https://doi.org/10.1051/e3sconf/202339303008

Spatiotemporal dynamic interpolation simulation and prediction method of fine particulate matter based on multi-source pollution model

Yinchi Ma*

Beijing Academy of Agriculture and Forestry Sciences, Beijing, China

Abstract—Taking the air pollution monitoring data of 34 air monitoring stations in Beijing from February 8 to February 9, 2020, as an example. A spatiotemporal dynamic interpolation model of PM2.5 based on a multi-source pollution model was established. Based on the hourly spatiotemporal data of the day, the dispersion and attenuation of non-point source pollution in Beijing were interpolated. An improved hybrid genetic algorithm was used to solve the parameters of the air pollution model. The spatiotemporal Kriging model was used to predict the PM2.5 concentration diffusion on an hourly scale. The data of this area were analyzed quantitatively and qualitatively. The prediction data based on the spatiotemporal data before the current time was verified by the actual monitoring data. The results show that the model and method constructed in this paper could simulate and predict PM2.5 concentration on an hourly scale well, which could provide a good reference for the analysis, simulation, and prediction of air pollution.

1. INTRODUCTION

It is generally believed that the higher amounts of sulfate, nitrate, ammonium and organic matter in PM2.5 are due to the heavy traffic or vehicle emission and the burning of solid fuels in most parts of China [1]. From the end of January to the middle of February in 2020, it was in the period of the Chinese Spring Festival. It was also in the period of epidemic prevention and control in Beijing. There were few motor vehicles on the road, and most factories and construction sites had stopped production and work. Moreover, there was little dust. Besides, the policy of coal to gas and the ban on fireworks has been implemented in recent many years. The overall level of social activities in Beijing had also been reduced. The emission of pollutants had also been reduced. However, during this period, there were still two heavy pollution weather processes. According to the data, during these heavy pollution weather processes, Beijing, Tianjin, and the Hebei region just encountered unfavorable weather conditions, including high temperature, high humidity in winter, almost no wind or very small wind.

In recent years, there are more and more researches on the harm of air pollution to human health. Air pollution has become a public health problem. High concentration particulate matter (PM2.5) is associated with lung cancer, cardiovascular diseases, respiratory diseases, and metabolic diseases [2-4]. The scientific prediction of PM2.5 concentration can help to reduce health risks and economic losses. Many scholars have done much research on the simulation and prediction of air pollution at home and abroad. The concentration of PM2.5 in the air is unstable with time, and most of it comes from artificial pollution [5]. Chen [6] and Xie [7] predicted PM2.5 using data of different scales. Liu [8] proposed a hybrid model composed of five algorithms: wavelet packet decomposition (WPD), gradient lifting regression tree (GBRT), linear programming lifting (LPBoost), multi-layer perceptron (MLP), and Dirichlet process hybrid model (DPMM) to simulate and predict the pollutant data collected in Tangshan at four different time intervals.

The existing air quality prediction methods include the deterministic method, statistical method, machine learning, and deep learning method. Many methods such as machine learning, neural network, and random forest are used in PM2.5 prediction. Junfei [9] proposed a PM2.5 prediction method based on image contrast-sensitive features and weighted banded neural network (WBBNN). Danesh [10] combined the satellite aerosol optical depth (AOD), land use, and meteorological data to create a daily PM2.5 prediction model by using the integrated machine learning method for Greater London from January 1, 2005, to December 31, 2013. The prediction method was carried out on 3960 grid cells in the ratio of 1 km × 1 km. In the prediction of PM2.5, the data used are more and more diverse, such as satellite, meteorological data, ground-based PM2.5, and geographic data [11] [12] [13].

Most of the above prediction methods only analyze the temporal dimension of PM2.5 in the long time series or the spatial dimension of PM2.5 distribution in different sampling points. The time scale is generally in the annual change, monthly change, or daily change, and the ability to predict the spatial variability is weak. At the same time,

*e-mail: mayinchi@bjfishery.com

the linear relationship is analyzed, but the law of atmospheric turbulence in the process of pollution diffusion and aggregation in a short time is ignored. Moreover, the prediction modeling analysis of PM2.5 concentration from aggregation, stability to diffusion process is not realized.

At present, real-time air pollution monitoring has become an important method to control pollution. However, due to the high cost of construction, the monitoring range of the monitoring station is limited, and it is unable to monitor the concentration of all air pollutants in every corner in real-time. Therefore, the overall understanding of the spatial and temporal distribution of air pollution is usually lacking. A comprehensive spatial and temporal analysis and air pollution control cannot be carried out. The satellite remote sensing and spatial interpolation/extrapolation techniques have been widely used to solve this problem in recent years. Spatiotemporal geo-statistics has become a research hotspot in the atmospheric environment and dynamic distribution of biological population because it can use spatiotemporal interpolation, random simulation estimation, and quantify spatiotemporal changes to reflect the process of geographical spatial changes dynamically.

In this paper, we used spatiotemporal data to simulate the distribution characteristics of PM2.5 in time and space. Then we analyzed the spatiotemporal distribution of PM2.5 and predicted the diffusion of PM2.5 with time by using the diffusion law of the atmospheric diffusion model and the correlation of data in adjacent space and time. The cross-validation was carried out by the measured data with the spatiotemporal interpolation results of the same day and the measured data with the forecast results based on the previous day's data in the hourly time scale of the second day. The reliability and superiority of the proposed method were proved.

2. MATERIAL AND METHOD

2.1 Hypothesis of the problem

Although due to meteorological factors, a stable state does not exist. However, the distribution of air pollutants does not change rapidly; there is always a process of change and diffusion. To facilitate research and analysis, we can assume that it is stable in a certain period. Then, according to the diffusion model, the distribution of air pollutants at the next set time point is dynamically predicted. This paper assumes that:

First, the influence of air temperature, air pressure, and humidity on PM2.5 distribution is not considered; only wind speed and direction are considered.

The second, Ignoring the vertical distribution of PM2.5, the three-dimensional problem is transformed into a two-dimensional problem.

The third, the pollutant obeys the diffusion law, and the change of diffusion process on the concentration axis is Gaussian distribution.

Fourth, when the wind speed and direction are added, the wind is stable in the set period, and the mutation factors and the influence of urban buildings on the wind direction are ignored.

The fifth, 34 air monitoring stations in Beijing are regarded as point spread pollution sources; the algorithm assumes eight pollution sources in eight directions around Beijing to simulate external pollution sources.

The sixth, under the same conditions, the attenuation coefficients of matter in different directions are the same.

2.2 Establishment of the model

Gaussian model can take wind direction and wind speed into account but cannot attenuate. The partial differential model cannot define wind direction and wind speed, but it can attenuate. For the assumption in this paper, these 34 monitoring points are not real pollution sources. Wind speed and direction have a great influence on the Gaussian model, and the Gaussian model cannot be attenuated. The Gauss large-space point-source diffusion model is compared with the partial differential diffusion model for Beijing under the stable weather during the epidemic period. The partial differential diffusion model with stronger adaptability and attenuation is selected.

At the same time, another assumption is made that in selecting monitoring points, it must be considered that the pollution here is serious and it is the source area. Therefore, PM2.5 of the whole city comes from these 34 pollution sources and external pollution sources. Diffusion and attenuation are considered based on these pollution sources, and spatiotemporal dynamic interpolation is performed for other spatiotemporal data.

Suppose a pollution source from which a matter begins to spread around with the change of time t. The diffusion coefficients along the x, y, and z directions are constants a, b, and c, respectively. Attenuation (e.g., absorption, metabolism, etc.) makes the mass decrease proportional to the concentration. Moreover, the concentration of this matter in the surrounding space is zero before diffusion.

The partial differential diffusion attenuation model shown in formula (1) is a linear parabolic equation with constant coefficients. It is the mathematical model of the attenuation diffusion process.

$$\frac{\partial u}{\partial t} = a^2\frac{\partial^2 u}{\partial x^2} + b^2\frac{\partial^2 u}{\partial y^2} + c^2\frac{\partial^2 u}{\partial z^2} - k^2\mathrm{u} \quad (1)$$

The distribution of PM2.5 within 0-200m near the ground is almost unchanged. Atmospheric monitoring stations are usually set near the ground. Therefore, the PM2.5 diffusion model can be flattened. That is to say, we only consider the distribution of PM2.5 in the horizontal plane and ignore the distribution of PM2.5 in the vertical direction. Therefore, for each pollution source with a total pollutant amount of G located at (x_0, y_0), the general analytical solution of diffusion is:

$$\mathrm{u(x,y,t)} = \frac{G}{8\pi ab\sqrt{\pi t}}\exp\{-\frac{(x-x_0)^2}{4a^2t} - \frac{(y-y_0)^2}{4b^2t} - k^2t\} \quad (2)$$

Therefore, it is necessary to calculate the parameters a, b, and k in the above formula (2).

Generally, the pollution sources of a city are divided into external pollution and local pollution. During the most serious period of the epidemic in February in Beijing, according to the prevention and control requirements, most people work at home and go out less. A large number of the service industry, industrial enterprises, and other industries shut down production. Only the infrastructure works. In theory, air pollution is low. It is convenient to analyze the sources and characteristics of air pollution in Beijing under objective conditions. Equation (2) is the prediction model of PM2.5 point source diffusion affected by a single source. In fact, in the case of multi-source emission, the concentration contribution to any point in the evaluation area results from the superposition of multiple pollution sources. Therefore, it is necessary to establish a multi-source PM2.5 diffusion prediction model under the joint influence of multi-source based on the PM2.5 point source diffusion model under the influence of a single source. The PM2.5 diffusion prediction model established in this way is closer to the actual situation.

In this paper, 34 air monitoring stations in Beijing are regarded as point source pollution sources. The sources of PM2.5 concentration measured by any monitoring station include the contribution of the monitoring station itself, the diffusion superposition contribution of other monitoring stations, and the contribution of external continuous pollution sources. The results of pollution superposition can be obtained by chemical mass conservation model as formula (3):

$$Z^* \times W = \Delta Z \qquad (3)$$

Z^* is the concentration matrix of each pollution source contributing to the monitoring station. W is the weight matrix of the contribution of each pollution source to the monitoring station. ΔZ is the sum matrix of the contribution of each pollution source to the monitoring station.

Therefore, the spatial and temporal concentrations of PM2.5 at various monitoring stations in Beijing are as follows:

$$Z^{\circ}(\mathrm{i,t}) = \mathrm{Z(i,t-1)} + \Delta \mathrm{Z} \qquad (4)$$

$Z^{\circ}(\mathrm{i,t})$ is the pollutant concentration value calculated according to formula (2) at time t of the monitoring station i. $\mathrm{Z(i,t-1)}$ is the actual monitoring value of pollutants at the previous time point of the station. ΔZ is the contribution of multiple pollution sources to the station during the diffusion process.

When there are N pollution sources in M periods, for the single monitoring station point (x_0, y_0), formula (3) can be expanded as follows:

$$\begin{bmatrix} Z^*_{11}\, Z^*_{12} \dots Z^*_{1N} \\ Z^*_{21}\, Z^*_{22} \dots Z^*_{2N} \\ \vdots \\ Z^*_{M1}\, Z^*_{M2} \dots Z^*_{MN} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_N \end{bmatrix} = \begin{bmatrix} \Delta Z_1 \\ \Delta Z_2 \\ \vdots \\ \Delta Z_M \end{bmatrix} \qquad (5)$$

Z^*_{MN} is the contribution concentration of pollution source N to the monitoring station when the time is M. WN is the weight of pollution source N. ΔZ_M is pollution concentration received from external of the monitoring station when the time is M. Therefore, for any time m, the received concentration of the monitoring station is the linear superposition of the contribution concentration of each pollution source. It is shown in formula (6):

$$\Delta Z_m = \sum_{n=1}^{N} Z^*_{mn}\, W_n \qquad (6)$$

Z^*_{mn} is the contribution concentration of each pollution source. It is calculated through formula (2) and shown in formula (7):

$$Z^*_{mn} = \frac{M_n}{8\pi ab\sqrt{\pi m}} \exp\{-\frac{(x-x_0)^2}{4a^2 m} - \frac{(y-y_0)^2}{4b^2 m} - k^2 m\} \qquad (7)$$

Through above formula, we can see that the unknown parameters in this method are a, b, k, and N pollution sources, namely N + 3 unknown parameters. For formula (5), W is solvable when M > N. W can be calculated by matrix inversion when M = N. When M < N, it can only be solved by an optimization algorithm, such as least square algorithm, genetic algorithm, simulated annealing algorithm and so on.

Generally, the relationship between the pollution source and the concentration received by the sensor has the strongest correlation with the distance and wind direction. Generally, the closer the distance is, the higher the contribution concentration is, and the greater the influence weight is. In this paper, it is assumed that the relationship is an inverse distance weighted linear relation.

The effect of multiple pollution sources on the same monitoring station (x_0, y_0) are mutually superimposed and influenced. To simplify the calculation process, this paper is based on the principle that the closer the distance, the greater the impact on the monitoring station. The formula for calculating the weight matrix W is set as formula (8):

$$W_n = \frac{\frac{1}{distance(0n)}}{\sum_{j=1}^{N} \frac{1}{distance(0j)}} \qquad (8)$$

$distance(0n)$ is the distance between the pollution source n and the monitoring station (x_0, y_0).

The effect of each pollution source on the monitoring site (x_0, y_0) can be obtained by weighting the distance between them.

In this paper, the improved hybrid genetic algorithm proposed in previous research solves N + 3 parameters. The designed fitness formula minF(S) is as follows:

$$\mathrm{minF(S)} = \sum_{t=1}^{M} \sum_{i=1}^{R} (\mathrm{Z(i,t-1)} + \Delta Z_m - Z(i,t)) \qquad (9)$$

R is the number of pollution monitoring stations. The number of available monitoring stations in Beijing is 34. M is the period. The data used in this paper is 24 hours a day. M is set as 24.

Interpolation is divided into two steps: first, quantitative analysis of the spatial structure of sample points, and then, prediction of unknown points. The general formula of spatiotemporal interpolation is as follows:

$$Z^*(s,t) = \sum_{i=1}^{N} \lambda_i Z(s_i, t_i) \qquad (10)$$

For any space-time point Z (s, t) that needs dynamic interpolation simulation, the weight λ_i is calculated by constructing the following weight matrix (formula 11) with the origin of Z (s, t) coordinate：

$$\mathrm{A\lambda = b} \qquad (11)$$

In formula 11, $\mathrm{A} = \begin{bmatrix} \gamma_{11(h_s h_t)} \gamma_{12(h_s h_t)} \cdots \gamma_{1n(h_s h_t)}\ 1 \\ \gamma_{12(h_s h_t)} \gamma_{22(h_s h_t)} \cdots \gamma_{2n(h_s h_t)}\ 1 \\ \vdots \\ \gamma_{1n(h_s h_t)} \gamma_{2n(h_s h_t)} \cdots \gamma_{nn(h_s h_t)}\ 1 \\ 1 \quad 1 \dots 1 \quad 0 \end{bmatrix}$, $\lambda = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \\ Tc \end{bmatrix}$, $\mathrm{b} =$

$\begin{bmatrix} \gamma_{01(h_s,h_t)} \\ \gamma_{02(h_s,h_t)} \\ \vdots \\ \gamma_{0n(h_s,h_t)} \\ 1 \end{bmatrix}$. λ is the weight matrix. γ is the spatiotemporal coefficient of variation. $h_s = \sqrt{\Delta x^2 + \Delta y^2}$ is the spatial distance between sample points. h_t is the time distance between sample points.

Spatial statistics assumes that regionalized variables have spatial limitations, continuity, and anisotropy. The continuity of different regions is described by a semi-variogram. The main semi-variogram models are the spherical model, exponential model, Gaussian model, power function model, and parabolic model. When a single model cannot express, two separate models can be used for simulation and combined. To explain the physical meaning of space, time, and space-time, the spatiotemporal variogram used in this paper is the Bilonick spatiotemporal separation model:

$$\gamma(h_s, h_t) = \gamma_s(h_s) + \gamma_t(h_t) + \gamma_{st}(h_{st}) \quad (12)$$

$\gamma_s(h_s)$ is the spatial variogram. $\gamma_t(h_t)$ is the time variogram. $\gamma_{st}(h_{st})$ is the spatiotemporal variogram. $h_{st} = \sqrt{{h_s}^2 + (\varphi \cdot h_t)^2}$. φ is the spatiotemporal set divergence ratio. After merging and simplifying, the spatial-temporal variogram is as follows:

$$\gamma(h_s, h_t) = C_0 + C_t * \left(\frac{3h_t}{2a_t} - \frac{{h_t}^3}{2{a_t}^3}\right) + a_s * h_s + C_{st}\left(\frac{3h_{st}}{2a_{st}} - \frac{{h_{st}}^3}{2{a_{st}}^3}\right) \quad (13)$$

C_0, C_t, a_t, a_s, C_{st}, a_{st} and φ are the seven parameters to be solved.

3. TEST AND ANALYSIS

3.1 Overview of Data Sources

The terrain of Beijing is high in the northwest and low in the southeast. Xishan is in the west, belonging to Taihang Mountains. Jundu mountain is in the north and northeast, belonging to Yanshan Mountains. The highest peak is Dongling Mountain in Mentougou District, West Beijing. The lowest ground is the southeast boundary of Tongzhou District. The average altitude of Beijing is 43.5 meters. The altitude of Beijing Plain is 20-60m, and that of the mountainous area is 1000-1500m. The change process of PM2.5 includes occurrence and evolution. Moreover, the evolution includes diffusion and attenuation. There are 35 air quality monitoring stations in Beijing, including 12 urban environmental assessment stations, 11 suburban environmental assessment stations, seven control points and regional stations, and five traffic pollution monitoring stations. They cover six districts and ten counties in Beijing, and the coverage can reflect the air quality of the whole Beijing area. In the data obtained in this paper, due to too much missing data of the botanical garden monitoring station, the monitoring station was deleted, and the remaining 34 monitoring stations were used. The data source of this paper is the website of the Beijing environmental protection monitoring center (https://beijingair.sinaapp.com). The daily hourly mean PM2.5 concentration data of 34 monitoring stations in Beijing in February 2020 were published.

3.2 Simulation and Prediction

It is assumed that the diffusion of pollutants is from high concentration areas to low concentration areas under meteorological conditions in this paper. Therefore, when superimposing multi-source diffusion, it is necessary to consider comparing between the monitoring station concentration and the diffusion point concentration. If the concentration of the diffusion point is low, the effect on the increase of the concentration of the monitoring station is small. It can be seen that three main factors affect the efficiency of pollutant diffusion calculation: the number of pollution sources, the size of the calculated pollution area, and the spatial resolution of the model. In this paper, monitoring stations and external pollution sources are set as point source pollution sources for a pollution accumulation process from 13:00 on February 8 to 6:00 on February 9. Then the number of pollution sources is N = 42 (including eight external pollution sources), the time is M = 24, and the number of monitoring stations is R = 34. The number of unknown parameters is 45. Because M < N in formula (5), we can only use an optimization algorithm to solve unknown parameters. In this paper, the improved hybrid genetic algorithm is used to calculate the amount of pollution.

Because the external pollution is not clear, eight external pollution sources are initially assumed in eight directions of Beijing. When the pollution source is set in the next iteration, the direction with a proportion less than 0.1 is excluded. Iteration stops until the result of the next iteration is consistent with that of the previous iteration. Thus, the position of the pollution source can be roughly determined. For example, in this experiment, the proportion of pollution in the north and northeast directions is less than 0.1 in the first iteration (Figure. 1). The pollution sources in these two directions are eliminated. At the beginning of the second iteration, there are 40 pollution sources, including six external pollution sources. The results of the second iteration are shown in Figure 2, and the pollution sources in the northwest and East directions are eliminated again. In the third iteration, there are 38 pollution sources, including four external pollution sources, and so on. It can be known that the main external pollution sources in Beijing during the epidemic period were the pollution sources in the southwest and Southeast directions (Figure. 3). After data inquiry, there are two large point source emissions of heavy industry in the southwest and southeast of Beijing. Therefore, it is inferred that the pollution is due to the diffusion and convergence of pollutants from these two directions to the urban area of Beijing. Moreover, the pollutants from these two directions are superposed with the pollution emissions from the urban area of Beijing. With the low wind speed, high humidity, and low atmospheric boundary layer height, the heavy pollution process of particulate matter is caused by the heterogeneous explosive growth of nitrate.

 https://doi.org/10.1051/e3sconf/202339303008

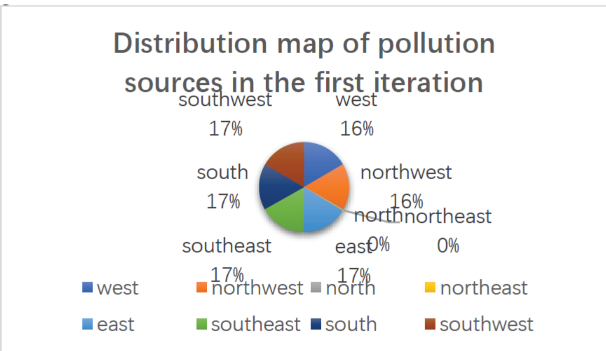

Figure 1. Distribution map of pollution sources in the first iteration

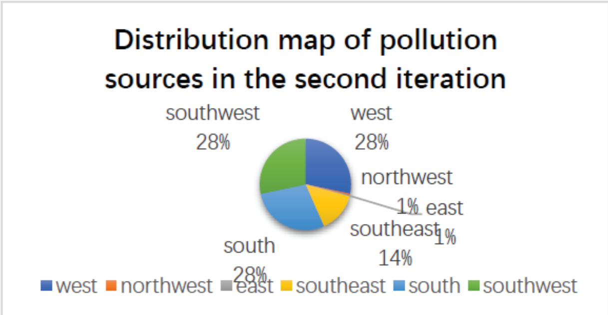

Figure 2. Distribution map of pollution sources in the second iteration

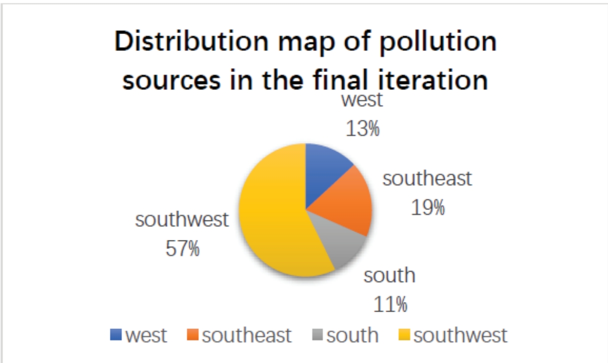

Figure 3. Distribution map of pollution sources in the final iteration

Prior knowledge (remote sensing data) can roughly infer that the external pollution sources in Beijing come from the southwest and Southeast directions, which is roughly consistent with the experimental results.

In this paper, the spatiotemporal Kriging method can be used to interpolate the spatiotemporal distribution of pollutants at any time and show the spatiotemporal change process of pollution superposition. In this paper, if we directly use all the data for spatiotemporal dynamic interpolation, we will find that the calculation matrix is particularly large. Even so, there is still a large error between the final predicted value and the actual monitoring value. Through the analysis, it is found that the correlation of some data between the time series data of different monitoring stations and the monitoring stations series data of different times is not strong, and some are even negative. It is mainly due to the large administrative scope of Beijing and the inconsistency of atmospheric turbulence activities between the central area of Beijing and the suburbs out of the sixth ring road. Therefore, to improve the accuracy of spatiotemporal interpolation and prediction, the selection of interpolation samples for each contribution point is improved in the spatiotemporal Kriging method. Among them, the process of selecting the point with the greatest contribution to each prediction point is as follows:

The first step is to screen candidate points. The nearest monitoring station $Z^{\circ}(i, t0)$ to the prediction point Z（x0, y0, t0） is obtained by calculating the space-time distance. Calculate the correlation coefficient R (y0, yi) between the time series y0 of the i station and the time series yi of other stations (Fig. 7), and calculate the correlation R (t0, Tj) between the spatial series data Ti at t0 and other spatial series data. Sort R (y0, yi) (i=1,2……34) and R (t0, Tj)（j=1,2……24） in reverse order. C points of time or space series with the largest correlation coefficient are selected as candidate points.

Step 2: select k points that contribute the most. Among the C candidate points obtained in the previous step, the space-time distance h_{st} between the prediction point and the candidate point is calculated. All the space-time distances are sorted, and P points with the smallest space-time distance are selected as P points with the greatest contribution.

To predict the future time point t + n of the interpolation point, that is, the time value in the space-time distance increases, the PM2.5 concentration of the future time point n can be predicted.

For the currently predicted time-space point Z (x0, y0, t0), substitute the surrounding data selected in the above steps that make the greatest contribution to the time-space point Z (x0, y0, t0) into formula (10) to calculate the estimation of the currently predicted time-space point. The program interpolates the other time and space points in turn and obtains the PM2.5 concentration distribution map of the Beijing urban area on the hourly time slice.

From 1:00 on February 8 to 24:00 on February 9, the positions of 34 monitoring stations were interpolated as unknown points. RMSE (14), MAE (15), and P (16) were used to evaluate the results of interpolation and prediction.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{M}\sum_{i=1}^{R}(Z^{*}(i,t)-Z(i,t))^{2}} \tag{14}$$

$$\mathrm{MAE} = \frac{\sum_{t=1}^{M}\sum_{i=1}^{R}|Z^{*}(i,t)-Z(i,t)|}{n} \tag{15}$$

$$\mathrm{p} = 1-\frac{\sum_{t=1}^{M}\sum_{i=1}^{R}|Z^{*}(i,t)-Z(i,t)|}{\sum_{t=1}^{M}\sum_{i=1}^{R}Z(i,t)} \tag{16}$$

$Z^{*}(i,t)$ is the interpolation or prediction value. $Z(i,t)$ is the monitoring value. M is the number of times. R is the number of monitoring stations. Moreover, n is the number of interpolation data or prediction data.

Using the Spatio-temporal data of 24 hours on February 9th, the PM2.5 concentration of next 1 hour was predicted (Figure. 5). Moreover, cross-validation (TABLE I) was carried out.

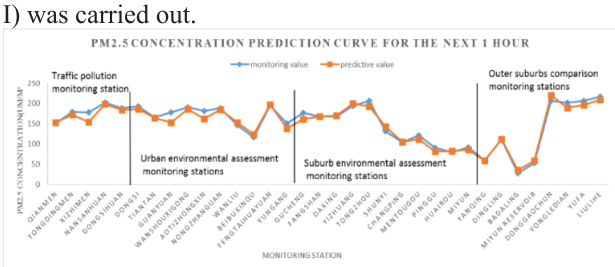

Figure 4. PM2.5 concentration prediction curve for the next 1 hour

TABLE I. CROSS-VALIDATION RESULTS OF 6-HOUR PREDICTION

Time	RMSE	MAE	P
next 1 hour	3.87	16.83	94.73%

4. CONCLUSIONS

In this paper, a model was established to simulate the diffusion process of air pollution in Beijing during the epidemic period in February. The improved hybrid genetic algorithm is used to settle the parameters in the multi-pollution source diffusion model, and the model parameters are obtained. Using the spatiotemporal Kriging method, the PM2.5 concentration at different time points in other locations can be obtained by hourly spatiotemporal interpolation of the heavy pollution process, which can simulate the evolution process of heavy pollution and can be used to interpolate and predict the PM2.5 concentration hourly in the future. Quantitative and qualitative analysis was taken for the result data. Using the actual monitoring data to verify the prediction data, to prove the effectiveness of the spatiotemporal dynamic interpolation simulation and prediction method. The results show that the model and method constructed in this paper can well simulate and predict the change of PM2.5 concentration on the hourly scale and can provide a reference for the analysis, simulation, and prediction of air pollution.

In this paper, the basic parameters lack a unified and perfect theoretical basis when the modified hybrid genetic algorithm is used to fit the variation function of Time-Space Theory. At the same time, it is difficult to get the best fitting result in a single fitting of the spatiotemporal theoretical model, which requires multiple fitting or manual debugging. At present, the fitting result is only a better fitting result. Therefore, in the next step of research, we can combine this algorithm with other intelligent optimization algorithms to train the data of longer time series to improve the prediction accuracy.

REFERENCES

1. Gautam, S., Patra, A. K., & Kumar, P.Status and chemical characteristics of ambient PM2.5 pollutions in China: a review. Environment, Development and Sustainability. 2018.
2. Guo H, Li W, Wu J. Ambient PM2.5 and Annual Lung Cancer Incidence: A Nationwide Study in 295 Chinese Counties. International Journal of Environmental Research and Public Health. 2020,17(5):1481.
3. Barrett, Julia R.. Subtle but Potentially Serious: Long-Term Ambient PM2.5 Exposure and Risk of Cardiopulmonary Mortality. Environmental Health Perspectives. 2020,128(2): 024001-1-024001-2.
4. Luo Yi, Teng Mengfan, Yang Kun, Zhu Yu, Zhou Xiaolu, Zhang Miao, Shi Yan, Research on PM2.5 estimation and prediction method and changing characteristics analysis under long temporal and large spatial scale - A case study in China typical regions, Science of The Total Environment, 2019, (696-):696.
5. Chen, Sheng, Gao, Miao, Hai, Gao, & Gao. The Effects of the Trans-Regional Transport of PM2.5 on a Heavy Haze Event in the Pearl River Delta in January 2015. Atmosphere, 2019,10(5):237.
6. Chen, X., Yin, L., Fan, Y., Song, L., Ji, T., Liu, Y., … Zheng, W. Temporal evolution characteristics of PM2.5 concentration based on continuous wavelet transform. Science of The Total Environment, 2019,134244.
7. Xie, Y., Wang, Y., Zhang, K., Dong, W., Lv, B., & Bai, Y. Daily Estimation of Ground-Level PM2.5 Concentrations over Beijing Using 3 km Resolution MODIS AOD. Environmental Science & Technology, 2015,49(20), 12280–12288.
8. Liu, H., Duan, Z., & Chen, C. A hybrid framework for forecasting PM2.5 concentrations using multi-step deterministic and probabilistic strategy. Air Quality, Atmosphere & Health, 2019,12(7),785–795.
9. Junfei Qiao, Zengzeng He, Shengli Du. Prediction of PM2:5 concentration based on weighted bagging and image contrast-sensitive features. Stochastic Environmental Research and Risk Assessment. 2020,34(3-4):561-573.
10. Danesh Yazdi, M., Kuang, Z., Dimakopoulou, K., Barratt, B., Suel, E., Amini, H., … Schwartz, J. Predicting Fine Particulate Matter (PM2.5) in the Greater London Area: An Ensemble Approach using Machine Learning Methods. Remote Sensing, 2020,12(6), 914.
11. Hua, Z., Sun, W., Yang, G., & Du, Q. A Full-Coverage Daily Average PM2.5 Retrieval Method with Two-Stage IVW Fused MODIS C6 AOD and Two-Stage GAM Model. Remote Sensing, 2019.11(13),1558.
12. Han, & Tong.Satellite-Based Estimation of Daily Ground-Level PM2.5 Concentrations over Urban Agglomeration of Chengdu Plain. Atmosphere, 2019,10(5),245.
13. Zamani Joharestani, M., Cao, C., Ni, X., Bashir, B., & Talebiesfandarani, S. PM2.5 Prediction Based on Random Forest, XGBoost, and Deep Learning Using Multisource Remote Sensing Data. Atmosphere, 2019,10(7),373.