Pol. J. Environ. Stud. Vol. XX, No. X (XXXX), 1-16
DOI: *10.15244/pjoes/211642* ONLINE PUBLICATION DATE: 2026-03-26

Original Research

Optimizing Real-Time $PM_{2.5}$ Predictions with Deep Belief Networks: Synergistic Role of Kriging Interpolation

Jin Xiao[1], Chao Bian[1, 2O*], Guangqiu Huang[2]

[1]Yinchuan University of Science and Technology, Ningxia 710021, China
[2]Xi'an University of Architecture & Technology, Xi'an 710055, China

Received: 31 March 2025
Accepted: 02 October 2025

Abstract

The aim of this study was to improve the accuracy and stability of fine particulate matter ($PM_{2.5}$) concentration prediction by integrating a deep belief network (DBN) and the kriging interpolation algorithm. First, we utilized a binary particle swarm optimization algorithm to construct the evaluation model, thereby identifying the main factors that most significantly influence the prediction of the $PM_{2.5}$ concentration. Second, we thoroughly analyzed the relationship between the real-time $PM_{2.5}$ concentration in Xi'an City during different months (or days) and various traditional variograms to obtain training samples for DBN. Finally, on the basis of the selected functions and sample data, we obtained corresponding kriging model interpolation results. The experimental results showed that, compared with the prediction errors of the traditional, spatiotemporal, and empirical kriging interpolation methods, the proposed method yielded a prediction error of only 4.677%. This error was calculated using the Mean Absolute Percentage Error (MAPE), which reflects the average percentage deviation between predicted and actual $PM_{2.5}$ concentrations. MAPE was chosen to facilitate comparison with previous studies in air quality prediction. This method increases not only the accuracy of the interpolation algorithm but also its robustness. The core contribution of this study is the introduction of a DBN to optimize the kriging interpolation algorithm, significantly improving the accuracy and stability of $PM_{2.5}$ concentration prediction. We combined climate and meteorological data with a deep learning technique to enable the model to adaptively select the most appropriate variogram, thereby more accurately simulating the air quality during different seasons and in various environments. In summary, this study provides a new and effective method for accurate air quality simulation, which has significant implications for advancing meteorological research and improving environmental pollution conditions.

Keywords: kriging interpolation algorithm, binary particle swarm optimization algorithm, environmental air quality simulation, $PM_{2.5}$ concentration prediction, deep belief network (DBN)

*e-mail: 15526387@qq.com
°ORCID iD: 0000-0003-2376-5025

Introduction

$PM_{2.5}$ has emerged as one of the most critical air pollutants threatening public health and environmental sustainability. On October 17, 2013, the World Health Organization officially classified fine particulate matter ($PM_{2.5}$) as a Group 1 carcinogen. With an aerodynamic diameter of ≤2.5 μm, $PM_{2.5}$ can remain suspended in the atmosphere for extended periods and penetrate deep into the human respiratory system, leading to severe respiratory and cardiovascular diseases. Consequently, the accurate prediction and effective control of $PM_{2.5}$ concentrations have become a focal point in atmospheric science and public health research.

Despite efforts to monitor air quality through fixed ground monitoring stations, the uneven spatial distribution and limited coverage of these stations hinder the precise characterization of regional $PM_{2.5}$ distribution patterns. This limitation has prompted researchers to explore interpolation algorithms as essential tools for estimating $PM_{2.5}$ concentrations across both spatial and temporal dimensions [1]. These methods, especially when combined with machine learning techniques such as random forests and deep learning, have demonstrated improved accuracy and robustness in complex pollution scenarios [2].

Kriging interpolation, widely used since the 1950s, is a geostatistical method that captures spatial correlations but often neglects temporal variability. To address this gap, recent studies have proposed integrating kriging with deep learning to better model spatiotemporal dynamics. For instance, Samal et al. [3] developed a hybrid framework combining CNN, BiGRU, and ANN with ordinary kriging, achieving high performance in regional air pollution prediction. Similarly, Nag et al. [4] proposed a two-stage framework using deep neural networks, with LSTM and ConvLSTM modules for temporal forecasting. Wu et al. [5] introduced a GTWR-XGBoost hybrid model, effectively incorporating both spatial heterogeneity and nonlinearity, demonstrating significant improvements in prediction accuracy across multiple datasets. These works underscore the potential of combining statistical and AI-based models to enhance predictive performance.

In parallel, many empirical studies have investigated the complex interactions between $PM_{2.5}$ and meteorological or topographic factors. Xu et al. [6] analyzed monthly and seasonal trends in Bijie and established correlations between pollutant levels and meteorological conditions. Kong et al. [7] reported that during the heating season in Shenyang, temperature and humidity are positively correlated with $PM_{2.5}$, while wind speed is negatively correlated. Zhang et al. [8] revealed seasonal variability in the chemical composition of $PM_{2.5}$ in Huainan. In a more recent study, Lu et al. [9] applied random forest algorithms and satellite imagery to simulate spatiotemporal $PM_{2.5}$ variation in the Central Plains Urban Agglomeration, demonstrating the advantages of data-driven models in high-resolution air quality estimation.

Furthermore, studies such as Xu et al. [10], Tang et al. [11], and Du et al. [12] provide robust evidence of spatial and temporal $PM_{2.5}$ variations across different provinces in China. These investigations strengthen the understanding of regional pollution dynamics and highlight the need for location-specific prediction models.

Recent advancements in machine learning have further propelled $PM_{2.5}$ prediction research. Zhao et al. [13] introduced a CNN-GRU-LSTM architecture that significantly improved prediction accuracy. Rakholia et al. [14] utilized ensemble learning methods to adapt to evolving target distributions in urban air quality forecasts. Yue et al. [15] evaluated global $PM_{2.5}$-related mortality trends and called for improvements in both pollution control and healthcare systems to meet SDG3.9 goals. The study by Ding et al. [16] provided a mechanistic link between $PM_{2.5}$ exposure and cardiovascular disease through a novel adverse outcome pathway model. At the policy level, Prosdocimi et al. [17] assessed the effectiveness of air quality regulations in Venice, while Chiesa and Vigliotti [18] compared mechanical ventilation strategies for indoor air quality in Italian schools. Zhou et al. [19] comprehensively reviewed deep learning models for $PM_{2.5}$ prediction, and Yu et al. [20] proposed a spatial-weighted EMD-LSTM model for short-term forecasting, further enriching the model landscape.

In addition to the aforementioned approaches, several recent studies provide valuable insights that complement our work. For example, Abirami and Chitra [21] proposed a spatiotemporal deep learning model for regional air quality forecasting, achieving high prediction accuracy. Simeonov et al. [22] conducted a multivariate statistical assessment of air quality and demonstrated the effectiveness of dimensionality reduction in capturing pollution dynamics. Furthermore, Srisang et al. [23] explored the calibration of $PM_{2.5}$ IoT sensors and addressed implementation issues using machine learning techniques. These studies reinforce the importance of integrating statistical, spatial, and intelligent methods in building robust air pollution prediction systems.

In this study, the use of a deep belief network (DBN) in conjunction with kriging interpolation was proposed to predict the $PM_{2.5}$ concentration more accurately. Compared with traditional methods, the proposed method comprehensively considers climatic and seasonal factors and other important factors influencing $PM_{2.5}$ levels. First, we selected factors that strongly influence the $PM_{2.5}$ concentration and incorporated them into the sample set of the deep neural network model. Second, we selected the $PM_{2.5}$ concentration in Xi'an City in different months (or days) and the prediction accuracy of the kriging model with different classical variograms as the research objective to construct the sample set of the deep neural network. Finally,

on the basis of the selected functions and sample data, we obtained the corresponding kriging interpolation results. The proposed method demonstrated the superiority of DBNs in feature extraction and pattern learning. When combined with kriging interpolation, the proposed method can be used to more effectively correct interpolation errors and accurately capture the nonlinear relationships in the data. In general, compared with traditional kriging interpolation techniques, the proposed method provides more reliable and highly generalizable interpolation results.

Material and Methods

Key Theories for Real-Time $PM_{2.5}$ Prediction

Increasing the $PM_{2.5}$ Prediction Accuracy Via Binary Particle Swarm Optimization

Increasing the $PM_{2.5}$ prediction accuracy is a key objective in the selection of crucial factors. The presence of an excessive number of factors can trigger the dimensionality problem, resulting in a significant increase in computational complexity. The factors influencing $PM_{2.5}$ include primarily anthropogenic and natural sources of pollution, with countless subfactors within these categories. Therefore, reducing the dimensionality of features by selecting the most important factors is essential. In addition, the spatial resolution of the $PM_{2.5}$ dataset is constrained by the distribution of urban monitoring stations in Xi'an. The average inter-station distance is approximately 5-7 km, which provides moderate coverage for city-scale interpolation. While this resolution may limit granularity, the proposed DBN-kriging approach enhances adaptability to spatial heterogeneity by learning from multiscale environmental contexts.

The degree of association is a measure of the degree of correlation between factors in two systems, varying over time or across different entities. During the evolution of systems, if two factors exhibit consistent trends, i.e., a high degree of synchronous variation, their association is considered strong; otherwise, it is considered weak. Kennedy and Eberhart developed the binary particle swarm optimization (BPSO) method for discrete combinatorial optimization problems, such as the knapsack problem, economic planning, and graph and image processing, which achieved notable results. The degree of association can be computed by selecting a few features through the BPSO algorithm and analyzing them with an established response surface model on the basis of the outputs.

In the BPSO algorithm, particles comprise binary codes, with each binary digit indicating whether a specific feature factor is selected. The binary positions are updated via a velocity iteration equation, as shown in Equation (1), after which the velocity is converted to a probability, with a value of 0 or 1 at each position, as described in Equation (2). This is the most significant difference between the BPSO algorithm and traditional particle swarm optimization algorithms.

$$s_i = \frac{1}{1+\exp(-v_i)} \tag{1}$$

$$x_i = \begin{cases} 1 & rand < s_i \\ 0 & rand \geq s_i \end{cases} \tag{2}$$

In this case, the response surface model function is denoted as $y = f(x)$. The optimization objective of the PSO algorithm can be expressed as $rmse(f(x) - y_0)$, where y_0 is the actual output used to construct the response surface model. A flowchart of the feature selection process using the PSO algorithm is shown in Fig. 1.

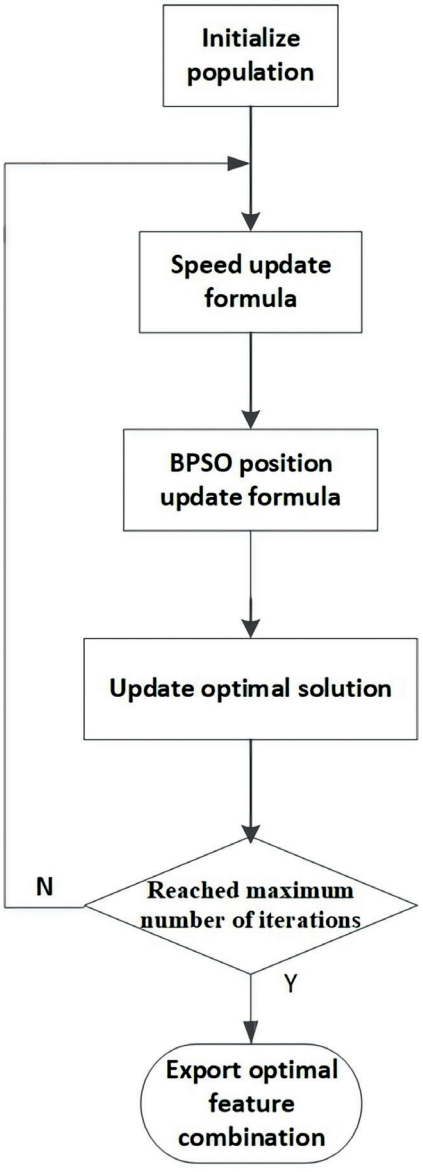

Fig. 1. Particle swarm optimization feature selection algorithm flowchart.

To evaluate the performance of different feature subsets during the BPSO process, a fitness function combining feature relevance and redundancy was designed. The fitness function is expressed as:

$$\text{Fitness} = \alpha \cdot \text{Accuracy} + (1-\alpha)\cdot(1-\text{Redundancy}) \quad (3)$$

where "Accuracy" reflects the predictive performance of the model using the selected features, and "Redundancy" represents the average mutual information (MI) among the selected features. The weight parameter α was set to 0.7 to prioritize informative features while penalizing overlap. This formulation ensures the selected feature subset maximizes predictive relevance while minimizing information redundancy, thus enhancing generalization and computational efficiency.

Kriging Algorithm

As a core technique in geostatistics, kriging interpolation is widely used for spatial data analysis and model construction. This method is also commonly referred to as a spatial local interpolation method, in which the attribute values at unknown points within a given range of sampling points can be estimated via spatial correlation range analysis.

This method is based on variogram theory and structural analysis, which enable it to provide unbiased and optimal estimates of regionalized variables within a limited area. In other words, this method can effectively reveal and account for the statistical properties of spatial data, offering solutions to complex geographical and environmental problems.

Assume that $Z(x)$ is a second-order stationary random function, with values of $Z(x_1)$, $Z(x_2)$, ..., $Z(x_n)$, at n positions. The estimate at point x_0 via the kriging algorithm is as follows:

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \quad (4)$$

For the above Equation to hold, $E[Z(x)] = m$ and $\sum_{i=1}^{n} \lambda_i = 1$ must also hold.

A variogram is a mapping function that characterizes the similar relationships among the features of samples in interpolation data. The quality of the variogram settings directly affects the interpolation effectiveness of the kriging algorithm. The variogram can be expressed as Equation (5). Variogram models are generally divided into those with sills and those without sills.

$$\gamma(x_i, x_j) = \frac{1}{2} E[Z(x_i) - Z(x_j)]^2 \quad (5)$$

Currently, classical spherical functions, exponential functions, and Gaussian functions can be used as variogram models with sills. The spherical function model is given by Equation (6), where a is the range parameter, C_0 is the nugget constant, and C is the sill parameter:

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 + C\left(\frac{3h}{2a} - \frac{h^3}{2a^2}\right) & h \in (0, a] \\ C_0 + C & h > a \end{cases} \quad (6)$$

DBN

A DBN is a type of deep neural network proposed by Hinton in 2006. This network is based on the greedy layer-by-layer pretraining strategy of a restricted Boltzmann machine (RBM), which greatly increases the training efficiency and largely alleviates the local optimum problem. A DBN comprises multiple RBMs and neural network layers and has been widely considered one of the deep neural network models with the best prediction performance. The DBN training process involves the unsupervised learning of RBMs and the supervised learning of the neural network. In addition, there is a transfer learning relationship between RBMs and neural network layers, which is an important difference between DBNs and other neural networks. This structure greatly improves the learning efficiency of the neural network, facilitates the effective extraction of the key features of the data, and avoids overfitting, as shown in Fig. 2.

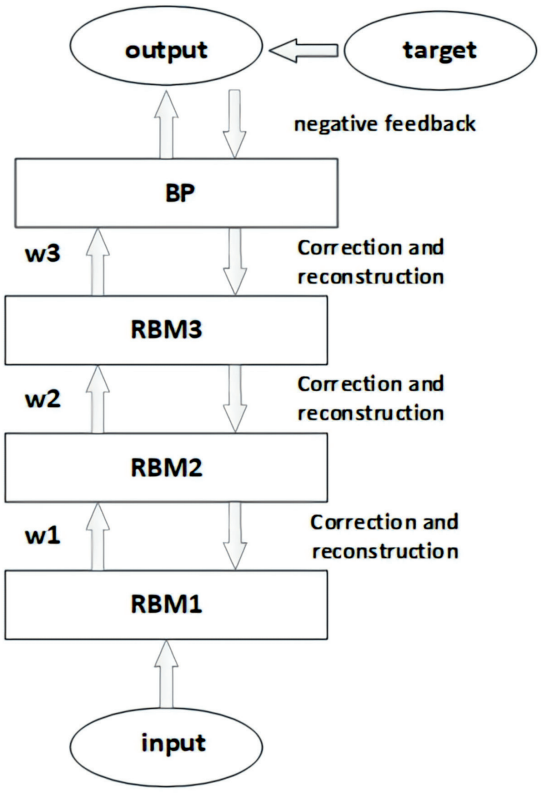

Fig. 2. Flowchart of the DBN algorithm.

Construction of a Real-Time $PM_{2.5}$ Prediction Model Optimized by a DBN

Data Sources

To evaluate the performance of the algorithm, we obtained real-time atmospheric environment monitoring data from the National Urban Air Quality Real-Time Release Platform of the Ministry of Environmental Protection. We selected daily air quality data for Xi'an City from January 2020 to January 2021 as the research object, and the relevant data are shown in Table 1. At the data processing stage, we used a dimensional normalization technique to eliminate dimension and scale differences among the various data types and to ensure the consistency and comparability of the data in the subsequent analysis, thus providing more sound and reliable research results. To ensure robust model evaluation and prevent overfitting, we adopted a two-fold validation strategy. Temporally, the dataset was split chronologically, using 70% of the data for training and the remaining 30% for testing, with the testing period strictly following the training period to avoid temporal data leakage. Spatially, we implemented a leave-one-station-out cross-validation approach: for each iteration, one monitoring site was held out for testing, while the remaining sites were used for training. This spatiotemporal validation framework ensures that the model's performance is assessed under realistic prediction scenarios and enhances its generalization capability.

Factors Influencing the $PM_{2.5}$ Concentration

Xi'an has a warm, temperate, subhumid, and continental monsoon climate, and a relationship may exist between the climatic characteristics and the $PM_{2.5}$ concentration in this area. In 2020, the annual average temperature in Xi'an fluctuated between 13.6ºC and 15.3ºC, with distinct climate features during each season. In winter, the environment is cold with weak winds, frequent haze, and relatively low precipitation. In spring, the temperature is comfortable, the environment is dry, and strong winds and variable weather patterns occur. The hot summer season is characterized by abundant rainfall, frequent thunderstorms, and notable dry conditions. Autumn is characterized by cool weather, a sudden decrease in temperature, and occasional rain. Therefore, in cities such as Xi'an, i.e., those with distinct seasonal climate characteristics, selecting parameters related to the climate during each season for in-depth analysis, which has been explored in detail, is crucial. For example, Zhang et al. [24, 25] proposed the use of the gray relational analysis method to measure the importance of various relevant factors for selecting key features. On this basis, several features with high correlations with the $PM_{2.5}$ concentration were selected.

The factors influencing $PM_{2.5}$ concentrations, as shown in Fig. 3, can be measured via two criteria: gas components and physical climate characteristics, namely, $X = \{Y_1, Y_2\}$. For each criterion, multiple indicators can be used for analysis. The gas components are SO_2, NO_2, CO, and O_3, represented by $Y_1 = \{Z_1, Z_2, Z_3, Z_4\}$, and the climatic factors are temperature, humidity, atmospheric pressure, rainfall, and wind speed, represented by $Y_2 = \{Z_5, Z_6, Z_7, Z_8, Z_9\}$.

Table 1. Selection of sample data from Xi'an City.

No.	SO_2 (μg/m³)	NO_2 (μg/m³)	CO (mg/m³)	O_3 (μg/m³)	Temperature (°C)	Relative Humidity	Atmospheric pressure (hPa)	Rainfall (mm)	Wind speed (m/s)	$PM_{2.5}$ (μg/m³)
1	17	50	1.396	9	3	0.55	950	0	1.2	75
2	28	50	1.886	2	4	0.41	956	0	2.3	165
3	17	57	1.578	5	5	0.42	930	0	0.8	115
4	22	70	1.427	5	2	0.49	999	0	1.6	126
5	17	53	1.284	4	0	0.56	968	3.2	3.2	99
6	20	51	1.214	5	-1	0.62	966	5.1	4.6	125
7	17	62	1.777	4	-2	0.66	949	7.8	5.8	98
8	11	45	1.407	2	0	0.68	920	1.6	3.6	82
9	20	56	1.542	3	1	0.52	988	0	1.9	107
:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:
N	16	52	0.979	7	2	0.41	976	0	1.8	100

Note: The "No." column represents the sample number or the index of the observational data. Relative humidity is reported as a fraction (0–1).

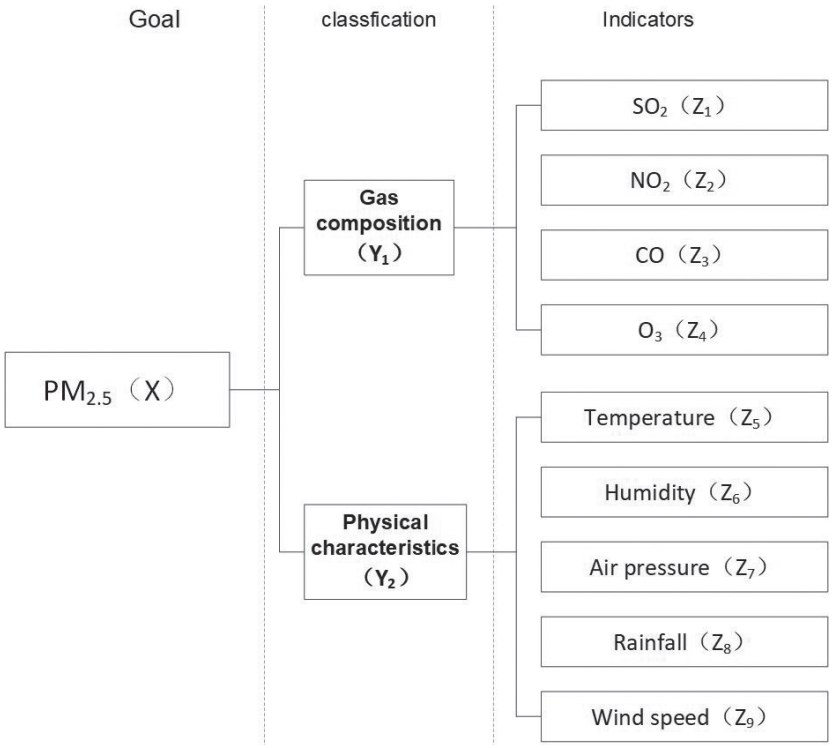

Fig. 3. $PM_{2.5}$-related factors.

After the application of the BPSO method, the kriging algorithm, as a type of interpolation method, can be used to construct the response surface model. The particle dimension is set to 9, which allows the selection of n optimal indicators from 9 indicators as a judgment basis. This approach not only facilitates the analysis of the relationships between the influencing factors and $PM_{2.5}$ levels but also reduces the dimensionality of the matrix data to accelerate algorithm execution.

To explore the degree of influence of each feature factor on the $PM_{2.5}$ concentration, the initial binary population of the BPSO algorithm can be considered. On the basis of the Monte Carlo method, the occurrence frequency of each feature is recorded, and combined with the calculated objective function values, their degrees of influence are assessed. The specific process is as follows:

Step 1: A large population size is selected for the particle swarm, and the

positions of the binary particles in the swarm are initialized.

Step 2: The objective function value f for each individual binary particle is calculated.

Step 3: The occurrence frequency p_i of the i-th feature for all the particles, as well as the corresponding objective function f_i for these particles, is recorded.

Step 4: The degree of influence for the i-th feature factor is computed as $\frac{\sum_1 f_i}{p_i}$, with low values indicating a high level of importance for the influencing factors.

Construction of the DBN and Kriging Algorithm

The kriging interpolation algorithm significantly depends on the choice of variogram, which affects the interpolation accuracy and stability. To increase the adaptability and accuracy of the algorithm when sample data are processed with notable feature differences, we explored strategies to intelligently select a variogram suitable for different sample features.

In this study, we considered that different seasons present distinct climatic and environmental characteristics. Therefore, regarding the characteristics of each season, the selection of an appropriate kriging variogram is the key to improving the interpolation accuracy. We integrated DBN into the kriging interpolation algorithm to accurately select the most suitable variogram for each sample. The task of variogram selection was formulated as a supervised classification problem. Each sample in the training set was labeled according to the variogram model (e.g., spherical, exponential, Gaussian) that achieved the best interpolation accuracy for that sample. DBN was trained to learn the nonlinear mapping from environmental variables to variogram labels. During prediction, the trained DBN assigns the most suitable variogram class for new input data, which is then used in kriging interpolation.

The specific operational procedure and method are as follows:

1) Dataset initialization: Assume a dataset containing N samples, where each sample may represent the data of a month or a day. At the initial stage, data preprocessing

is performed to reduce the data range of the different physical variables (such as temperature, humidity, and wind speed) and to remove potential noise or outliers.

2) Dataset organization: Different variograms are selected and applied to the kriging interpolation algorithm. Among the interpolation results, the samples with the optimal effect comprise the dataset for the DBN, and the output of the dataset is the number of variograms, such as 1, 2, and 3. Through layer-by-layer training, DBN can learn multilevel representations of environmental data. In this study, the DBN architecture consisted of three hidden layers with 100, 50, and 25 neurons, respectively. Each hidden layer was initialized using a Restricted Boltzmann Machine (RBM) and pre-trained using contrastive divergence. Sigmoid activation functions were applied throughout the layers. The final fine-tuning stage was performed using backpropagation with a learning rate of 0.01 and a maximum of 200 epochs. To mitigate overfitting, dropout with a rate of 0.2 was applied between layers, and early stopping was employed by monitoring validation loss with a patience threshold of 10 epochs.

3) Optimization and fine-tuning: At the optimization stage, the DBN facilitates the use of techniques such as backpropagation (BP) and gradient descent to adjust the weights and biases to optimize the internal parameters, capture the physical relationship between the environmental variables and the $PM_{2.5}$ concentration, and reduce prediction errors.

4) Integration of the DBN: The trained DBN is integrated into the kriging interpolation model to classify each set of new input data and select the most suitable variogram. To predict the $PM_{2.5}$ concentration, new environmental data can be input into DBN. Then, an appropriate variogram can be selected and used for kriging interpolation.

5) Calculation and comparison: On the basis of the selected function and sample data, the interpolation results of the corresponding kriging model can be obtained and compared.

The key advantage of this model is its ability to assign an appropriate variogram to each sample, which can represent data from a specific month or day. Therefore, the corresponding function can be accurately selected for interpolation on the basis of daily data, which can improve the accuracy and robustness of the interpolation algorithm.

The pseudocode is as follows:

Collect and clean the data
Perform principal component analysis feature selection
Initialize and train the DBN {Input: Cleaned feature selection data
Output: Deep feature representation and optimal variogram}
Apply the kriging interpolation algorithm {Input: Optimal variogram and spatial data
Output: Spatial interpolation result}
Predict the $PM_{2.5}$ concentration via the combination of the DBN and kriging interpolation algorithm
Evaluate the prediction performance {If the performance is satisfactory:
Apply the model for real-time prediction and analysis
Otherwise:
Adjust the parameters and conduct retraining}

Results and Discussion

Model Implementation and Case Study

To evaluate the performance of the model, we obtained real-time atmospheric environment monitoring data from the National Urban Air Quality Real-Time Release Platform of the Ministry of Environmental Protection and conducted a detailed simulation analysis using daily data for Xi'an City from January 2020 to January 2021. We not only analyzed the data in detail but also compared the prediction results with those of the ordinary, spatiotemporal, and improved kriging algorithms. The purpose of this comparative analysis was to confirm the significant advantages of the proposed model in terms of accuracy and reliability.

During data processing, we carefully organized and categorized the daily $PM_{2.5}$ data within the selected time range and performed a detailed statistical analysis, including measures of central tendency (such as the mean and median), variability (such as the standard deviation), and seasonal trends. Through data observation and analysis, we determined that the $PM_{2.5}$ concentration is the lowest in summer and highest in winter, with the values in spring and autumn occurring between those in the other seasons. To visualize this notable seasonal variation trend more clearly, we constructed a box plot, as shown in Fig. 4.

Fig. 4. Description: Each box represents the distribution of daily $PM_{2.5}$ concentrations for a given month, with the sample size for each month indicated. The plot displays the median, interquartile range, and potential outliers. The data analysis revealed that the $PM_{2.5}$ concentration was lowest in summer and highest in winter, with values in spring and autumn falling between those in the other seasons.

In this study, we rigorously preprocessed the entire set of sample data (365 samples in total) to ensure the data quality and analysis accuracy. We employed a dimensional normalization technique to standardize the original data to eliminate unit and scale differences among the various data types, as well as to ensure consistency and comparability of the data in the subsequent analysis.

Based on the results of data preprocessing, the Binary Particle Swarm Optimization (BPSO) algorithm was further employed for dimensionality reduction. During the optimization process, the selection of parameters significantly affects both the accuracy of convergence and the computational efficiency. Among these, the population size and the number of iterations are the most critical parameters. Therefore, a sensitivity

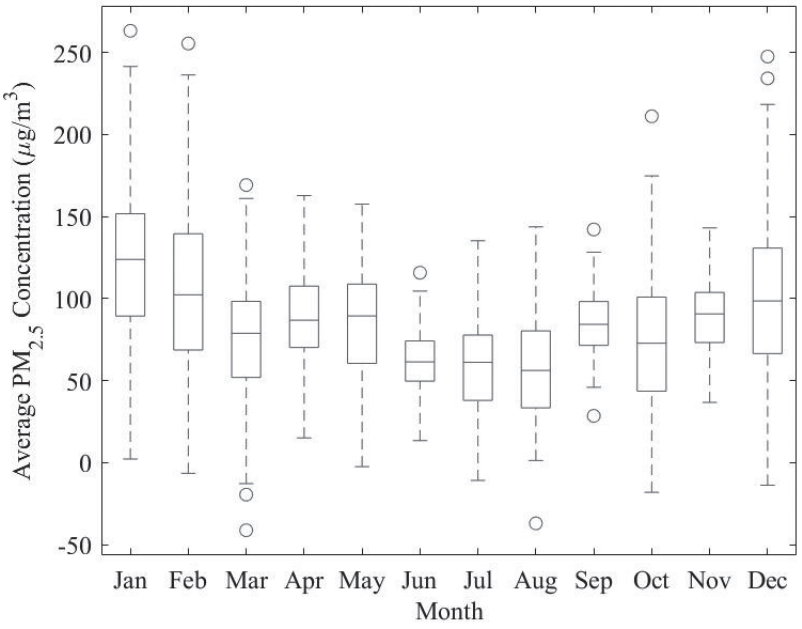

Fig. 4. Seasonal variation in the $PM_{2.5}$ concentration box plot (January 2020-January 2021).

analysis was conducted by setting the number of iterations from 50 to 500 with an increment of 50, and the population size from 20 to 200 with an increment of 20. The optimization results are illustrated in Figs 5 and 6.

The analysis indicated that when the number of iterations reached 400, the mean squared error (MSE) attained its minimum value, and further increases in the iteration count no longer resulted in performance improvements, while the growth in computational time slowed down. Similarly, when the population size was set to 170, the objective function also achieved its lowest value, and the computational cost tended to stabilize. Therefore, 400 iterations and a population size of 170 were selected as the optimal parameter configuration in this case.

On the basis of the data preprocessing results, we employed the BPSO method to reduce the dimensionality of the data. We set the population size of the particle swarm to 50 and the number of iterations to 100 for optimization. The iteration curve and the degree of influence of the variables are shown in Figs 7 and 6, respectively. The optimization results obtained with the particle swarm indicated that selecting rainfall and wind speed as features minimized the objective function, which coincides with the results shown in Fig. 8.

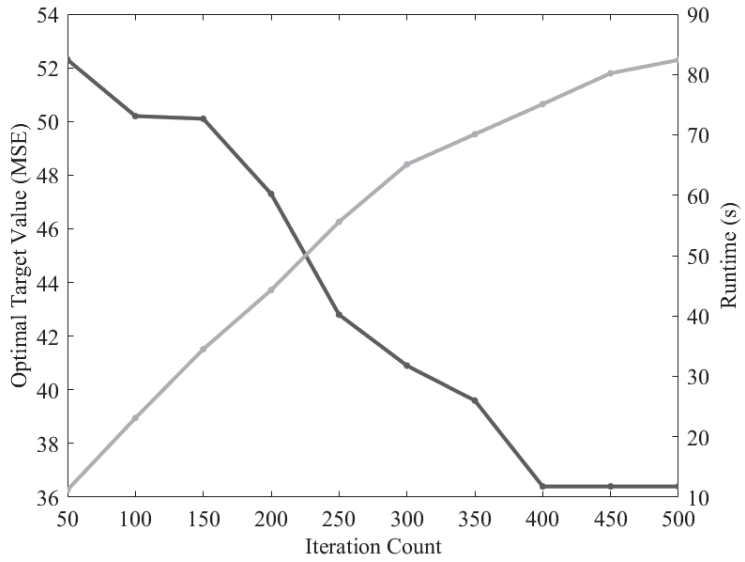

Fig. 5. Impact of Iteration Count on Objective Function (MSE) and Running Time.

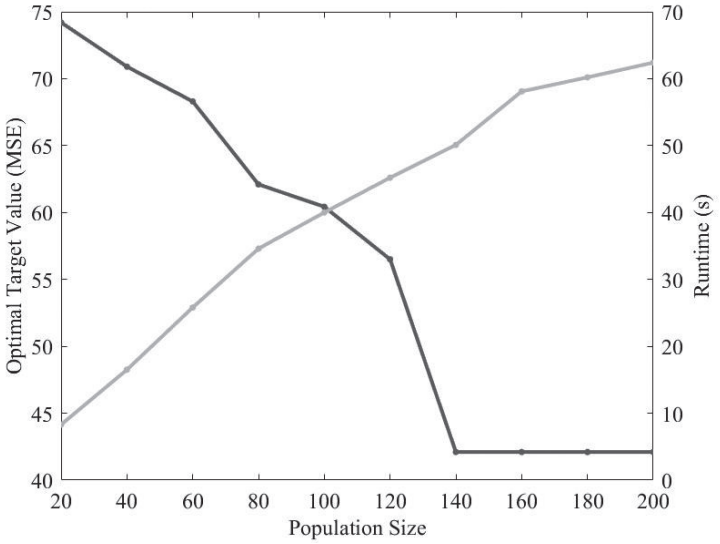

Fig. 6. Impact of Population Size on Objective Function (MSE) and Running Time.

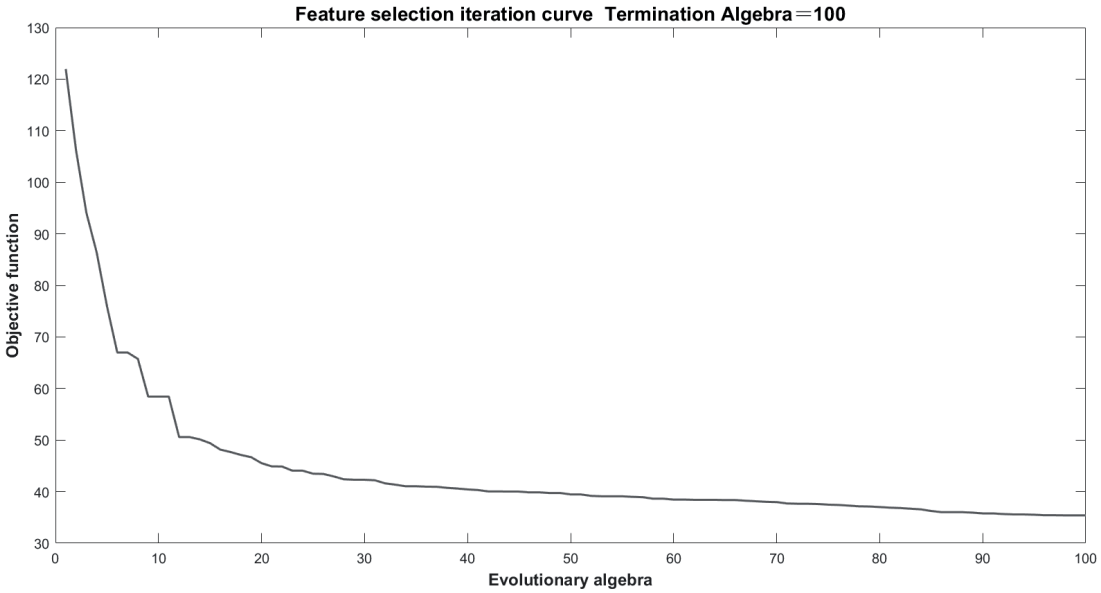

Fig. 7. Iteration curve.

Fig. 7. Description: This plot shows the iteration curve for the binary particle swarm optimization (BPSO) method during the optimization process. The x-axis represents the number of iterations (ranging from 1 to 100), whereas the y-axis represents the value of the objective function at each iteration. The curve depicts how the objective function value changes with each iteration, indicating the convergence behavior of the optimization process. The curve shows a decreasing trend, which suggests that the optimization successfully minimizes the objective function as the iterations progress.

Fig. 8. Description: This plot shows the degree of influence of the different input variables (such as rainfall, wind speed, and temperature) on the optimization process via the BPSO method. The x-axis represents the different variables, whereas the y-axis indicates the degree of influence or weight assigned to each variable in minimizing the objective function. This figure clearly shows that rainfall and wind speed have the greatest influence, which aligns with the optimization result that selecting these two variables minimizes the objective function.

After data preprocessing, we used three classical functions, namely, spherical, exponential, and Gaussian functions, for in-depth training and interpolation of the kriging model. Our main goal was to explore more accurate and practical spatial data interpolation and prediction strategies with the help of these carefully selected functions.

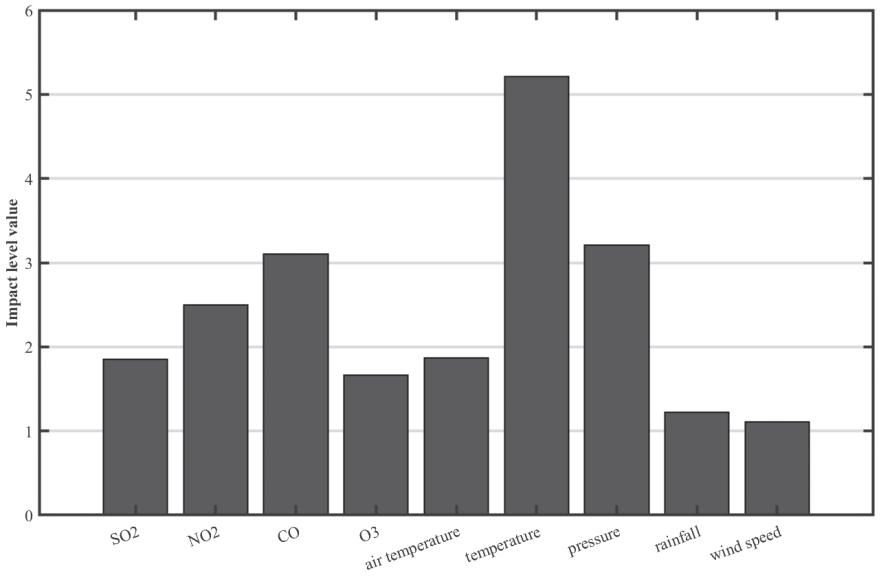

Fig. 8. Degree of influence.

To evaluate the performance and accuracy of the prediction model, we considered the mean square error (MSE) as the core evaluation metric. The MSE effectively represents the average of the squares of the differences between the predicted values and the actual values, and it is widely used and recognized as an important quantitative evaluation metric. In the conventional evaluation process, we compare the MSE values generated by different algorithms (such as linear regression, decision trees, and neural networks) for a unified dataset. A smaller MSE generally indicates a greater level of consistency between the predicted results and the actual outcomes.

However, to determine the most suitable algorithms and model, we used a multidimensional and multilevel evaluation and decision-making mechanism. In addition to performing an in-depth analysis of model prediction accuracy, we comprehensively considered key factors such as model complexity, training time, and interpretability. Ensuring that the selected model is accurate, practical, and of extensive application value is necessary.

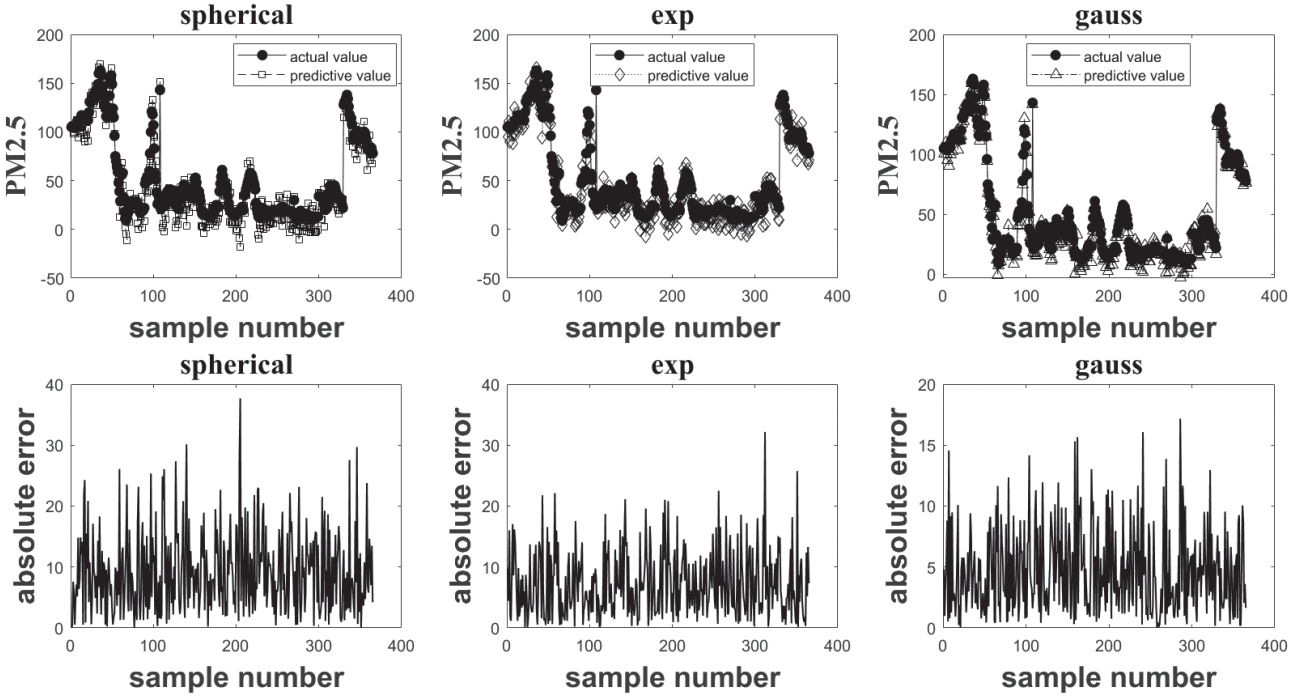

Fig. 9. Interpolation results using different variograms.

A detailed comparison of the interpolation results and the actual values is shown in Fig. 9.

The experimental results showed that the kriging model constructed using the three distinct variograms (spherical, exponential, and Gaussian functions) still has room for improvement in terms of its interpolation accuracy. Specifically, the interpolation errors generally ranged from 0 to 30. Among these three methods, the model constructed with the Gaussian variogram exhibited relatively small interpolation errors, but its accuracy should be improved further.

For in-depth data analysis, we set the error threshold to 10. According to this criterion, we selected samples with an error less than the threshold from the three sets of results and obtained 126, 56, and 234 samples. The total number of samples satisfying the error threshold significantly exceeds the total number of samples (365), a phenomenon that should be investigated further.

To investigate this phenomenon and its underlying factors in depth, we constructed a relationship curve between different error thresholds and the number of samples, as shown in Fig. 10. During a careful analysis of this curve, we observed that when the error threshold was 0.13 times the maximum error value, the total number of samples that satisfied this condition slightly exceeded the original total number of samples. This finding provides an important reference for further optimizing the model and selecting more suitable interpolation algorithms.

Fig. 10. Description: This plot shows the change in the total number of samples meeting the error threshold condition as the threshold increases. The x-axis represents the error threshold, and the y-axis represents the number of samples with absolute errors below the threshold for three kriging models (spherical, exponential, and Gaussian). A regression line is added to highlight the general trend, showing that as the threshold increases, the number of qualifying samples also increases. When the threshold is 0.13 times the maximum error, the number of qualifying samples slightly exceeds the total number of samples.

After organizing the samples that satisfied the specified criterion, we compiled the data into a dataset and categorized them into three class labels according to their corresponding variograms (spherical, exponential, and Gaussian functions). The data with these labels were then input into our designed DBN for in-depth analysis and processing. During the configuration phase of the neural network, on the basis of previous experience and the results of a preliminary experiment, we set the number of iterations to 2000 and set the error accuracy to 0.001. In addition, we divided the training and test set samples via a spatiotemporal slicing method, where the training set consisted of data from the previous months, covering the $PM_{2.5}$ concentration trends across different regions, while the test set was composed of data from the subsequent months for validation. This spatiotemporal division ensures that the model can consider both spatial dynamics and temporal dynamics to increase its generalizability and predictive power. During neural network training, we used tansig and sigmoid transfer functions. The application of these classical functions not only ensures the stability of model training but also greatly improves the training efficiency and model performance. The performance of the DBN was compared with that of the classical support vector machine (SVM) and BP neural network, as shown in Fig. 11. We observed that the accuracy of the DBN significantly exceeded that of the SVM and BP neural network models, which

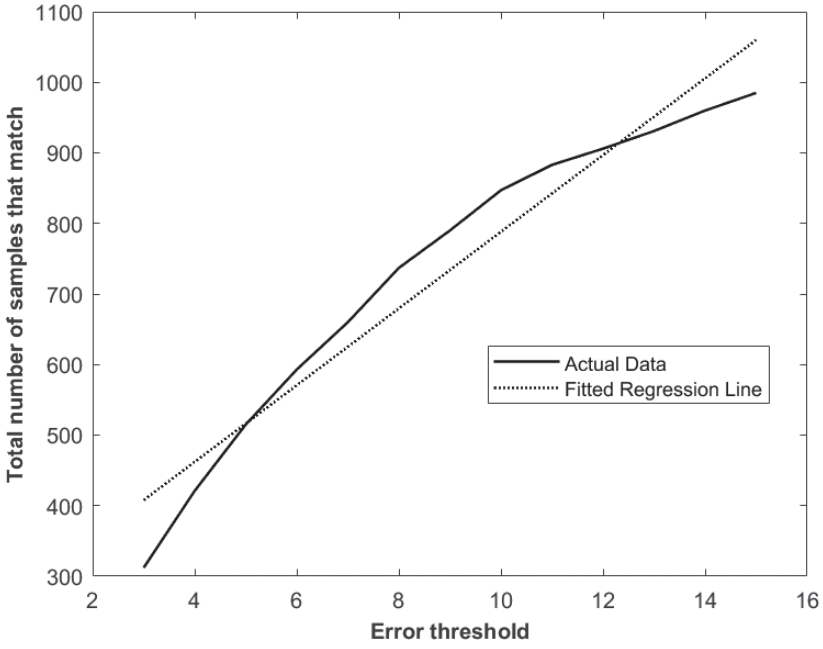

Fig. 10. Variation in the total number of samples with respect to the error threshold and regression.

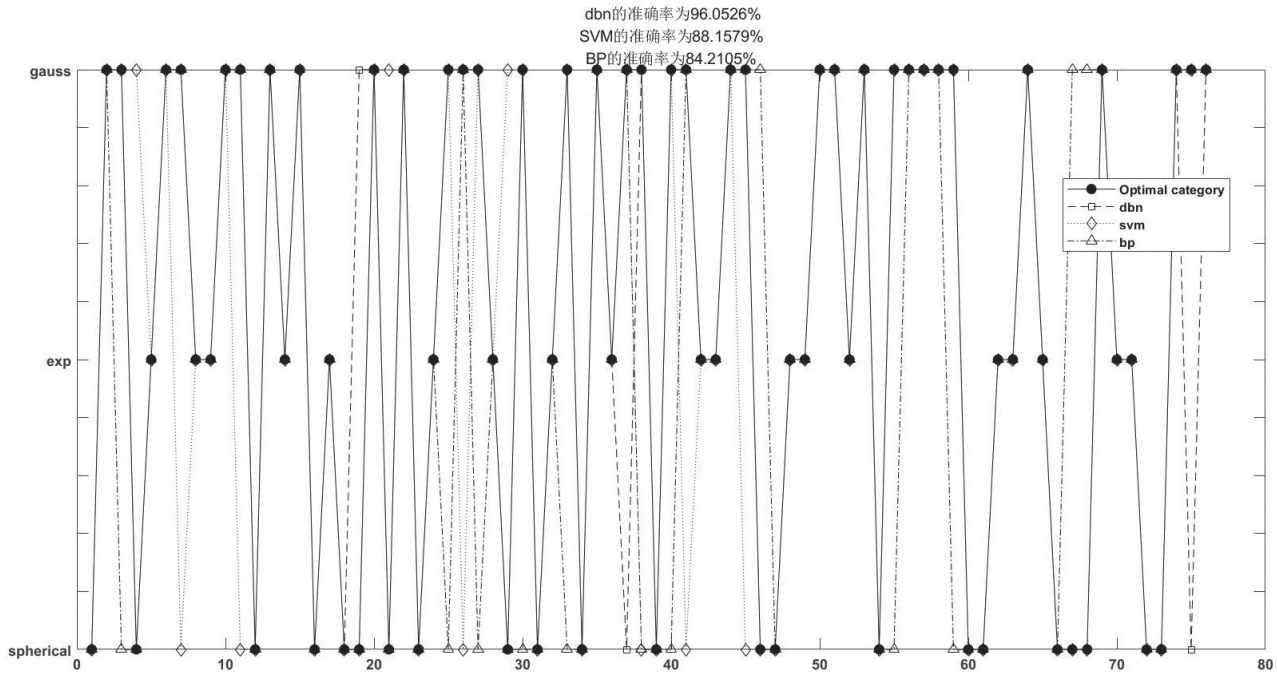

Fig. 11. Comparison of the accuracy of the DBN with that of different transfer functions.

not only validates the excellent performance of the DBN in processing complex datasets but also reveals its superior generalizability. This result also suggests that the DBN can be used to more accurately and reliably select the most suitable variogram for the kriging model, thereby greatly improving the accuracy and reliability of air quality prediction.

Next, we integrate the optimized DBN into the kriging interpolation method to improve the accuracy and computational efficiency of data interpolation. To comprehensively evaluate the effectiveness and potential advantages of this integrated strategy, we conducted a series of comparative analyses, with a focus on exploring the differences in prediction performance between the DBN kriging model and the other two advanced improved kriging algorithms, as well as the traditional kriging algorithm. The specific results are shown in Figs 12 and 13.

In this process, we not only quantitatively evaluated the errors and accuracy but also thoroughly investigated the stability and adaptability of each method in resolving different datasets and environmental conditions. Our goal was to reveal the performance of these algorithms in actual application scenarios to more accurately identify and quantify the advantages and potential limitations of the DBN kriging model.

Fig. 12. shows the error distributions of the four algorithms during different seasons in the form of pie charts. Notably, with the exception of the DBN kriging model, the other three algorithms exhibited high error variances and uneven error distributions. This result highlights the distinct limitations of these three algorithms in adapting to $PM_{2.5}$ fluctuation patterns during different seasons. In contrast, the DBN kriging model demonstrated a more balanced and stable error distribution across the four seasons, indicating that this model provides excellent adaptability and can be used to automatically select the most suitable variogram on the basis of specific environmental and atmospheric conditions for each month, thereby ensuring accuracy and consistency in year-round predictions.

Compared with the self-adaptive butterfly optimization algorithm (SABOA), our model not only inherits its excellent search accuracy and iterative capability but also imparts broader applicability and flexibility to the model by introducing an adaptive response null-space planning and control strategy. Exploring the effects of this model in actual application scenarios and future research directions is important. Compared with previous studies, the DBN kriging model effectively improves the performance and accuracy of the system through adaptive techniques. This finding has been confirmed by the design scheme of the PAM4 transceiver, which reveals the enormous potential of adaptive technology in improving system performance. Through these integrations and innovations, our model not only overcomes the limitations of previous algorithms and achieves broader, more accurate, and more stable $PM_{2.5}$ predictions but also creates new research avenues in the fields of environmental science and atmospheric research. The novelty of the proposed model is reflected in its adaptability and generalizability, enabling it to maintain notable prediction accuracy and stability during different seasons and in variable environments.

Fig. 13. confirms the excellent performance of our model. In the evaluation based on various statistical

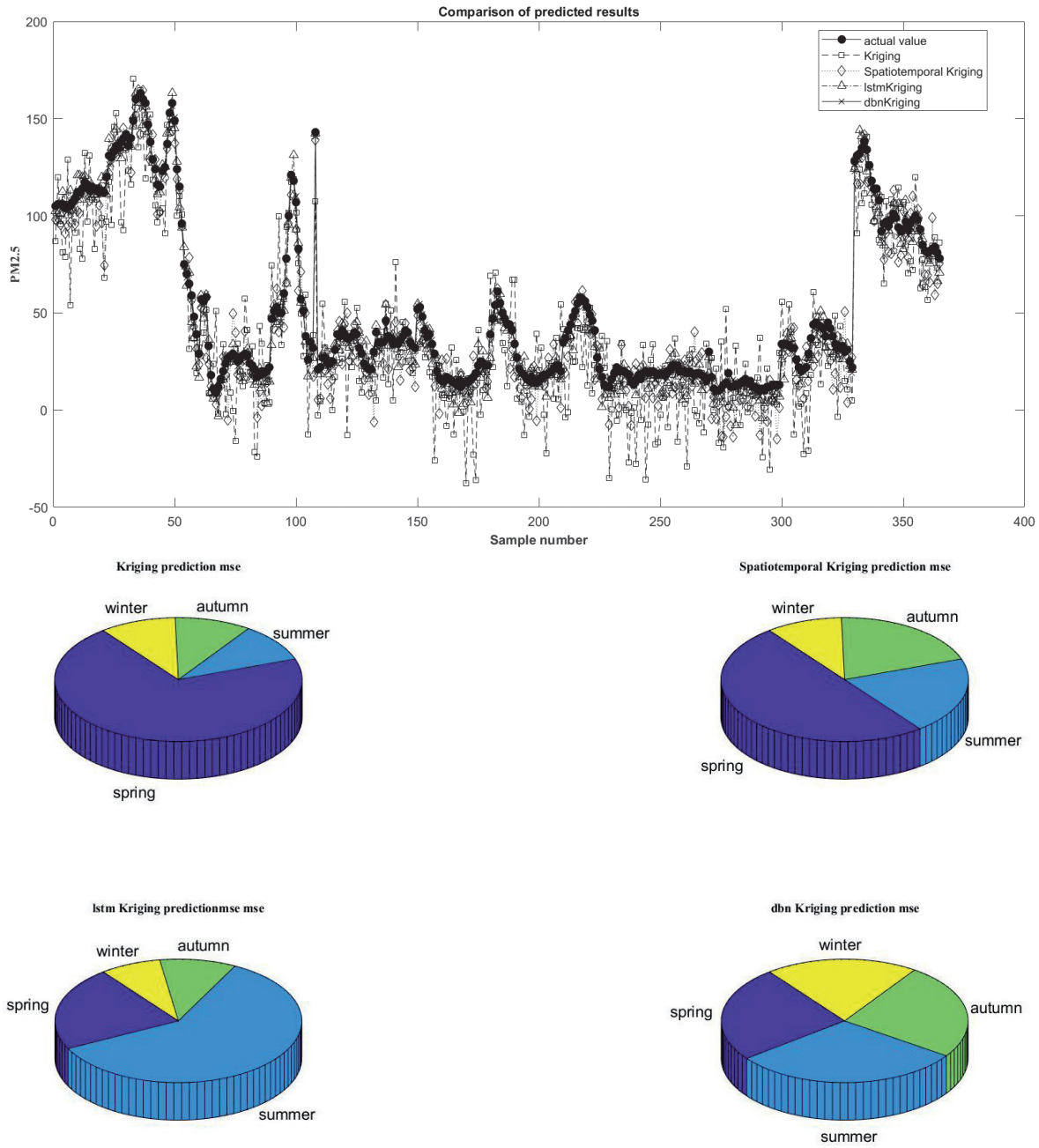

Fig. 12. Comparison of the interpolation results obtained via the different algorithms.

metrics, the traditional kriging algorithm performs relatively poorly, with its large mean square error significantly limiting its practicality. In contrast, both the spatiotemporal kriging algorithm and the improved kriging algorithm yielded significant accuracy improvements, with similar performance levels. In this study, we employed a kriging algorithm optimized by a DBN. This model specifically accounts for the differences and complexity of daily real-time data and optimally selects the appropriate variogram in an adaptive manner, thus achieving more accurate spatial interpolation. This optimization strategy is highly effective and eventually reduces the mean absolute percentage error (MAPE) to 4.677%, indicating that our model achieves superior interpolation performance compared to that of all the comparison algorithms.

Synergistic Effects of DBN and Kriging Interpolation

The integration of deep belief networks (DBNs) with kriging interpolation demonstrated significant improvements in predicting $PM_{2.5}$ concentrations. Unlike traditional linear models, the DBN-kriging framework effectively captures the complex nonlinear relationships between air quality and meteorological variables. This synergy enhances both the spatial interpolation accuracy and the temporal prediction

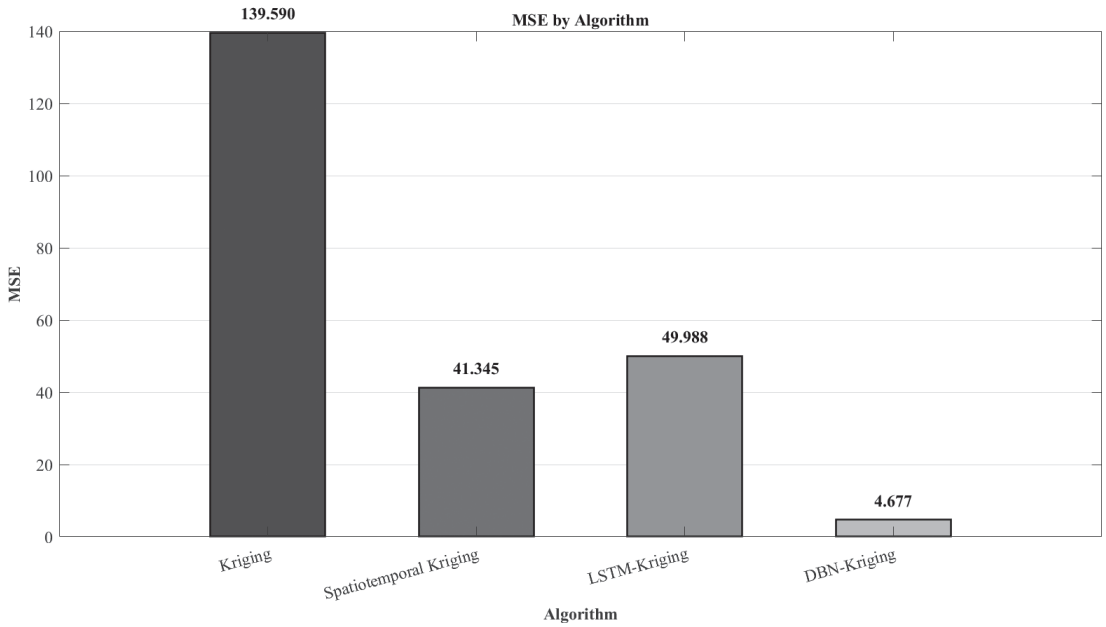

Fig. 13. Histogram of MAPE values for different algorithms.

stability, addressing key limitations in conventional environmental modeling approaches. To evaluate real-world deployment feasibility, the computational efficiency of the proposed model was also considered. Under a standard computing environment (Intel Core i7, 16 GB RAM), the average training time for the DBN was approximately 20 minutes for one year of daily $PM_{2.5}$ data, while the inference (prediction) time for new samples was under 0.5 seconds. These results indicate that the model can support near-real-time prediction with proper optimization and integration, making it suitable for online environmental monitoring systems or early warning applications.

Feature Optimization and Dimensionality Reduction

To mitigate the impact of noise and redundant variables, principal component analysis (PCA) was employed. The PCA process enabled the model to extract the most relevant features for $PM_{2.5}$ prediction, enhancing computational efficiency while preserving essential information. Furthermore, the binary particle swarm optimization (BPSO) algorithm was utilized to rank the influence of different environmental factors, resulting in a more interpretable and adaptable response surface model tailored to local conditions.

Cross-District Applicability within Xi'an

To clarify the scope, our evaluation focuses on cross-district conditions within Xi'an. Scenario tests using synthetic or reweighted environmental profiles (to mimic differing temperature, humidity, and rainfall regimes) indicate that the framework remains functional under varied settings. However, these scenario tests are not observational validations in other cities and do not reflect the full complexity of coastal or heavy-industrial emission sources. Inter-city generalizability therefore remains unverified and will require external observational datasets or transfer-learning experiments in future work.

Comparison with Related Studies

A comparative review of recent literature further highlights the strengths of the proposed method. Bhattacharya et al. [26], Hörmann et al. [27], and Yadav et al. [28] emphasized the seasonal variability in $PM_{2.5}$ concentrations and the effects of climate and lockdown measures. However, their methods did not incorporate adaptive interpolation mechanisms or real-time variogram selection.

In contrast, our approach addresses these gaps by integrating deep learning and geostatistical modeling. The ability of the DBN-kriging framework to dynamically adapt to regional characteristics, combined with the noise-reduction capabilities of PCA, provides a comprehensive solution for accurate, high-resolution air quality prediction.

Conclusions

This study developed a deep belief network (DBN) integrated with kriging interpolation for accurate $PM_{2.5}$ concentration prediction. By combining DBN-assisted variogram selection with PCA-based feature optimization, the framework captures nonlinear relationships between $PM_{2.5}$ and meteorological/air-quality covariates and consistently improves accuracy over conventional interpolation baselines.

Using multi-station data across districts within Xi'an, the model demonstrated stable performance under heterogeneous local conditions, indicating intra-city (cross-district) applicability. We also conducted scenario analyses constructed from environmental profiles with differing temperature, humidity, and rainfall regimes; these tests suggest flexibility to climatic variation but do not constitute inter-city observational validation or fully represent coastal or heavy-industrial emission mixtures.

Therefore, inter-city generalizability remains to be validated. As future work, we will evaluate transferability with external observational datasets and transfer-learning protocols in representative coastal and industrial cities.

Overall, the proposed DBN–kriging framework outperforms traditional approaches by better capturing complex nonlinear interactions, providing a methodological basis for air-quality modeling, climate-aware pollution-control strategies, and intelligent environmental monitoring.

Acknowledgments

The authors would like to thank the anonymous reviewers and the editorial team of thePolish Journal of Environmental Studiesfor their constructive comments and suggestions, which have helped improve the quality of this work. The air quality and meteorological data were provided by Qingyue Data. This study was supported by the National Natural Science Foundation of China (No. 71874134), the Science and Technology Project of Shaanxi Province (No. 2019JZ-30), and NYG2024287.

Conflict of Interest

The authors declare no conflict of interest.

References

1. LI J., HEAP A.D. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecological Informatics. **6** (3), 228, **2011**.
2. CHEN T., GUESTRIN C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'16), pp. 785, **2016**.
3. SAMAL K.K.R., BABU K.S., DAS S.K. Spatial-temporal prediction of air quality by deep learning and kriging interpolation approach. ICST Transactions on Scalable Information Systems. **10** (7), 1, **2023**.
4. NAG P., SUN Y., REICH B.J. Spatio-temporal DeepKriging for interpolation and probabilistic forecasting. Spatial Statistics. **57** (4), 100773, **2023**.
5. WU D., WANG H., TANG C., LI L. Estimation of hourly $PM_{2.5}$ concentrations in Sichuan Province based on GTWR-XGBoost combined model. Research of Environmental Sciences. **36** (4), 649, **2023** [In Chinese].
6. XU M., ZHANG C., XU L., ZHANG R., ZHI J. Analysis of atmospheric pollution characteristics and influencing factors in Bijie city area. Safety and Environmental Engineering. (5), 64, **2018**
7. KONG F., WU M., QU H. Variation regularity and influencing factors of $PM_{2.5}$ concentration during heating season in Shenyang City. Journal of Ecology. **43** (8), 2313, **2023** [In Chinese].
8. ZHANG J., DU S., ZHOU X., WANG Q., ZUO D., YANG Y., QIN Y., QIN J., YANG Y., SHI S., TAN J., PENG S. Seasonal distribution of chemical components of $PM_{2.5}$ in different functional areas of Huainan City and source apportionment. Safety and Environmental Engineering. **30** (1), 212, **2023** [In Chinese].
9. LU Y.M., ZENG S.P., ZENG J., WANG S., SONG Y.Z. High-resolution spatiotemporal simulation of $PM_{2.5}$ concentration based on random forest: A case study in the core area of the Central Plains Urban Agglomeration. China Environmental Science. **43** (8), 3300, **2023** [In Chinese].
10. XU S., ZOU B., WANG M. Comparison of neural network and kriging methods for spatial estimation of $PM_{2.5}$ concentration. Geomatics and Information Science of Wuhan University. (10), 1642, **2020.**
11. TANG Y., YANG F., ZHAN Y. High-resolution spatio-temporal distribution and correlation analysis of $PM_{2.5}$ and PM_{10} in the Sichuan basin. China Environmental Science. (12), 4950, **2019** [In Chinese].
12. DU Y., HUANG Q. Spatio-temporal distribution characteristics of $PM_{2.5}$ in Henan Province and its relationship with vegetation coverage. Sciences Journal of Ecology and Environment. **28** (11), 2257, **2019**.
13. ZHAO L., LI Z., QU L. A novel machine learning-based artificial intelligence method for predicting the air pollution index $PM_{2.5}$. Journal of Cleaner Production. **468**, 143042, **2024**.
14. RAKHOLIA R., LE Q., VU K., HO B.Q., CARBAJO R.S. Accurate $PM_{2.5}$ urban air pollution forecasting using multivariate ensemble learning accounting for evolving target distributions. Chemosphere. **364**, 143097, **2024**.
15. YUE H., HE C., HUANG Q., ZHANG D., SHI P., MOALLEMI E.A., XU F., YANG Y., QI X., MA Q., BRYAN B.A. Substantially reducing global $PM_{2.5}$-related deaths under SDG3.9 requires better air pollution control and healthcare. Nature Communications. **15**, 2729, **2024**.
16. DING R., HUANG L., YAN K., SUN Z., DUAN J. New insight into air pollution-related cardiovascular disease: An adverse outcome pathway framework of $PM_{2.5}$-associated vascular calcification. Cardiovascular Research. **120** (7), 699, **2024**.
17. PROSDOCIMI I., MASIOL M., TATTARA G. Air pollution in Venice and in its mainland: A first assessment of air quality control policies. Environmental and Ecological Statistics. **31** (2), 273, **2024**.
18. CHIESA G., VIGLIOTTI M. Comparing mechanical ventilation control strategies for indoor air quality: Monitoring and simulation results of a school building in northern Italy. Energy and Buildings. **322**, 114665, **2024**.
19. ZHOU S., WANG W., ZHU L., QIAO Q., KANG Y. Deep-learning architecture for $PM_{2.5}$ concentration prediction: A review. Environmental Science and Ecotechnology. **21**, 100400, **2024**.

20. YU Q., YUAN H.W., LIU Z.L., XU G.M. Spatial weighting EMD-LSTM based approach for short-term $PM_{2.5}$ prediction research. Atmospheric Pollution Research. **15** (10), 102256, **2024**.
21. ABIRAMI S., CHITRA P. Regional air quality forecasting using spatiotemporal deep learning. Journal of Cleaner Production. **283**, 125341, **2021**.
22. SIMEONOV V., TSAKOVSKI S., LAVRIC T., SIMEONOVA P., PUXBAUM H. Multivariate statistical assessment of air quality: a case study. Microchimica Acta. **148** (3), 293, **2004**.
23. SRISANG W., JAROENSUTASINEE K., JAROENSUTASINEE M., KHONGTHONG C., PIAMONTE J.R.P., SPARROW E.B. $PM_{2.5}$ IoT sensor calibration and implementation issues including machine learning. Emerging Science Journal. **8** (6), 2267, **2024**.
24. ZHANG J., FENG W., SHAN Y. Characteristics of $PM_{2.5}$ and PM_{10} pollution in Xinyang city and their relationship with meteorological factors. Green Technology. (22), 5, **2018** [In Chinese].
25. ZHANG J., LIU Y., CUI L.L., LIU S.Q., YIN X.X., LI H.C. Ambient air pollution, smog episodes and mortality in Jinan, China. Scientific Reports. **7** (1), 11209, **2017**.
26. BHATTACHARYA T., NARAYAN T., CHAKRABORTY S., KONAR S., SINGH S. Statistics as a technology to predict the seasonal variation of air pollution. International Journal of Innovative Technology and Exploring Engineering. **9** (3), 1426, **2020**.
27. HÖRMANN S., JAMMOUL F., KUENZER T., STADLOBER E. Separating the impact of gradual lockdown measures on air pollutants from seasonal variability. Atmospheric Pollution Research. **12** (2), 84, **2021**.
28. YADAV A.K., SHIRIN S., EMMANOUIL C., JAMAL A. Effect of seasonal and meteorological variability of air pollution in Singrauli coalfield. Aerosol Science and Engineering. **6** (1), 61, **2022**.