A Municipal PM2.5 Forecasting Method Based on Random Forest and WRF Model

Nan Jiang , Fei Fu , Hua Zuo , Xiuping Zheng and Qinghe Zheng

***Abstract*—In the recent years, air pollution is a very serious problem in China and elsewhere, and it is a factor that significantly affects the quality of human health. Fine particulate matter (PM2.5) is considered to be the culprit of haze weather. Therefore, research affects the quality of human life on PM2.5 forecasting has received increasing attention. Knowing this information in advance is very important to protect humans from health problems. This paper proposes a new prediction method, using the predicted value of the Weather Research and Forecasting (WRF) model as input data, increasing the atmospheric inversion factor as an additional input factor and constructing a municipal atmospheric pollutant response model through a random forest algorithm. we use 3-fold cross-validation (CV) to evaluate model performance. The result of the experiment shows, compared with the traditional atmospheric simulation method, this method has practical application significance. The simulation results have improved timeliness and accuracy. It provides a simple and effective method for PM2.5 prediction.**

***Index Terms*—air pollution prediction (forecasting), PM2.5, random forest, weather research and forecasting (WRF) model,**

I. Introduction

With the rapid industrialization and economic development, the environmental problems of cities (especially big cities) are highly valued [1, 2]. Air pollution, as a major problem of environmental pollution [3], has a very negative impact on human health and social development [5]. Exposure to fine particulate matter (PM2.5) [8, 9] remains a worldwide public health issue [6].

We can ensure good air quality in cities in a number of ways [7]: continuously monitor air quality by using smart sensors and developing systems that predict air quality. Affected by many factors, the content of air pollutants often changes dramatically [8], sometimes it can be normal, and at another time it can get worse. Therefore, it is very important to provide an accurate method to predict the air quality in the future [9]. Most air pollution simulation methods rely on mathematical simulations [10]. These methods depend on atmospheric dynamics captured by various generation models, simulating the physical and chemical diffusion process of atmospheric pollutants in the air, and the dynamic changes of pollutants are calculated by computer [11, 12].

A. Challenges

Although the atmospheric transport model prediction method has achieved long-term development in recent decades, it still lacks performance in some aspects [14]:

Firstly, numerical prediction methods are often applied to large-area air quality transmission simulations. For urban-scale air quality predictions, the prediction results are often unsatisfactory due to complex surface conditions.

Secondly, the numerical forecasting model requires detailed pollution source data and meteorological data like the air pollutant emission inventory. In practical applications, many cities do not have complete data. These data are often difficult to collect and the accuracy of prediction is easily affect.

Thirdly, the numerical forecasting model has a high computational complexity, a large computational power requirement, and a long calculation time.

At present, with the rise of artificial intelligence [16,17,18], statistical prediction methods based on machine learning models have become one of the effective tools for studying air pollution. Compared with the traditional method (mechanical prediction method of atmospheric transmission model), the deep learning model is simple and fast. In the case of detailed statistics, it is easier to excavate the response relationship between the factors affecting the concentration of pollutants [19]. However, as a statistical simulation method, the defects are also obvious [29], that is, the statistical forecasting mostly uses single-site method, and cannot reflect the mutual influence characteristics between the stations. Moreover, when the pollution source changes greatly, the statistical model is difficult to reflect the actual pollution situation [20].

At this stage, the two methods are relatively independent. In view of the advantages and disadvantages of the above two methods [21], it is hoped to provide a more feasible and effective choice for air pollution forecasting in small and medium-sized cities in China.

B. Contributions

This paper proposes a random forest prediction method based on WRF model. The model uses the stereo weather forecast data obtained by WRF model as an input factor to establish a random forest algorithm to complete the prediction. Aiming at the imperfect characteristics of many air quality forecasting theories and methods in small and medium-sized cities, the forecasting system does not need to carry out complex numerical calculations, and has lower computational power requirements, and its timeliness is also significantly

Manuscript received August 18, 2019; revised December 22, 2019. This work was supported by Qingdao Environmental Protection Special Fund.

Nan Jiang , Fei Fu , Hua Zuo , Xiuping Zheng are with the Qingdao Research Academy of Environmental Sciences, Qingdao 266000, China.

Qinghe Zheng (corresponding author) is with School of Information Science and Engineering, Shandong University, Qingdao 266237, China (e-mail: 15005414319@ 163.com).

better than numerical forecasting.

C. Organization

The remainder of the paper is organized as follows. Section II gives a brief review to the related work. Section III introduces the data sources and research process. In Section IV, we introduce the WRF model and random forest. Experimental results and comparisons are presented in Section V. Finally, we conclude our work and future direction in Section VI.

II. Related Work

In the past few years, a large number of studies have been proposed for air quality prediction worldwide. Pablo [13] presents a WRF-Chem model to predict high pollution events. The forecasting system is based on accurately simulating carbon monoxide (CO) as a PM10/PM2.5 surrogate. Since during episodes and within the city there is a high correlation (over 0.95) among these pollutants. Therefore, the simulation of the system does not involve the aerosol chemical model [39]. Zhang [36] developed a source-oriented version of the weather research and forecasting model with chemistry (SOWC, hereinafter). SOWC separately tracks primary particles with different hygroscopic properties rather than instantaneously combining them into an internal mixture. The systems more realistically predict radiative feedbacks from anthropogenic aerosols compared to models that make internal mixing or other artificial mixing assumptions [37]. Yahya [41] has constructed an online-coupled weather research and forecasting model which with chemistry with the model of aerosol dynamics, Reaction, Ionization and dissolution (WRF/Chem-MADRID). The model performs well for O3 and satisfactorily for PM2.5 in terms of both discrete and categorical evaluations. Several authors have used neural networks for predicting PM2.5. Wang [38] put forward an urban air quality forecasting system which based on the new generation of weather research forecast and chemistry model WRF-Chem and a regional haze weather forecasting system based on regional atmospheric environment modeling system (RegAEMS). The results of WRF-Chem performed well in urban air quality forecast on surface concentrations of air pollutants such as SO2, NO2 and PM10. RegAEMS presents relatively good ability in forecast on regional haze weather. Zang [42] develops and demonstrates a multi-scale, three-dimensional variational data assimilation (MS-3DVAR) system for WRF/Chem aerosol simulation. They have implemented MS-3DVAR in WRF/Chem for regional particulate matter (PM) air quality modeling. The results have shown that MS-3DVAR not only significantly improve the prediction of PM2.5 concentration, but also concentrations of individual species. In [24], An empirical study of PM2.5 forecasting using neural network have been proposed to forecast PM2.5 value hourly. Kowalski presented and compared some linear procedures for PM10 and PM2.5 forecasting. Another work [22] uses a deep hybrid model for weather forecasting. It doesn't forecast PM2.5 but predicts other variables like wind and temperature. In [6], an enhanced PM2.5 air quality forecast model based on nonlinear regression (NLR) and back-trajectory concentrations has been developed for use in the Louisville, Kentucky metropolitan area. Wang [25] proposed a statistically reliable and interpretable national modeling framework based on Bayesian downscaling methods with the application to the calibration of the daily ground PM2.5 concentrations across the Continental U.S. using satellite-retrieved aerosol optical depth (AOD) and other ancillary predictors in 2011. In [28], authors have proposed a hidden Markov model to predict 24-hour-average PM2.5 concentrations. SHEN [23] establish the relationship between PM2.5, satellite TOA reflectance, observation angles, and meteorological factors in a deep learning architecture (denoted as Ref-PM modeling). Taking the Wuhan Urban Agglomeration (WUA) as a case study, the results demonstrate that compared with the AOD-PM modeling, the Ref-PM modeling obtains a competitive performance. In [26], This article proposes a space-time LUR model at a regional scale by incorporating aerosol optical depth (AOD) data from the Moderate Resolution Imaging Spectroradiometer (MODIS) [27]. Although machine learning is widely used in many tasks, there is no widely exploited when it comes to PM2.5 predict for small intervals of time. In this article, we attempt to use machine learning algorithm for hourly prediction.

III. Dataset

Qingdao is located in the southeastern part of the Shandong Peninsula. It is a coastal hilly city with two sides facing the Yellow Sea in the southeast. The terrain is east to the west, and the north and south sides are raised and the middle is flat. The total area is 11282 square kilometers, and the total resident population in 2017 is 9.2905 million. Due to the direct regulation of the marine environment and the influence of the southeast monsoon, current and water mass from the ocean, the urban area has a significant marine climate. In spring, the temperature rises slowly, one month later than inland; in summer, it is humid, hot and rainy, but not hot and hot; in autumn, it is cool, with little precipitation and strong evaporation; in winter, the wind is warm and low, with a long duration. The annual average is 22 days. The annual average rainfall is 662.1mm, and the rainfall in spring, summer, autumn and winter accounts for 17%, 57%, 21% and 5% of the annual rainfall respectively. The annual precipitation is 1272.7 mm (1911) at most, 308.2 mm (1981) at least. The annual variability of precipitation is 62%. The annual average snowfall days is only 10 days. The annual average pressure is 1008.6 mbar. The annual average wind speed is 5.2m/s, with the south east wind as the dominant wind direction. The annual average relative humidity is 73%, the highest is 89% in July and the lowest is 68% in December. This paper chooses the single-site model as the experimental data to establish a forecasting model. The site is located in Fig. 1, and it shows the distribution of monitoring sites in the study area.

The data used in this article includes the following three types of data：

1. Historical meteorological data
2. Historical air pollutant data
3. Meteorological simulation data

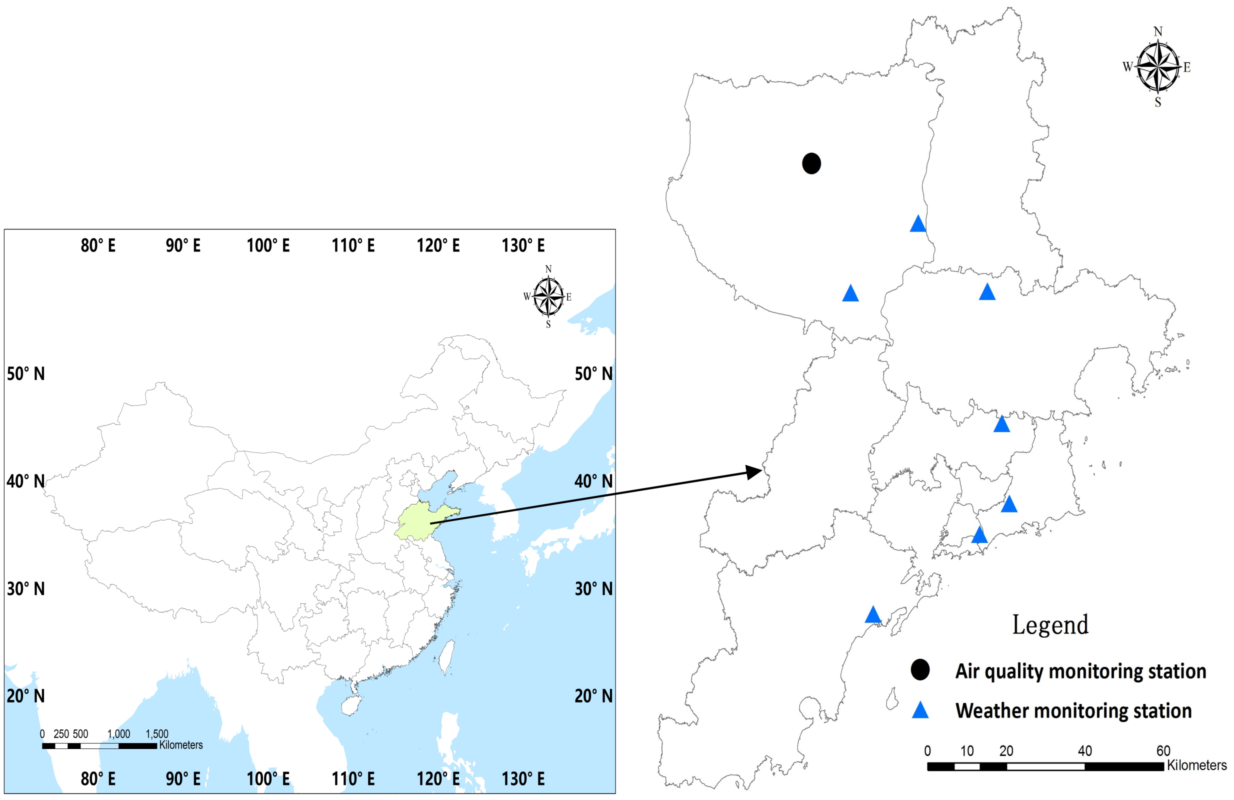

Fig. 1 Research area.

The historical air pollutant data are the hourly average pollutant concentration data from 1/1/2019 to 2/1/2019, including the mean value of PM2.5, PM10, SO2, NO2, O3 concentration. The historical meteorological data are the monitoring data of meteorological stations in the whole city, including hourly wind speed, air temperature, relative humidity and air pressure of each meteorological station. The weather simulation data is a grid calculated by processing the historical weather data through the WRF model.

The hourly air pollutant data are matched with the grid point data of the nearest WRF model at the time node, and the meteorological data of the corresponding station are obtained. Model using 80% of the dataset as the training set and 20% as the test set.

IV. Forecasting Method

The main objective of this paper is to obtain the regression relation between the concentration of PM2.5 and other factors by the random forest algorithm. This process can be summarized as the following steps: Obtain hourly atmospheric PM2.5 pollutant data and historical meteorological data from the site (temperature, wind speed, humidity, atmospheric thermal stability factor, etc.); Forecasting weather factor conditions by time through WRF model. Using the PM2.5 concentration as the model dependent variable, the influence factor of the PM2.5 concentration in the region was taken as input; the regression relationship was generated by the random forest algorithm. Finally, we construct the model to predict PM2.5 and compare the errors of different factors on the verification set. The entire running flow is shown in Fig. 2.

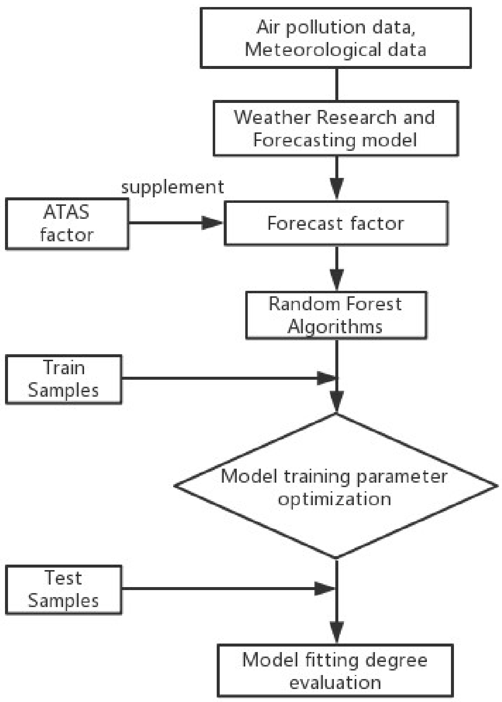

Fig. 2. Experimental process.

A. *Weather research and forecasting (WRF) model*

It is the latest regional atmospheric dynamic-chemical coupling model developed by the center for atmospheric research in the United States. This model can realize the complete coupling of chemical model and meteorological model in space-time resolution, that is, to realize the real on-line transport simulation of pollutants without interpolation [30]. WRF model not only has good prediction effect, but also has wide applicability and convenient expansibility [31].

In the field of air pollution prediction, WRF/chem and WRF-CMAQ based on WRF model are widely used.

Based on the mesoscale meteorological model such as WRF and air quality model CAMX, the simulation area and grid are delimited with Qingdao as the center, and a model system suitable for Qingdao air quality simulation is built. The specific parameters are as follows:

Simulation range: The simulation grid adopts four layers of grid nesting settings, the outermost area includes all the land areas of China, the grid resolution is 27km, the number of grid rows is 158, and the number of columns is 204; the second layer is 9km, the number of grid rows is 146, and the number of columns is 146, and the third layer is part of Qingdao and its surrounding cities, the grid resolution is 3km, and the grid rows are 204. The number is 74, the number of columns is 92, and the simulation range is shown in the figure below.

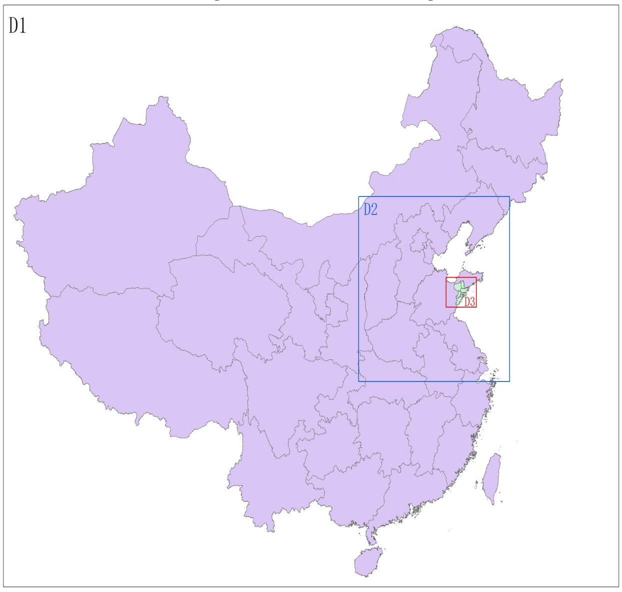

Fig. 3. Three layer nesting diagram of simulation scope.

WRF model parameters: The mesoscale meteorological model WRF provides the meteorological field for the air quality model CAMX, and the WRF model and CAMX model use the same simulation period and spatial projection coordinate system. The simulation range of WRF is slightly larger than that of CAMX. There are 35 air pressure layers in the vertical direction, and the layer spacing increases gradually from bottom to top. The parameterization scheme of mesoscale meteorological model WRF is optimized through sensitivity experiment comparison. NCEP and FNL reanalysis data is used for large-scale background field and boundary condition, and ADP high altitude and ground observation data are used for objective analysis and optimization of the first guess field. The simulated wind speed, temperature, relative humidity and other data are compared with the observation data. The simulation data and observation data are highly correlated. The parameterization scheme is shown in the Table 1.

CAMX model parameters: The CAMX model uses cb05 as the gas-phase chemical mechanism, CF as the aerosol chemical mechanism, ppm as the horizontal advection scheme, and Zhang03 as the dry deposition scheme. The default boundary field is used as the boundary field, and the default initial field is used as the initial field on the first day, and then the simulation results of the previous day are used as the initial field.

Emission inventory: The list of air emission sources in Qingdao is based on the recommended methods and data of the technical guide for the preparation of the list of air pollutant emission sources issued by the Ministry of ecological environment, and is calculated by combining the data obtained through enterprise research, sampling analysis and other means, including the fixed combustion source, mobile source and process. Sources, solvent use sources, dust sources, storage and transportation sources, farming and animal husbandry sources, waste disposal sources, biomass combustion sources, other sources, etc. the list of anthropogenic emissions in areas outside Qingdao was processed by MEIC in 2016. The list of natural emissions in the whole simulation area was processed by Megan model, and the leaf area index data and vegetation type data retrieved by MODIS satellite were used. According to the weather conditions simulated by WRF, the spatial fusion of the emission data from man-made sources and natural sources is carried out to obtain the emission inventory with high spatial-temporal resolution under the benchmark scenario applicable to CAMX model.

TABLE 1
WRF parameterization scheme

Parameterization scheme	Selected scheme name
Microphysical process scheme	Lin
Long wave radiation scheme	RRTM
Short wave radiation scheme	Dudhia
Near formation plan	MM5 Similarity
Land surface process plan	5-layer thermal diffusion
Boundary layer scheme	YSU
Cumulus convection scheme	Kain-Fritsch

B. *Random forest algorithm*

The random forest algorithm was proposed by Breiman and Cutler in 2001 as an integration of decision tree algorithms, improving model longitude by generating a large number of classification trees. Compared with traditional machine learning methods such as neural networks, random forest algorithms are fast and perform well when dealing with big data. In addition, random forests can well reflect the nonlinear effects between variables [32].

The random forest [33] regression algorithm is composed of decision trees related to random vector θ,The numerical variables of the model are taken as dependent variables, and it is assumed that the training set is extracted independently from the distribution of random vector Y and X .The homogenization error of any numerical predicted value is $E_{X,Y}\left(Y-h(X)\right)^2$ .The mean value of the k regression was used as the prediction result for the model. The main steps of the algorithm are as follows:

(1) Record the original dataset as $T=\{(x_1,y_1),(x_2,y_2),\cdots,(x_3,y_3)\}$,the generated random

vector sequence is recorded as $\theta_i\left(i=1,\cdots k\right)$,subsample set is randomly obtained from the original data T by bootstrap sampling, which is recorded as $T_i\left(i=1,2,\cdots k\right)$;

(2) Establish a regression model for each subsample set separately $\{h(X,\theta_i),i=1,\cdots,k\}$, where matrix X is the independent variable for modeling, assume that parameter set $\{\theta_k\}$ is independently and identically distributed.

(3) After the k -round training, the average of the k -round results is used as the prediction result of the random forest model. The forecast result is:

$$f_r\left(x\right)=\frac{1}{k}\sum_{i=1}^{k}h_i\left(x\right) \tag{1}$$

In the formula, $f_r\left(x\right)$ represents the result of RFR model, h is the result of a single regression tree model. The random sample set can be obtained by Bootstrap sampling, and the regression tree modeling is carried out, which increases the difference between the models and improves the extrapolation prediction ability of the model.

The random forest algorithm can obtain better prediction results without complex parameter adjustment, can effectively process the non-linear process, and shows great advantages in preventing over-fitting. This paper chooses the random forest algorithm as the PM2.5 concentration prediction study.

C. Forecast factor analysis

The WRF model is used to simulate the meteorological situation of Qingdao. Figures B1-B6 in Appendix B are the forecast process form each algorithm. For air pollution, Surface meteorological factors (wind speed, temperature, relative humidity, air pressure) have great influence on the transfer, transmission and diffusion of PM2.5 in the region. Considering extracting the basic surface meteorological factors affecting PM2.5 concentration from the acquired experimental data.

Table 2 lists the basic surface meteorological factors contained in the experimental data. Figures A1-A5 in Appendix A are the variation range of each factor. Appendix C is the correlation between the factors.

TABLE 2
Summary of basic weather forecast factors

Factor	Range	Mean Value
Wind speed /ms	[0.111-11.68]	3.38
temperature /℃	[-4.4-15.3]	3.49
Relative humidity /(%)	[7.75-79.71]	40.18
Air pressure /hPa	[1004.75,1023.38]	1013.57
RAIN/mm	[0, 0.000324]	0.000008

The diffusion process of atmospheric pollutants is not only related to surface meteorological factors, but also the vertical pressure state affects its changes. Previous studies have shown that high air pressure instability will affect the flow of the atmosphere, making it difficult to spread pollutants on the surface. This paper uses A as a parameter to measure the state of pressure [34].

The atmospheric temperature of the monitoring point at 500 *hPa*, 700 *hPa*, 850 *hPa* and 925 *hPa* is obtained by the WRF model as T_{925} , T_{850} , T_{700} and T_{500} ,Its corresponding atmospheric dew point temperature is T_{d925} ,T_{d850} ,T_{d700} and T_{d500} .

$$A=T_{925}-T_{500}-\left[(T_{925}-T_{d925})+(T_{850}-T_{d850})+(T_{700}-T_{d700})\right] \tag{2}$$

When A becomes larger, the atmospheric state becomes unstable, and the surface diffusion conditions become better. On the contrary, the atmospheric state is stable and the surface diffusion conditions are poor.

D. Forecast accuracy evaluation index

The *K*-fold cross-validation method is an effective data verification method. It cuts the data samples into smaller subsets, making the best use of the finite number of samples to reduce errors. And effectively avoid over-learning and under-learning. The method is as follows: The original data set is divided into *K* groups, one of which is used as a test set each time, and the remaining *K*-1 group is used as a training set to train the model, and a total of *K* models are obtained. Finally, the average of the classification rates of the *K* models is used as the true classification rate of the hypothesis function. This paper takes *K*=3 to construct and verify the model. The model error evaluation indicators used in this paper include root mean square error (RMSE), coefficient of determination (MAE) [35]. The calculation was made by using the following equations:

$$RMSE=\sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i-\hat{y}_i)^2} \tag{3}$$

$$MAE=\frac{1}{m}\sum_{i=1}^{m}\left|y_i-\hat{y}_i\right| \tag{4}$$

where y_i is the measured value of air quality, $\hat{y}_i$ is the predicted value of a certain time in model.

V. Experiments and Results Analysis

The PM2.5 concentration of the selected air quality monitoring station throughout the month is shown as figure 4. The average PM2.5 concentration per hour is 115 ug/m3, and the high PM concentration data makes the data fluctuate greatly. It can be seen from the figure that the proportion of heavy pollution weather is large. In a short time, the concentration of PM2.5 changes sharply, which may be caused by the interference of external measures.

In this paper, the short-term prediction is carried out, The predicted value is the concentration of PM2.5 after 1 hour. In addition to the comparison between machine learning models, the experiment also added the SVM model to do the contrast experiment, and the prediction result is shown in the Table 3. The results show that random forest is better than SVM in PM2.5 prediction. Although we only need to use the existing meteorological data and do not need to calculate the meteorological simulation data by WRF model, it is clearly that good prediction results can be obtained in this time span.

In order to test the prediction effect of the model in different time span, the data of the next 1-24 hours are selected to predict. Different from the PM2.5 concentration after the prediction of 1 h, the long-term prediction needs to simulate the future weather factor as an input factor through the WRF model.

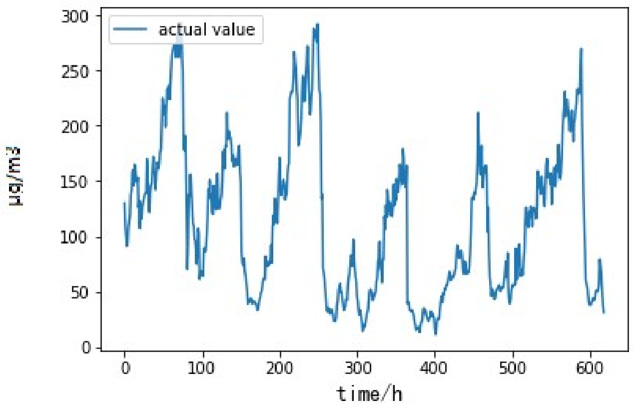

Fig. 4. PM2.5 concentration trend.

TABLE 2
Summary of basic weather forecast factors

Factor	Range	Mean Value
Random Forest	17	12.07
SVM	23	17

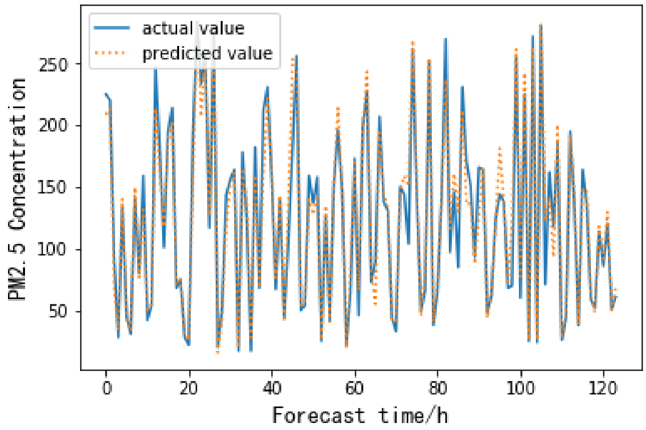

Fig. 5. Forecast result for 1 hour.

We construct two different combinations of forecasting factors and calculate their error evaluation indicators respectively.

The combination of two different predictors is as follows:

1. General weather forecasting factor.

2. General weather forecasting factor + atmospheric thermal stability state factor.

Figure 6 and figure 7 show the RMSE and MAE changes in test set of the PM2.5 concentration real-time forecasting model for two different combinations of forecasting factors.

Obviously, with the increase of time, the prediction error also increases. This is because each prediction result is affected by the previous prediction result, that is, error transmission.

It can be seen that the combined accuracy of the atmospheric thermal stability factor of the basic prediction factor is obviously higher than that of the basic prediction factor. It is indicated that the atmospheric thermal stability state factor is one of the important factors affecting the diffusion of pollutants. Taking the atmospheric thermal stability state factor as an additional factor can effectively improve the accuracy of model prediction.

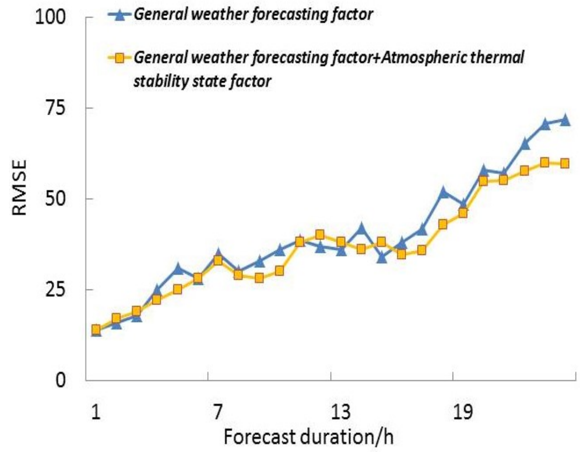

Fig. 6. RMSE trend.

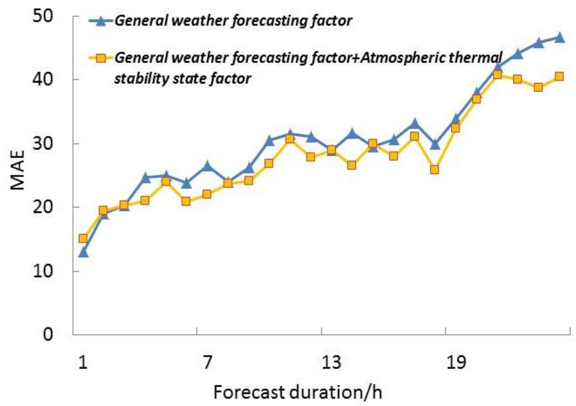

Fig. 7. MSE trend.

VI. Conclusion and Future Work

PM2.5 concentration forecast is an important aspect of current air pollution control. In this paper, we use the random forest method to fuse the WRF model, use the atmospheric pollutant concentration and basic meteorological factors training model, and add the atmospheric thermal stability factor as an additional factor to model and predict the municipal PM2.5 concentration. The experimental results show that the method has achieved certain results in the short-term and medium-term forecasting applications. However, under the prediction of long-term span, the prediction error increases with the influence of the law of error propagation. How to improve the long-term prediction accuracy will be the direction of the next study. Moreover, we plan to apply the proposed method to more time series structures in the future, such as long short-term memory (LSTM).

References

[1] Zhang C , Yuan D . Fast Fine-Grained Air Quality Index Level Prediction Using Random Forest Algorithm on Cluster Computing of Spark[C]// Ubiquitous Intelligence & Computing & IEEE Intl Conf on Autonomic & Trusted Computing & IEEE Intl Conf on Scalable Computing & Communications & Its Associated Workshops. IEEE, 2016.

[2] Cobourn W G . An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations[J]. Atmospheric Environment, 2010, 44(25):3015-3023.

[3] Yu Z, Yi X, Ming L, et al. Forecasting Fine-Grained Air Quality Based on Big Data[C]// Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. 2015. Hu X, Belle J H, Meng X, et al.

Estimating PM2.5 Concentrations in the Conterminous United States Using the Random Forest Approach[J]. Environmental Science & Technology, 2017, 51(12):6936.
[4] Wang Y, Hu, Xuefei, Chang, Howard, et al. A Bayesian Downscaler Model to Estimate Daily PM2.5 Levels in the Conterminous US[J]. International Journal of Environmental Research and Public Health, 2018, 15(9).
[5] Wang Y, Hu, Xuefei, Chang, Howard, et al. A Bayesian Downscaler Model to Estimate Daily PM2.5 Levels in the Conterminous US[J]. International Journal of Environmental Research and Public Health, 2018, 15(9).
[6] Zheng Y , Li S , Zou C , et al. Analysis of PM2.5 concentrations in Heilongjiang Province associated with forest cover and other factors[J], 2019.
[7] Q. Zheng et al., "An end-to-end image retrieval system Based on gravitational field deep learning," in IEEE International Conference on Computer Systems, Electronics and Control (ICCSEC), Dalian, China, pp. 936-940, 2017.
[8] Mölders N, Tran H N Q, Cahill C F, et al. Assessment of WRF/Chem PM 2.5 forecasts using mobile and fixed location data from the Fairbanks, Alaska winter 2008/09 field campaign[J]. Atmospheric Pollution Research, 2012, 3(2):180-191.
[9] Q. Zheng et al., "Understanding and boosting of deep convolutional neural network based on sample distribution," in Proc. IEEE. ITENC, Chengdu, China, pp. 823-827, 2017.
[10] Xin-Hai L I. Using "random forest" for classification and regression[J]. Chinese Journal of Applied Entomology, 2013.
[11] Kebin H E, Hong H, Qiang Z. Urban air pollution in China: Current status, characteristics, and progress[J]. Annual Review of Energy & the Environment, 2011, 27(1):397-431.
[12] Keyong H , Qingyang X , Xia M , et al. Predicting monthly high-resolution PM 2.5, concentrations with random forest model in the North China Plain[J]. Environmental Pollution, 2018:S0269749118319997-.
[13] Q. Zheng et al., "Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process," IEEE Access, vol. 6, pp. 15844-15869, 2018.
[14] Q. Zheng et al., "Fine-grained image classification based on the combination of artificial features and deep convolutional activation features," in IEEE/CIC ICCC, pp. 1-6, Qingdao, China, 2017.
[15] Mao L , Qiu Y , Kusano C , et al. Predicting regional space–time variation of PM2.5with land-use regression model and MODIS data[J]. Environmental Science and Pollution Research, 2012, 19(1):128-138.
[16] Q. Zheng and M. Yang, "A Video Stabilization Method based on Inter-Frame Image Matching Score," Global Journal of Computer Science and Technology, vol. 17, no. 1, pp. 35-40, 2017.
[17] Q. Zhang et al., "Segmentation of hand gesture based on dark channel prior in projector-camera system," in IEEE/CIC ICCC, pp. 1-6, Qingdao, China, 2017.
[18] Spiridonov V , Jakimovski B , Spiridonova I , et al. Development of air quality forecasting system in Macedonia, based on WRF-Chem model[J]. Air Quality, Atmosphere & Health, 2019.
[19] Ling-Jyh Chen, Wuulong Hsu, MingWei Cheng,. Demo:LASS: A Location-Aware Sensing System for Participatory PM2.5 Monitoring[C]// International Conference on Mobile Systems. 2016.
[20] Perez P, Gramsch E. Forecasting hourly PM2.5 in Santiago de Chile with emphasis on night episodes[J]. Atmospheric Environment, 2015, 124:22-27.
[21] Pires J C M , Sousa S I V , Pereira M C , et al. Management of air quality monitoring using principal component and cluster analysis—Part I: SO2 and PM10[J]. Atmospheric Environment, 2008, 42(6):1249-1260.
[22] Tian J , Chen D . A semi-empirical model for predicting hourly ground-level fine particulate matter (PM2.5) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements[J]. Remote Sensing of Environment, 2010, 114(2):221-229.
[23] Hu X , Waller L A , Al-Hamdan M Z , et al. Estimating ground-level PM2.5 concentrations in the southeastern U.S. using geographically weighted regression[J]. Environmental Research, 2013, 121(Complete):1-10.
[24] Q. Zheng et al., "Static Hand Gesture Recognition Based on Gaussian Mixture Model and Partial Differential Equation," IAENG International Journal of Computer Science, vol. 45, no. 4, pp. 569-583, 2018.
[25] Tian J , Chen D . A semi-empirical model for predicting hourly ground-level fine particulate matter (PM2.5) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements[J]. Remote Sensing of Environment, 2010, 114(2):221-229.
[26] Q. Zhang, M. Yang, Y. Zhou, K. Kidiyo, Q. Zheng, and X. Zhang, "Segmentation of hand posture against complex backgrounds based on saliency and skin colour detection," IAENG International Journal of Computer Science, vol. 45, no. 3, pp. 435-444, 2018.
[27] Q. Zheng, X. Tian, M. Yang, and H. Wang, "The Email Author Identification System Based on Support Vector Machine (SVM) and Analytic Hierarchy Process (AHP)," IAENG International Journal of Computer Science, vol. 46, no. 2, pp. 178-191, 2019.
[28] Sun W , Zhang H , Palazoglu A , et al. Prediction of 24-hour-average PM2.5 concentrations using a hidden Markov model with different emission distributions in Northern California[J]. Science of The Total Environment, 2013, 443:93-103.
[29] Na Z, Yang Y, Zhou X. Application of geographically weighted regression in estimating the effect of climate and site conditions on vegetation distribution in Haihe Catchment, China[J]. Plant Ecology, 2010, 209(2):349-359.
[30] You W , Zang Z , Zhang L , et al. Estimating national-scale ground-level PM25 concentration in China using geographically weighted regression based on MODIS and MISR AOD[J]. Environmental Science and Pollution Research, 2016, 23(9):8327-8338.
[31] Q. Zheng, X. Tian, M. Yang, and S. Liu, "Near-infrared Image Enhancement Method in IRFPA Based on Steerable Pyramid," Engineering Letters, vol. 27, no. 2, pp. 352-363, 2019.
[32] H. Zhuang, M. Yang, Z. Cui, and Q. Zheng, "A method for static hand gesture recognition based on non-negative matrix factorization and compressive sensing," IAENG International Journal of Computer Science, vol. 44, no. 1, pp. 52-59, 2017.
[33] Q. Zheng, X. Tian, M. Yang, and H. Wang, "Differential Learning: A Powerful Tool for Interactive Content-Based Image Retrieval," Engineering Letters, vol. 27, no. 1, pp. 202-215, 2019.
[34] You W , Zang Z , Zhang L , et al. Estimating ground-level PM10 concentration in northwestern China using geographically weighted regression based on satellite AOD combined with CALIPSO and MODIS fire count[J]. Remote Sensing of Environment, 2015, 168:276-285.
[35] Q. Zheng et al., "A bilinear multi-scale convolutional neural network for fine-grained object classification," IAENG International Journal of Computer Science, vol. 45, no. 2, pp. 340-352, 2018.
[36] Q. Zheng, X. Tian, N. Jiang, and M. Yang, "Layer-wise learning based stochastic gradient descent method for the optimization of deep convolutional neural network," Journal of Intelligent & Fuzzy Systems, vol. 37, no. 4, pp. 5641-5654, 2019.
[37] Q. Zheng, X. Tian, M. Yang, Y. Wu et al., "PAC-Bayesian framework based drop-p1ath method for 2D discriminative convolutional network pruning," Multidimensional Systems and Signal Processing, early access, 2019. DOI: 10.1007/s11045-019-00686-z
[38] J. Li et al., "Dynamic Hand Gesture Recognition Using Multi-direction 3D Convolutional Neural Networks," Engineering Letters, vol. 27, no. 3, pp. 490-500, 2019
[39] Saide P E , Carmichael G R , Spak S N , et al. Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRF–Chem CO tracer model[J]. Atmospheric Environment, 2011, 45(16):2769-2780.
[40] Zhang H , Denero S P , Joe D K , et al. Development of a source oriented version of the WRF/Chem model and its application to the California regional PM10 / PM2.5 air quality study[J]. Atmospheric Chemistry and Physics, 2014, 14(1):485-503.
[41] Yahya K , Zhang Y , Vukovich J M . Real-time air quality forecasting over the southeastern United States using WRF/Chem-MADRID: Multiple-year assessment and sensitivity studies[J]. Atmospheric Environment, 2014, 92:318-338.
[42] Zang Z , Li Z , Chao Y , et al. Prediction of PM2.5 Using A Multi-Scale Three-Dimensional Variational Data Assimilation System with WRF/Chem during CalNex 2010[C]// Agu Fall Meeting. AGU Fall Meeting Abstracts, 2011.
[43] Shen H , Li T , Yuan Q , et al. Estimating regional ground-level PM2.5 directly from satellite top-of-atmosphere reflectance using deep learning[J]. 2017.

Nan Jiang was born in Jining, Shandong, China in 1992. He received his M.S. degree from Wuhan University of Surveying and mapping engineering in 2016 and Worked in Qingdao Research Academy of Environmental Sciences since 2016. His research direction is Geographic Information System and remote sense.

APPENDIX A

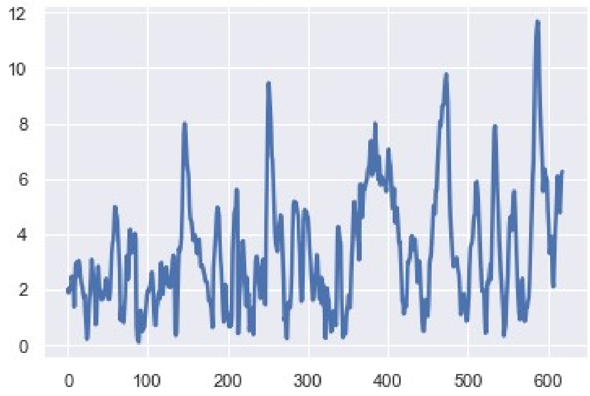

Fig. A1. The TEMP forecast result for 1 hour.

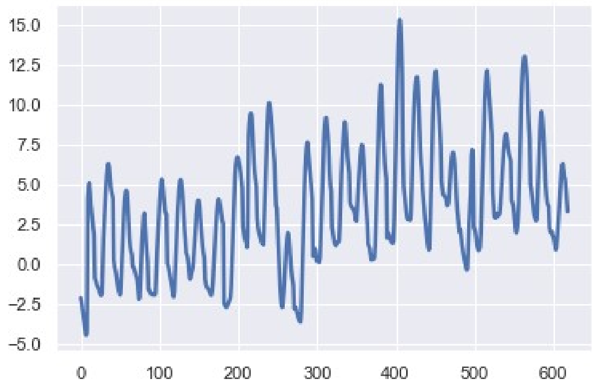

Fig. A2. The TEMP forecast result for 1 hour.

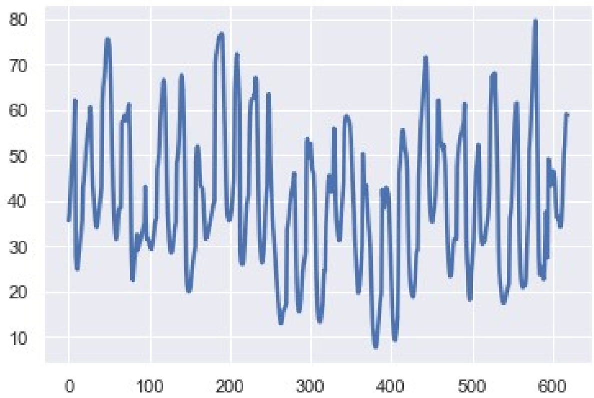

Fig. A3. The TEMP forecast result for 1 hour.

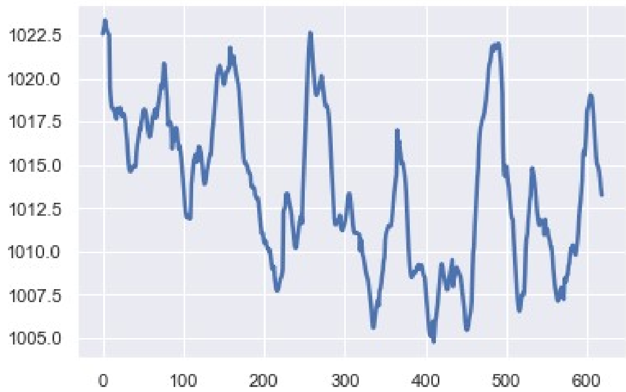

Fig. A4. The TEMP forecast result for 1 hour.

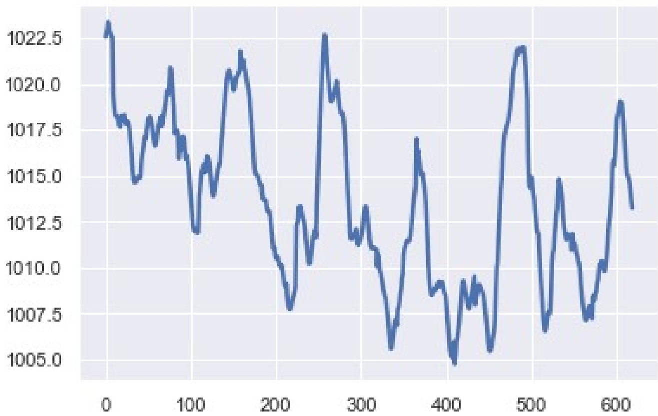

Fig. A5. The TEMP forecast result for 1 hour.

APPENDIX B

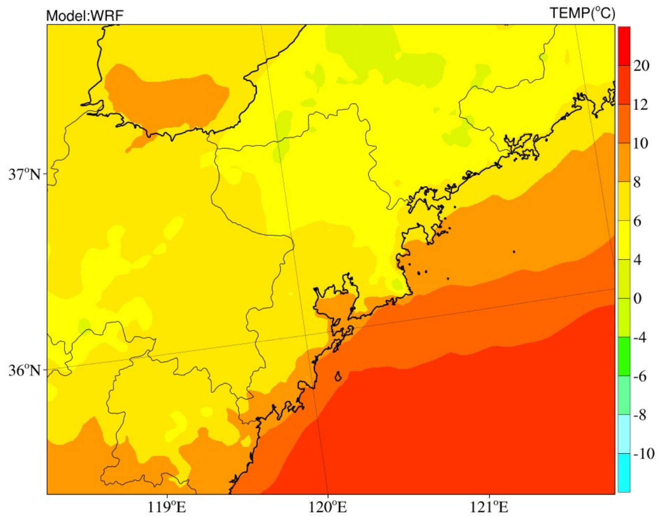

Fig. B1. The TEMP forecast result for 1 hour.

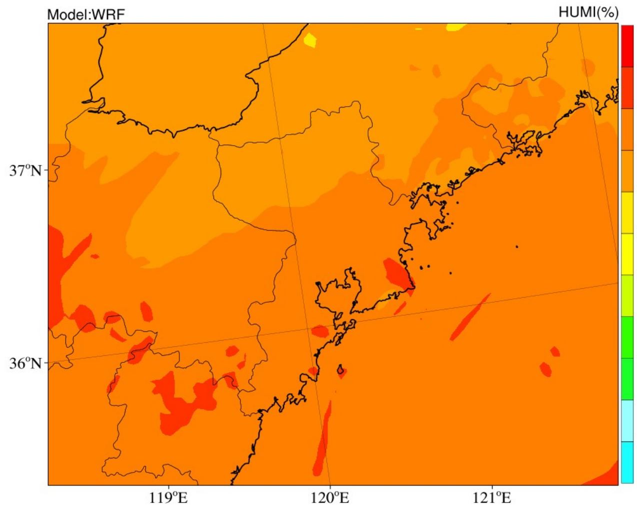

Fig. B2. The HUMI forecast result for 1 hour.

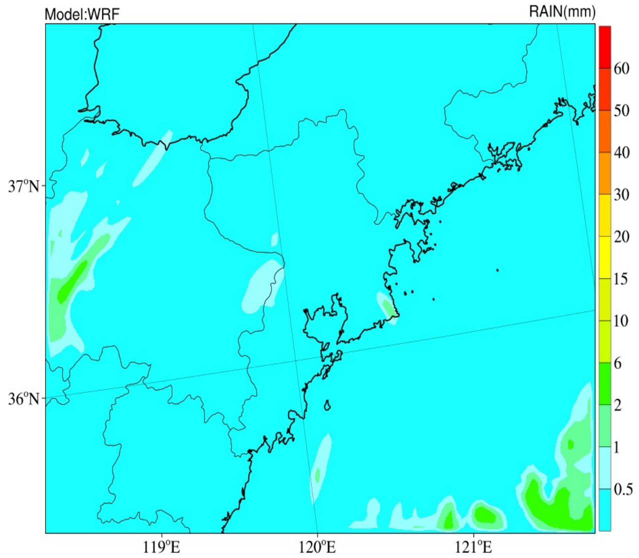

Fig. B3. The RAIN forecast result for 1 hour.

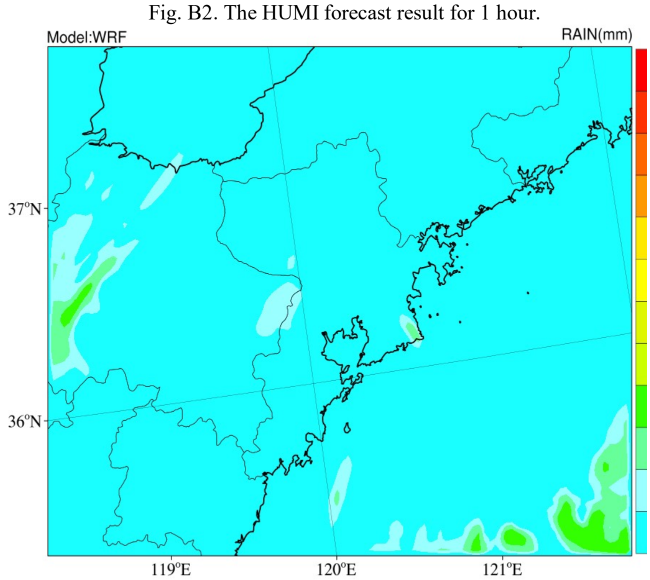

Fig. B4 The WIND.forecast result for 1 hour

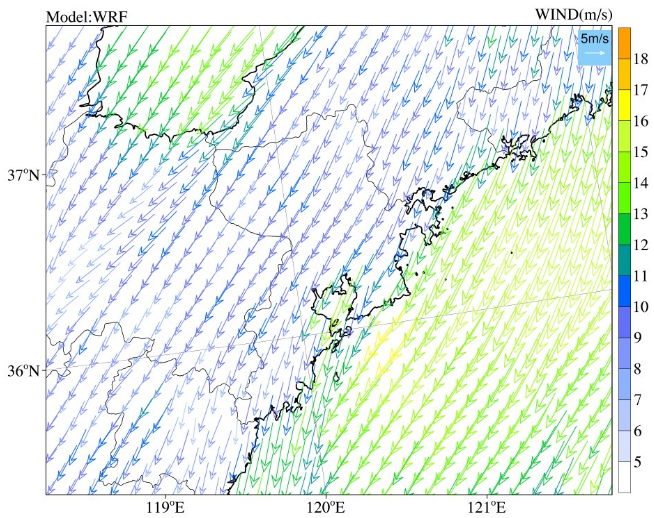

Fig. B5 The WIND.forecast result for 1 hour.

APPENDIX C

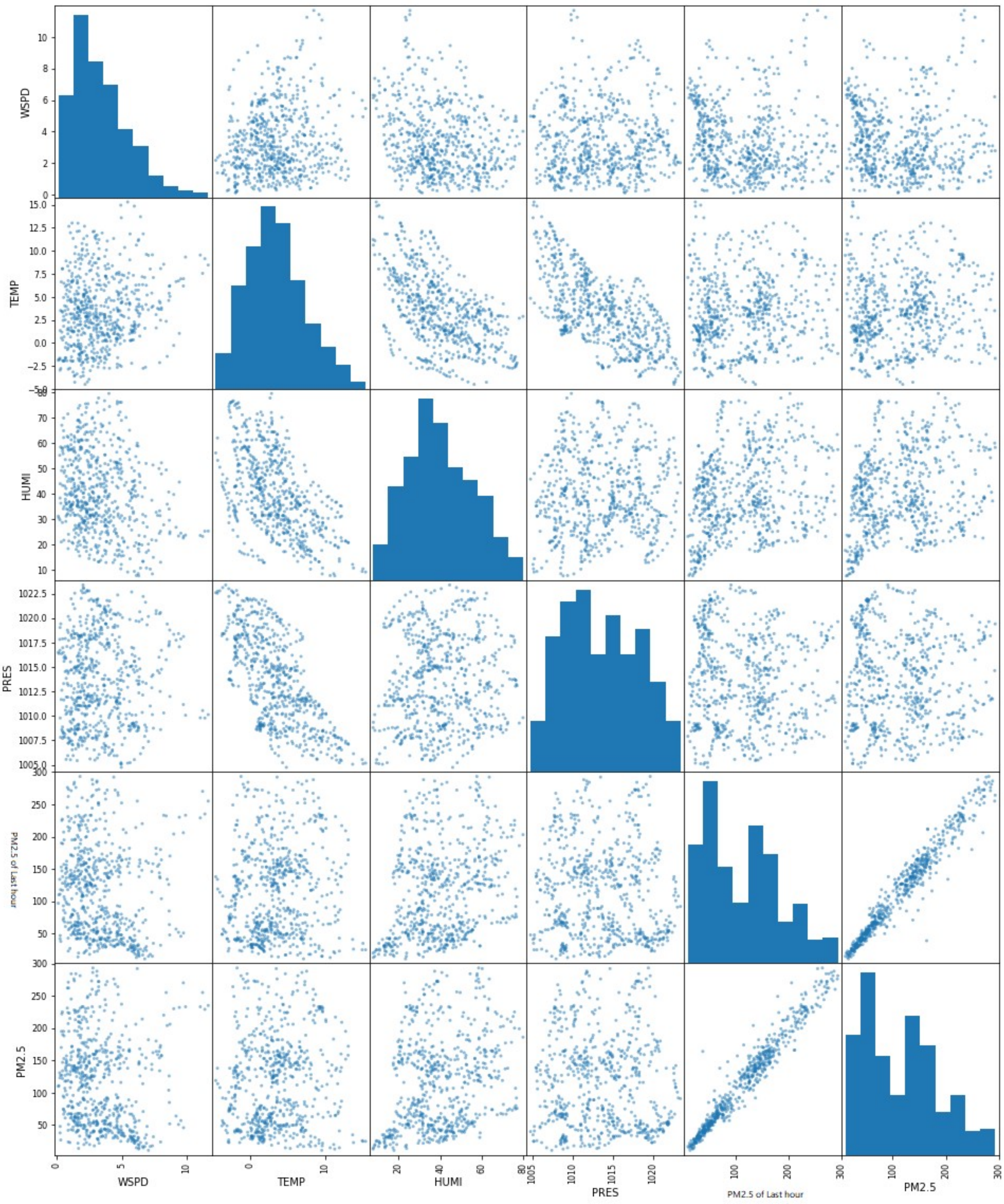

Fig. C1 Correlation analysis of various factors.