x concentrations in flue gas samples were presented in **Figure 3b**. NO_x concentrations were measured using nitrogen oxide analyzers. The red line represents the linear fit between the measured N_2O and NO_x concentrations.

As shown in **Figure 3a**, significant spatial heterogeneity in N_2O concentrations was observed between across sampling point on both Sides A and B of the SCR unit, with approximately 50% of locations exceeding atmospheric background levels. **Figure 3b** further reveals a strong liner correlation between N_2O and NO_x concentrations.
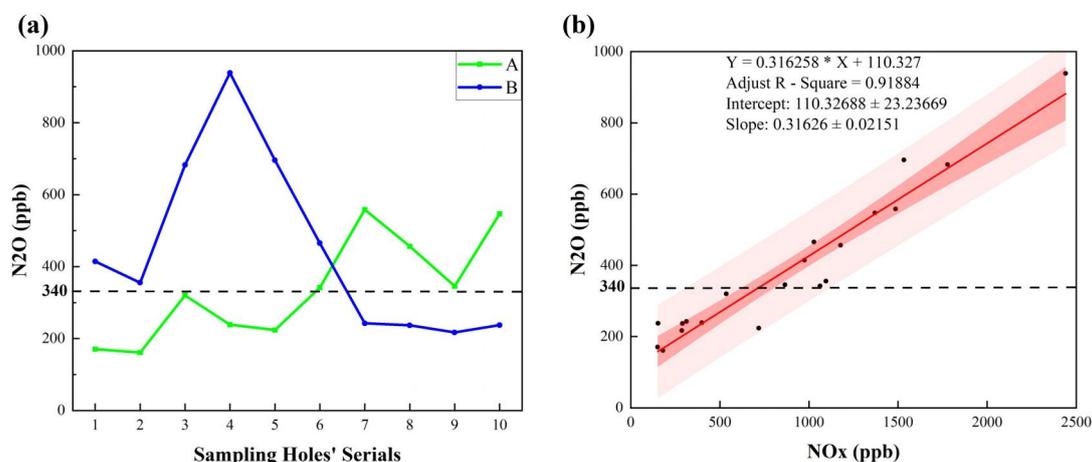
Figure 3 Flue Gas Sample Test Results. (a) N_2O Concentrations in Flue Gas From Different Sampling Holes

Complementing these findings, a parallel investigation was conducted at a coal-fired combined heat and power (CHP) facility (2 × 300 MW subcritical units). Flue gas samples were collected at the inlet and outlet of both Side A and Side B of the SCR unit associated with Unit #1 under two power generation loads: 210 MW and 105 MW. The results are summarized in Table 1. Key observations confirm consistent detection of both N_2O and NO_x at all sampling locations, alongside marked asymmetries in pollutant concentrations and removal efficiencies between SCR sides. Moreover, the concentrations of N_2O and NO_x emitted from the boiler were significantly higher under the higher power generation load compared to the lower load. Correspondingly, the removal efficiencies of these pollutants also varied with power load. Notably, the concentrations of N_2O and NO_x at the outlet were significantly lower than those at the inlet, unequivocally demonstrating SCR's capacity for simultaneous N_2O/NO_x abatement through synergistic removal mechanisms. This finding supports, to some extent, the results presented earlier.

Table 1 Flue Gas N_2O and NO_x Concentrations and Their Removal Efficiencies Under Different Power Generation Loads

Power Generation Load	Sides	Location	N_2O(ppb)	NO_x(ppb)	Removal Efficiency of N_2O (%)	Removal Efficiency of NO_x (%)
70%	A	Inlet	667	14745	87.56	88.75
70%	A	Outlet	83	1659		
70%	B	Inlet	598	13715	93.14	92.28
70%	B	Outlet	41	1059		
35%	A	Inlet	460	12710	84.78	77.07
35%	A	Outlet	70	2915		
35%	B	Inlet	288	11275	99.31	91.04
35%	B	Outlet	2	1010		

Note: "Power Generation Load" refers to the percentage of the actual load of the target unit during sampling relative to its designed maximum power generation capacity. Each installed SCR reactor has two sides, denoted in this study as "Side A" and "Side B". The terms "Inlet" and "Outlet" indicate the sampling locations: the former refers to the sampling port located upstream of the SCR reactor, while the latter refers to the port located downstream. "N_2O (ppb)" and "NO_x (ppb)" represent the relative concentrations of N_2O measured by gas chromatography and NO_x measured by a nitrogen oxide analyzer, respectively. The "Removal Efficiency" is calculated as: "Removal Efficiency" = ("Inlet Concentrations" – "Outlet Concentrations") / "Inlet Concentrations" × 100%, which applies to both N_2O and NO_x in this work.

4 CONCLUSION

In this work, a sampling-based measurement approach was employed to investigate the substantial spatial heterogeneity in N_2O concentration within ultra-low emission SCR systems while revealing their paradoxical capacity for simultaneous NO_x (88.75-91.04%) and N_2O (84.78-99.31%) abatement under optimized conditions, thereby establishing SCR's dual-function potential for carbon-pollutant synergy. These empirical insights advance co-governance strategies by demonstrating how existing pollution control infrastructure can mitigate greenhouse gases alongside conventional pollutants. Future research should integrate catalytic mechanism modeling with multi-plant field validation to develop sector-transferable frameworks applicable beyond power generation to energy-intensive industries like primary metal smelting and bulk chemical manufacturing.

COMPETING INTERESTS

FUNDING

REFERENCES

[1] Ghanem S, Ferrini S, Di Maria C. Air pollution and willingness to pay for health risk reductions in Egypt: A contingent valuation survey of Greater Cairo and Alexandria households. World Development, 2023, 172, 106373. https://doi.org/10.1016/j.worlddev.2023.106373.

[2] Zhao Bin, Wang Shuxiao, Hao Jiming. Challenges and perspectives of air pollution control in China. Frontiers of Environmental Science & Engineering, 2024, 18, 68. https://doi.org/10.1007/s11783-024-1828-z 2024.

[3] Feng Tong, Ma Jie, Yang Yuanjian, et al. Synergistic effects of air pollution control policies: Evidence from China. Journal of Environmental Management, 2025, 373, 123581. https://doi.org/10.1016/j.jenvman.2024.123581.

[4] Boraiy M, El-Metwally M, Wheida A, et al. Statistical analysis of the variability of reactive trace gases (SO_2, NO_2 and ozone) in Greater Cairo during dust storm events. Journal of Atmospheric Chemistry, 2023, 80, 227–250. https://doi.org/10.1007/s10874-023-09449-4.

[5] Jyethi D. Air Quality: Global and Regional Emissions of Particulate Matter, SO_x, and NO_x. Plant responses to air pollution, 2016, 5-19. https://doi.org/10.1007/978-981-10-1201-3_2.

[6] Elkaee S, Phule A, Yang J. Advancements in (SCR) technologies for NOx reduction: A comprehensive review of reducing agents. Process Safety and Environmental Protection, 2024, 184, 854-880. https://doi.org/10.1016/j.psep.2024.02.046.

[7] Di Natale F, Carotenuto C, Parisi A, et al. Wet electrostatic scrubbing for flue gas treatment. Fuel, 2022, 325, 124888. https://doi.org/10.1016/j.fuel.2022.124888.

[8] Zhu Minghui, Lai Junkun, Tumuluri U, et al. Nature of Active Sites and Surface Intermediates during SCR of NO with NH3 by Supported V_2O_5-WO_3/TiO_2 Catalysts. Journal of the American Chemical Society, 2017, 139(44): 15624–15627. http://doi.org/10.1021/jacs.7b09646.

[9] Busca G, Lietti L, Ramis G, et al. Chemical and mechanistic aspects of the selective catalytic reduction of NOx by ammonia over oxide catalysts: A review. Applied Catalysis B: Environmental, 1998, 18(1-2), 1-36. http://doi.org/10.1016/s0926-3373(98)00040-x.

[10] Song Xiaobo, Parker G, Johnson J, et al. A Modeling Study of SCR Reaction Kinetics from Reactor Experiments. SAE Technical Paper, 2013. https://doi.org/10.4271/2013-01-1576.

[11] Han Yinghui, Li Xiaolei, Zhao Yi. Abatement of SO_2-NO_x binary gas mixtures using a ferruginous highly active absorbent: Part I. Synergized effects and mechanism. Journal of Environmental Sciences, 2015, 30(1): 55-64.

[12] Devadas M, Kröcher O, Elsener M, et al. Influence of NO_2 on the selective catalytic reduction of NO with ammonia over Fe-ZSM5. Applied Catalysis B: Environmental, 2006, 67 (3-4), 187-196. http://doi.org/10.1016/j.apcatb.2006.04.015.

[13] Song Yubao, Liu Xinhui, He Chuan, et al. Study on the ammonium bisulfate deactivated SCR catalyst at low load operation. Electric Power, 2019, 52(1), 144-150.

[14] Wang Xin, Duan Rucheng, Li Zhuocan, et al. The critical role of oxygen vacancies in N_2O decomposition over cobalt-doped CeO_2 catalysts. Environmental Science & Technology. 2025, 59(11): 5839-5847.

[15] He Shuwei, Han Yinghui, Qin Xiaodong, Adavanced in accounting methodology of nitrous oxide emissions from the adipic industry. Journal Environmental Sciences. 2024, 157: 615-632

QUANTITATIVE STUDY ON METHANE EMISSION IN LIAONING PROVINCE BASED ON GAUSSIAN PLUME MODEL

YuChao Zhang

College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou 350028, Fujian, China.
Corresponding Email: zycfafu@163.com

Abstract: Methane, a critical greenhouse gas, significantly contributes to climate warming. This study quantifies methane emissions in Liaoning Province, China—a major industrial region—using TROPOMI satellite data and the Gaussian plume model. Filtered satellite observations and ERA5 meteorological parameters were integrated to retrieve emission rates and spatial patterns. Results indicate annual methane emissions of (720 ± 150) Gg yr^{-1}, with coal mining accounting for more than 60%. Seasonal analysis highlights elevated winter emissions linked to heating-related coal use. Satellite-derived emissions exceed EDGARv6.0 estimates by approximately 35%, suggesting underreported industrial sources, particularly mine ventilation leaks. The study demonstrates the effectiveness of combining remote sensing and dispersion modeling for regional methane monitoring. Key recommendations include enhancing emission inventories through high-resolution models and promoting methane recovery technologies in coal and steel sectors. These findings emphasize the need for targeted industrial emission controls to address climate and air quality challenges.
Keywords: Gaussian plume model; TROPOMI satellite data; Methane emissions; EDGAR

1 INTRODUCTION

As an important short-lived greenhouse gas, methane (CH_4) has a global warming potential about 28 times that of carbon dioxide (CO_2) on a 100 year scale, and its contribution to global warming can not be underestimated, second only to carbon dioxide[1]. In the atmosphere, methane not only affects the greenhouse effect through a series of complex photochemical reactions, but also participates in the formation of secondary pollutants such as tropospheric ozone, which has a negative impact on air quality. As an important heavy industry base in China, Liaoning Province has rich coal resources, and coal mining, steel production, chemical industry and other industries are booming. In terms of steel production, the crude steel output of Liaoning Province reached 70.682 million tons in 2024, accounting for 7.03% of the national crude steel output, ranking fourth in the country. As a large iron and steel enterprise in Liaoning Province, Angang Group's annual crude steel production capacity will reach 70million tons in 2024, accounting for 67% of the crude steel production in Liaoning Province and 51% of the crude steel production in Northeast China. In addition, the steel output of Liaoning Province in 2024 was 78.6768 million tons [2].While these industrial activities promote economic growth, they are also accompanied by a large number of methane emissions. In the process of coal mining, the methane in the coal seam will be released into the atmosphere with the coal mining; Steel production and chemical processes involve a variety of chemical reactions, which are also important sources of methane emissions [3]. These methane emissions have a significant impact on air quality and global climate change in Liaoning Province and even surrounding areas. Traditional bottom-up emission inventory preparation methods, such as Edgar (emissions database for Global Atmospheric Research), mainly rely on statistical data and emission factors [4]. This method has some limitations in practical application, it is difficult to accurately capture the dynamic characteristics of emission sources, and it is easy to omit some "super emission sources". Although the number of "super emission sources" is relatively small, the emission intensity is great, which has a significant impact on the total amount of methane emissions in the region and even the world. Therefore, it is of great practical significance to develop a more accurate and comprehensive methane emission monitoring and quantification method.

Based on TROPOMI satellite data and Gaussian plume model, a top-down quantitative method of methane emission is proposed in this study. This method aims to accurately identify the hot spots of methane emissions in Liaoning Province, quantify the emission rates of major industrial sources, clarify the contribution of different industrial activities to methane emissions, assess the consistency between the satellite retrieval results and the existing inventory, identify possible differences and analyze the reasons.

2 METHODOLOGICAL FRAMEWORK AND DATA ACQUISITION

2.1 Data Source and Pretreatment

TROPOMI data: This study obtained the XCH_4 data observed by TROPOMI satellite for some days from 2019 to 2021, with a spatial resolution of 49 km². To ensure the data quality, the observation data with quality mark (QA>0.8) and cloud fraction<0.2 were selected. The quality mark QA is used to measure the reliability of data. The higher the value of QA, the higher the reliability of data; Cloud cover will interfere with satellite observations, and low cloud cover data can more accurately reflect the methane column concentration on the ground. Through such screening, noise data and data affected by clouds can be effectively removed, and the accuracy of subsequent analysis can be improved [5]. The

following Figure 1 shows the spatial distribution of methane column average dry air mixing ratio in North China in a certain year using TROPOMI satellite data. It can be seen from the figure that the methane mixing ratio shows obvious spatial differences. The color bar displays data ranging from 1700 to 2100 (in 1E-9). The more red the color, the higher the methane mixing ratio, and the more blue the color, the lower the methane mixing ratio.
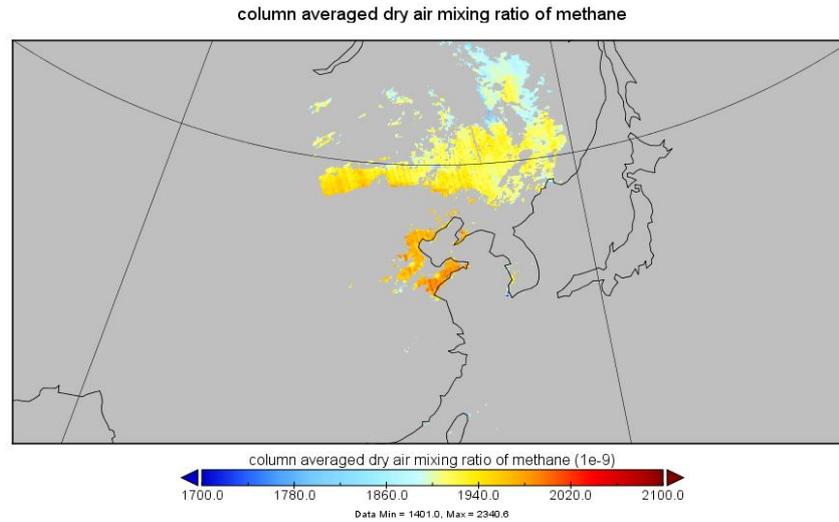
Figure 1 Data Visualization Display

Meteorological data: ERA5 reanalysis data is provided by the European Center for medium range weather forecasting (ECMWF), with high spatial-temporal resolution and extensive coverage of meteorological parameters. In this study, 10 m wind speed (U,V), boundary layer height and atmospheric stability classification are used. The wind speed and direction determine the transmission direction and diffusion speed of methane in the atmosphere; The height of the boundary layer affects the diffusion range of methane in the vertical direction; The classification of atmospheric stability is closely related to the diffusion degree of methane. Under different stability conditions, the diffusion parameters of methane are different. These meteorological data are very important for accurately simulating the diffusion process of methane.

Location of emission sources: in order to determine the key research areas, Fushun and Fuxin coal mine groups and other typical areas were selected in combination with the global energy Observatory (GEO) and the coal mine distribution map of Liaoning Province. The global energy Observatory provides basic information about energy related facilities around the world, while the coal mine distribution map of Liaoning Province shows the distribution of coal mines in the province in detail. Through the combination of the two, the geographical location of coal mines can be accurately determined, which provides the basis for the subsequent research on methane emission in these key areas.

2.2 Gaussian Plume Model Construction

Assuming that the diffusion of methane obeys the Gaussian distribution under steady-state conditions, its concentration enhancement C (x, y) can be expressed as:

$$C(x, y) = \frac{Q}{2\pi \ \sigma_y \sigma_z \mu} \exp\left(-\frac{y^2}{2\sigma_y^2} - \frac{(z-H)^2}{2\sigma_z^2}\right) \tag{1}$$

Where, Q is the emission rate (kg·h⁻¹), which is the key parameter to be retrieved in the model and reflects the mass of methane emitted by the emission source per unit time; σ_y and σ_z are horizontal and vertical diffusion parameters respectively, and their values are determined by Pasquill - Gifford stability classification [6]. The values of these two parameters are also different with different atmospheric stability, which determines the diffusion degree of methane in the horizontal and vertical directions; U is the wind speed (m·s⁻¹). The greater the wind speed is, the farther the transmission distance of methane is and the faster the diffusion speed is; H is the effective height of emission source, which can be assumed to be 50m for coal mine ventilation shaft. This height is determined by comprehensively considering the general building height of coal mine ventilation shaft and the actual position of methane emission outlet and other factors. Through this formula, the concentration distribution of methane in space can be described under given meteorological conditions.

2.3 Emission Rate Inversion

Deduction of background concentration: in order to accurately obtain the change of methane concentration caused by emission sources, the background concentration needs to be deducted. In this study, the wind direction in the Yellow

Sea (38°N-40°N, 120°E-122°E) to calculate the background XCH₄ concentration. The region is far away from the main methane emission source in Liaoning Province, and its methane concentration can represent the natural background level in the atmosphere. By subtracting the background concentration from the observed data, a more accurate enhancement of methane concentration caused by local emission sources can be obtained.

Plume identification: extract the methane enhancement along the wind direction section, and carefully eliminate the interference from other sources. During the growth of rice fields, methane emissions will occur, and the decomposition of organic matter in landfills will also release methane. These non target emission sources will interfere with the analysis of methane emissions from industrial sources. Through the identification and elimination of emission source types, combined with geographic information and related field knowledge, these interference factors are removed, so as to extract the methane plume emitted by industrial sources more accurately.

Least squares optimization: the least squares optimization method is used to minimize the mean square error between the predicted value of the model and the observed value through iterative solution, so as to obtain the best Q value. In the iteration process, the value of emission rate Q is continuously adjusted to calculate the mean square error between the methane concentration predicted by the model and the actual observed concentration. When the mean square error reaches the minimum, the corresponding Q value is the emission rate retrieved [7]. This method can make full use of the observation data and improve the accuracy of emission rate inversion.

2.4 Uncertainty Analysis

Sensitivity of meteorological parameters: the uncertainty of meteorological parameters will affect the emission rate inversion results. This study can test the influence of wind speed error ($\pm 20\%$) and stability classification deviation on Q value. The change of wind speed will directly affect the diffusion speed and transmission distance of methane. When there is error in wind speed, the diffusion range and concentration distribution of methane in the model will change, which will affect the inversion results of emission rate. The deviation of stability classification will also affect the values of diffusion parameters σ_y and σ_z, leading to the deviation of the model in the simulation of methane diffusion, and ultimately affecting the accuracy of Q value.

Selection of background area: the selection of background area also has a certain impact on the calculation results of emission rate. The research process can compare the emission differences caused by different background ranges. Here, Taking Bohai Bay and Inner Mongolia grassland as examples, their relative positions are different from those of Liaoning Province, and their atmospheric background methane concentrations are also different. Selecting different background areas to calculate the background concentration, the enhancement of methane concentration after deducting the background will be different, resulting in differences in the calculated emissions. By comparing the results of different background regions, we can evaluate the Uncertainty Impact of background region selection on the research results.

3 EMISSION QUANTIFICATION AND COMPARATIVE ANALYSIS

3.1 Emission Hot Spot Identification

Through the analysis of the data, it can be found that the high value area of methane enhancement is mainly concentrated in Fushun (123.8°E, 41.8°N), Fuxin (121.6°E, 42.0°N) coal mine cluster and Anshan Iron and steel industrial zone. Fushun and Fuxin coal mine groups are important coal producing areas in Liaoning Province. Coal mining activities are frequent, and a large amount of methane is released from the coal mining process. Anshan Iron and steel industrial zone has become an important area for methane emission due to chemical reactions in the process of steel production.

The study also found that the methane emission in Liaoning Province showed obvious seasonal characteristics of high in winter and low in summer. This is closely related to the enhancement of coal-burning activities in the heating season. In winter, in order to meet the heating demand, the amount of coal combustion increases significantly, which not only increases the methane emission in the process of coal mining, but also produces a certain amount of methane emission in the process of coal-fired heating. In summer, the demand for heating is reduced, the use of coal is reduced, and methane emissions are also reduced.

3.2 Quantification of Emission Rate

Fushun mining area, as the main coal producing area in Liaoning Province, plays an important role in the emission of methane in Liaoning Province. The ventilation system in the process of coal mining will discharge the methane in the coal seam. At the same time, if there is leakage in the process of gas drainage, it will also cause a large amount of methane to be discharged into the atmosphere. Accurate quantification of methane emissions from Fushun mining area plays a key role in evaluating the contribution of coal mining industry to methane emissions in the province.

Comparing the satellite retrieval results with the EDGARv6.0 list, it can be found that the satellite retrieval results will be overestimated by about 35% compared with the EDGARv6.0 list,corresponding to an EDGARv6.0 baseline estimate

of approximately 530 Gg yr⁻¹ for total provincial methane emissions. It can be inferred that the specific reason is mainly due to the mine ventilation and gas drainage leakage not reported in the existing list. In the traditional emission inventory preparation process, due to the limitations of data acquisition and insufficient understanding of some emission processes, some emission sources may be omitted or the emission intensity may be underestimated. The results of this study reveal the potential problems of the existing inventory in the quantification of industrial sources, and provide a reference for the subsequent improvement of the emission inventory compilation method. At the same time, relevant studies have pointed out that there is a deviation between the satellite retrieval results and the emission inventory, in which the satellite retrieval results are higher than the inventory emission intensity as a whole, the straight-line slope across the origin is 1.08, and the Pearson correlation coefficient is 0.7 [8].

3.3 Sources of Uncertainty

In the process of simulating methane diffusion, the Gaussian plume model ignores the influence of terrain disturbance and unsteady meteorological conditions, while the Gaussian plume model assumes flat terrain and steady meteorological conditions, which makes the simulated methane diffusion in these areas deviate from the actual situation and may lead to underestimation of emissions. In addition, the meteorological conditions in the actual atmosphere are constantly changing, and the diffusion law of methane under unsteady meteorological conditions is more complex. Gaussian plume model is difficult to accurately describe, which will also affect the accuracy of emission rate inversion. These are the main causes of uncertainty.

Of course, the quality of the data will also slightly affect the uncertainty of the results. The pixel resolution of TROPOMI satellite data is 49 km², which means that one pixel may cover multiple emission sources. In this case, the contributions of different emission sources will mix with each other, leading to the fuzziness of point source contributions. For some small-scale point sources with high emission intensity, their emissions may not be accurately identified and quantified under the existing resolution, which will affect the accurate assessment of the total amount and distribution of regional methane emissions.

4   CONCLUSIONS

This study confirmed the applicability of TROPOMI satellite and Gaussian plume model in regional methane emission monitoring. Through this method, the hot spots of methane emission in Liaoning Province can be identified, the emission rates of major industrial sources can be quantified, and the existing emission inventories underestimate the contribution of industrial sources. In order to further improve the accuracy and refinement of methane emission monitoring, the following measures need to be taken in the future:

Build a comprehensive observation network by combining satellite observation, ground monitoring stations, UAVs and other observation means, such as the functions of sentinel-5p satellite and ground sensors. Ground monitoring stations can provide high-resolution near ground methane concentration data, while UAVs can flexibly observe in specific areas to obtain more detailed emission source information. Through the fusion of multi platform data, the shortage of single observation method can be made up, and the monitoring accuracy of methane emission can be improved.

Increase the R&D and promotion of methane recovery technology for key industries such as coal mine and steel. For example, in the process of coal mining, more advanced gas extraction and utilization technologies can be adopted to use the extracted methane for power generation, heating, etc., which not only reduces methane emissions, but also realizes the comprehensive utilization of energy. In the process of steel production, the production process should be optimized to reduce the generation and emission of methane. At the same time, the Liaoning provincial government can refer to Shanxi coal mine gas power generation subsidy policy for the formulation and implementation of relevant programs.

Improve the preparation method of emission inventory based on the problems of existing emission inventory found in this study. Strengthen the research on the emission process of industrial sources, improve the determination method of emission factors, and give full consideration to unreported emission sources and leakage. At the same time, technical means such as big data and machine learning are used to improve the accuracy and timeliness of emission inventory preparation.

In view of the limitations of Gaussian model and TROPOMI satellite data resolution, research on high-resolution model is carried out. Combined with more refined terrain data and meteorological data, develop a model that can more accurately describe methane diffusion under complex terrain and unstable meteorological conditions, improve the simulation ability of small-scale emission sources, and provide more powerful technical support for the precise control of regional methane emissions.

Through the implementation of the above measures, it is expected to further improve the understanding and control ability of methane emissions in Liaoning Province, and contribute to coping with climate change and improving regional air quality.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Pang Xinrui. Analysis of methane emission control and emission reduction mechanism under the global carbon trading system. International Institute of green finance, Central University of Finance and economics, 2024.

[2] National Bureau of Statistics. Statistical Communique of the People's Republic of China on the National Economic and Social Development. stats.gov.cn, 2024.

[3] Fuchurui. Influencing factors and utilization ways of methane emission in coal industry. Coal Quality Technology, 2024, 39(4): 40-47.

[4] Liqingqing, Suying, Shang Li, et al. Comparative analysis of China's carbon emissions accounting based on international typical carbon database. Progress in Climate Change Research, 2018, 14(3): 275-280.

[5] Sadavarte P, Pandey S, Maasakkers J D, et al. Methane Emissions from Superemitting Coal Mines in Australia Quantified Using TROPOMI Satellite Observations. Environmental Science & Technology, 2021, 55(22): 16573−16580.

[6] Micallef A, Micallef C. The Gaussian Plume Model Equation for Atmospheric Dispersion Corrected for Multiple Reflections at Parallel Boundaries: A Mathematical Rewriting of the Model and Some Numerical Testing. Sci, 2024, 6(3): 48.

[7] Wangshuying, Gao Yongsheng. The mathematical principle and analogy analysis of the least square method and its related methods. Hans Press, 2024.

[8] Zhuwentao. Retrieval of carbon dioxide emissions from Chinese cities based on satellite remote sensing. Tsinghua University Thesis Service System, 2024.