ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

Based Mixed Mathematical Model of Contaminant Related Meteorological Factors Used In Indicator System

Yang Fan[1, a], Kaiyu Xie[1, b]

[1] Department of Computer, North China Electric Power University, Baoding 071000, China;

[a]fywithnv@163.com, [b]Verbinsiki@163.com

Abstract

This paper based on the pollutant concentration and meteorological data of the central and southern cities of Hebei Province, this paper analyzes the main meteorological indicators affecting the concentration of various pollutants, and proposes a multi-index screening method for the influence factors of pollutant concentration. First, combined with a variety of data visualization methods, it is found that the relationship between the same pollutant and its meteorological impact factors is different in different cities or different seasons. Subsequently, by plotting the linear correlation matrix heat maps of six pollutants and 13 meteorological indicators， it is found that the same meteorological impact factor has different effects on different pollutant concentrations. Huge difference. Moreover, the intrinsic function relationship between different pollutant concentrations and various meteorological factors is also complex and variable. Based on the above analysis, considering the comprehensive use of Pearson correlation coefficient, Kendall correlation coefficient and maximum information coefficient MIC, which have different correlations between measurement variables, an automatic and efficient multiindicator screening method for pollutant concentration meteorological impact factors is proposed.Finally, taking Shijiazhuang and Xingtai as examples, the screening results of major meteorological impact factors corresponding to 48 combinations of two different cities, four different seasons and six different pollutants were given.

Keywords

Kendall; MIC; Multi-indicator screening; Visualization.

1. Introduction

With the vigorous development of industry, transportation and construction industry, the problem of air pollution dominated by carbon dioxide, nitrogen oxides and particulate matter has become increasingly serious, and it has become a serious problem faced by the Chinese government and society [1]. The rapid economic development in the Beijing-Tianjin-Hebei region has made its environmental problems increasingly serious, and the problem of air pollution is particularly acute. Hebei Province is one of the important cities in the Beijing-Tianjin-Hebei economic circle. At the same time, the central and southern regions are also one of the most polluted areas in the country, and their air pollution characteristics are representative. In the past 10 years, with the advancement of numerical forecasting technology, the perfection and enrichment of detection methods, and the rapid development and application of high-performance computers, modern weather forecasting technology has made remarkable progress, including rapid update assimilation analysis and forecasting, ensemble forecasting, The application of new technologies such as probabilistic forecasting and digital forecasting has promoted the improvement of China's weather forecasting

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

business [2]. The continuous development of industrialization and urbanization has led to an increase in the economic level. At the same time, environmental pollution has become increasingly serious. The concentration of fine particles, nitrogen oxides and sulfides in the air is getting higher and higher, and smog is frequently caused in many cities across the country. This has caused great troubles in our daily life, and it has also aroused strong concern from the society.

Literature [3] analyzed the air pollution phenomenon in 42 cities in China in the past ten years. Through the analysis of pollution factors, pollution index and air quality level, the results show that the source of air pollution is mainly due to the large amount of coal combustion in China. The increase of sulfide, particulate matter and API in the air, and seasonal changes will also affect the air quality; the literature [4] analyzes the PM10 pollution index data of Rizhao City for one year, and the results show that the pollutant concentration is obvious. The seasonal changes are also influenced by customs; the literature [5] utilizes the hourly monitoring data of PM2.5 and PM10, the joint analysis of satellite and radar monitoring data, and the continuous wording of Shanghai in March-September 2013. The cause of pollution in 6d is explained: the continuous pollution of air is closely related to the high and low spatial configuration of the weather situation; the literature [6] discusses the relationship between static weather and air pollution in Beijing, for example, by adopting different meteorological elements. The threshold and weight of the static and stable weather index are improved to better establish the pollutant monitoring model; the literature [7] uses the WRF-Chem model to simulate 2 In October 014, a process of heavy pollution of PM2.5 in the Beijing-Tianjin-Hebei region showed that the distribution of pollutants was significantly positively correlated with the height of the boundary; the literature [8] used ECMWF numerical forecast products and dynamic statistics methods for Beijing-Tianjin-Hebei The quantitative analysis of smog and air pollution in 14 cities shows that this scheme can predict the development of smog pollution in the area 5 to 6 hours ahead of time, and forecast and predict the development of pollution in advance; the literature [9] proposes Statistical correlation and K-means' new hybrid gene selection algorithm to achieve effective differentiation of gene subset selection; the literature [10] established a new nonlinear artificial intelligence set combined with genetic algorithm and neural network prediction model, with strong self Adapting to learning ability, the forecast results are accurate and stable; the literature [11] combined with KNN data mining algorithm to construct the haze level forecasting classifier. At present, environmental meteorology has become a hot topic as a new subject. Based on reading relevant literatures, this paper will select meteorological factors with good correlation with pollutants, and construct a prediction model for the difference between two adjacent days of air pollutants. At the same time, the mathematical model of pollutants is established according to the seasonal differences, and the optimal hybrid model is selected to provide a reference for the government to determine reasonable air quality forecasting models and policies.

This paper will take the air quality monitoring data sets of the central and southern cities of Hebei Province in Shijiazhuang and Xingtai as examples, and combine various visualization methods and multiple variable correlation analysis methods to analyze the main meteorological indicators affecting the concentration of various pollutants. A screening method for the influence factor of pollutant concentration is proposed. The research and discussion of this subject has important theoretical and practical significance for the prediction and prevention of pollution in the central and southern parts of Hebei Province.

2. Properties

According to the daily concentration data of 2014-2016 and the meteorological data of 2014-2016 in a certain city of Hebei Province, the data set is analyzed and analyzed first, and then the meteorological factors of different pollutants in different cities and different seasons are extracted. Matlab is used. By programming to visualize the results, we use the fixed variable method to study the correlation between the various factors one by one, and obtain the following diagram:

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

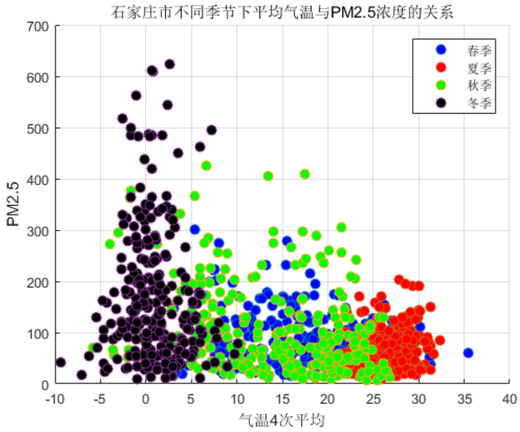

Fig.1 Relationship between average temperature and PM2.5 concentration in different seasons in Shijiazhuang

Figure 1 shows that there is a significant relationship between meteorological factors and pollutant concentrations in Shijiazhuang City under different seasons. According to theoretical analysis, the reason why the low temperature PM2.5 value is significantly lower than the higher temperature summer is that when the near-surface temperature is high, the atmospheric convection is intensified, which greatly reduces the PM2.5 concentration; When the inversion layer appears in the atmosphere, PM2.5 is not easy to diffuse, resulting in an increase in PM2.5 concentration; while the temperature in spring and autumn is relatively comfortable and the temperature is similar, the value of PM2.5 is relatively close and between summer and winter.

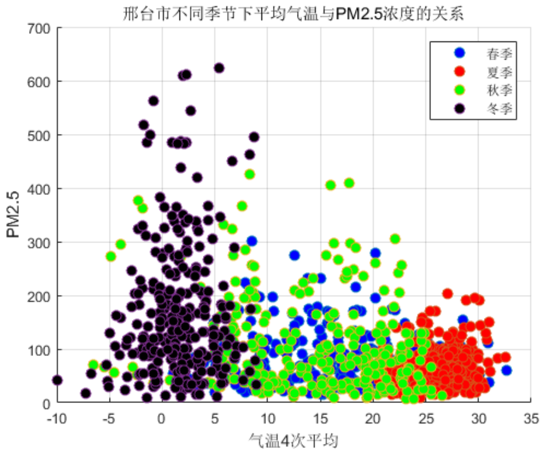

Figure 2 Relationship between average temperature and PM2.5 concentration in different seasons in Xingtai City

Comparing Fig. 2 with Fig. 1, it can be seen that the scatter plots of the average temperature and PM2.5 concentration in different seasons in Shijiazhuang City and Xingtai City are roughly the same as the whole, and there are some differences in local points. After analysis, it may be because the distance between Shijiazhuang City and Xingtai City is relatively close, and the temperature and other meteorological conditions are relatively close, so the PM2.5 concentration is similar in general.

Comparing Figure 3 with Figure 2, it can be seen that in the same city, the relationship between different meteorological factors is also different in different seasons. It can be seen that the influence of seasonal factors on meteorology is still more important.

International Core Journal of Engineering Volume 5 Issue 10, 2019
ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

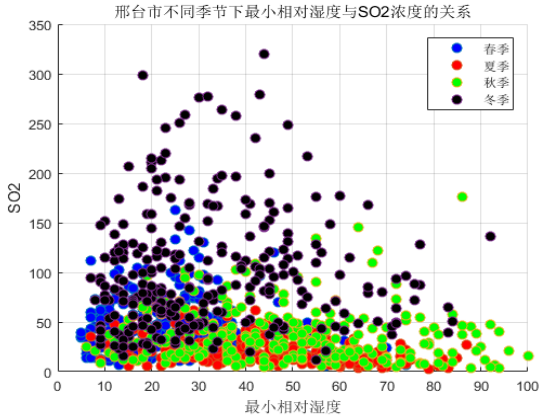

Fig. 3 Relationship between minimum relative humidity and SO2 concentration in different seasons in Xingtai City

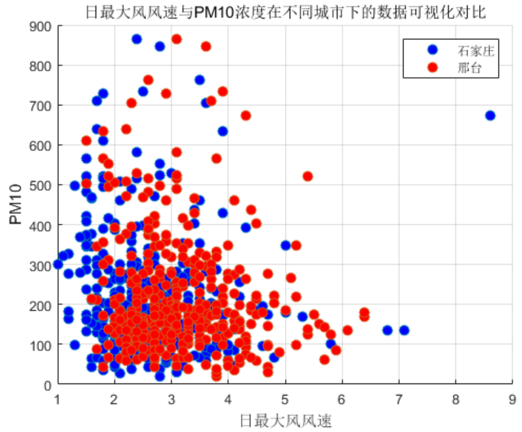

Figure 4 Comparison of data visualization of daily maximum wind speed and PM10 concentration in different cities

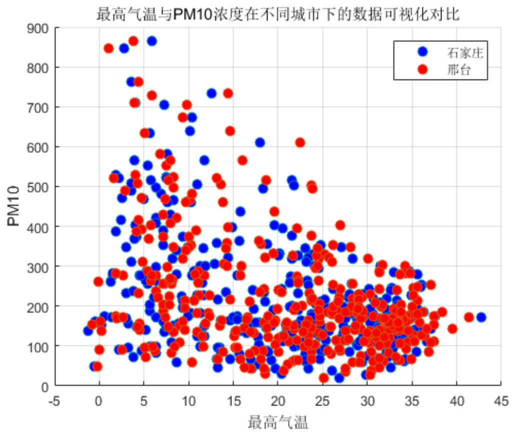

Figure 5 Comparison of data visualization of maximum temperature and PM10 concentration in different cities

Figure 1, Figure 2, and Figure 3 show that the impact of the season on air quality is more important. If we do not consider the seasonal impact, from the perspective of different cities, we can study the interaction relationship of meteorological factors in different cities. 4 and Figure 5. It can be obtained that even under the same meteorological factors, the relationship between meteorological factors in different cities will be different.

According to the relevant research of meteorological theory, we selected 11 indicators that may affect air quality. This index is divided into two aspects, one is the concentration of pollutants: including

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

suspended fine particles TSP, sulfur dioxide, nitrogen oxides, carbon monoxide, photochemical oxides， VOCs, greenhouse gases; on the other hand, meteorological conditions: including temperature, wind, rain, snow, seasons. We will first analyze each indicator intuitively, and finally use statistical methods to screen. Meteorological data of 7 cities in central and southern Hebei Province (including Shijiazhuang, Xingtai, Handan, Qinhuangdao, Langfang, Tangshan, Baoding) (including temperature average, maximum temperature, minimum temperature, 20-20 precipitation, local pressure 4 Sub-average, relative humidity 4 times average, minimum relative humidity, 10 minutes wind speed 4 times average, daily maximum wind speed, total sunshine hours), and mixed layer height, ventilation coefficient, static stability index, 6 in Shijiazhuang and Xingtai Contaminant concentration data. Since only Shijiazhuang and Xingtai provide pollutant concentration data, the research and prediction models in this paper are based on Shijiazhuang City and Xingtai City, and can be easily extended and applied to the analysis of other cities, including meteorological effects of pollutant concentrations. Screening of factors, prediction of pollutant concentrations, etc.

3. Establishment and Analysis of Indicator System and Mixed Mathematical Model of Contaminant Related Meteorological Factors

3.1 Establishment and solution of correlation model between correlation coefficient of pollutant concentration and various meteorological factors

In order to seek the relationship between meteorological factors and pollutants, we selected a typical meteorological factor to compare the concentration of pollutants. The experimental data of the comparison results are as follows:

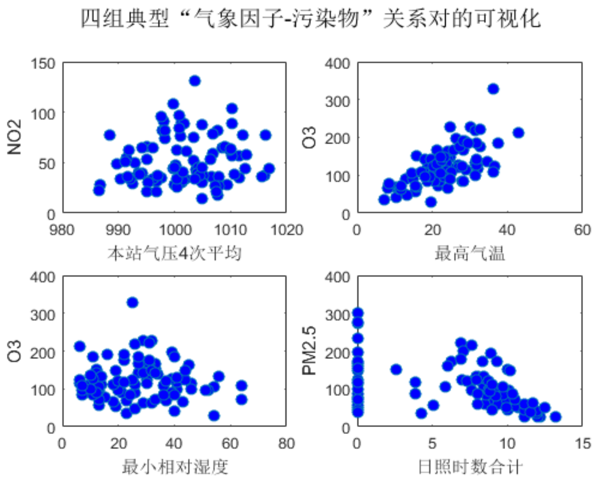

Figure 6 Comparison of the highest temperature and PM10 concentration in different cities

Figure 6 reflects the relationship between meteorological factors and pollutants. For example, the relationship between maximum temperature and ozone is generally positively correlated. Generally speaking, the higher the temperature, the higher the ozone concentration; When the relative humidity is related to ozone, it can be seen that the correlation appears as a negative correlation state; however, when considering the relationship between the total sunshine hours and PM2.5, it is found that there is a series of points and Focusing on the two correlations between 5-15, the reason may be that the weather conditions are cloudy, causing the number of sunshine to be 0, while in normal weather conditions, the average number of sunshine is around 10 hours, which will be presented as one. A negative correlation state. Based on the above considerations, we will measure the interrelationship between different pollutant factors in different cities and seasons from different angles and indicators.

3.1.1 Pearson

Euclidean Metric can find the distance between two vectors, ranging from 0 to positive infinity. However, since the Euclidean distance is given the same weight for each dimension by default in the

ISSN: 2414-1895
DOI: 10.6919/ICJE.201909_5(10).0043

calculation, the Pearson model with a range of values ranging from -1 to 1 is used to obtain the correlation trend.

$$\rho(x,y)=\frac{E[(X\text{-}\mu_x)(Y\text{-}\mu_y)]}{\sigma_x\sigma_y}=\frac{E[(X\text{-}\mu_x)(Y\text{-}\mu_y)]}{\sqrt{\sum_{i=1}^{n}(X_i\text{-}\mu_x)^2}\sqrt{\sum_{i=1}^{n}(Y_i\text{-}\mu_y)^2}}$$

There are two main steps in the calculation of the Pearson correlation coefficient: first, the correlation coefficient, followed by the significance level:

The calculation formula is

$$\gamma=\frac{\sum_{i=1}^{n}(x_i-\bar{x})\sum_{i=1}^{n}(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2(y_i-\bar{y})^2}}$$
$$=\frac{n\sum_{i=1}^{n}x_iy_i-(\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}y_i)}{\sqrt{n\sum_{i=1}^{n}x_i{}^2-(\sum_{i=1}^{n}x_i)^2}\sqrt{n\sum_{i=1}^{n}y_i{}^2-(\sum_{i=1}^{n}y_i)^2}}$$

In the calculation, the hypothesis H0 is first proposed: the correlation coefficient ρ=0 of the population A; H1: the correlation coefficient ρ≠0 of the population A, then the statistic of the test is calculated:

$$T=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}\sim t(n-2)$$

Finally, the level of significance can be determined based on the look-up table.Using the Pearson correlation coefficient, we modeled and programmed the summer pollutant concentrations and meteorological factors in the dataset in Xingtai City, and obtained the thermal correlation curves of linear correlation coefficients of different meteorological factors. Figure 7:

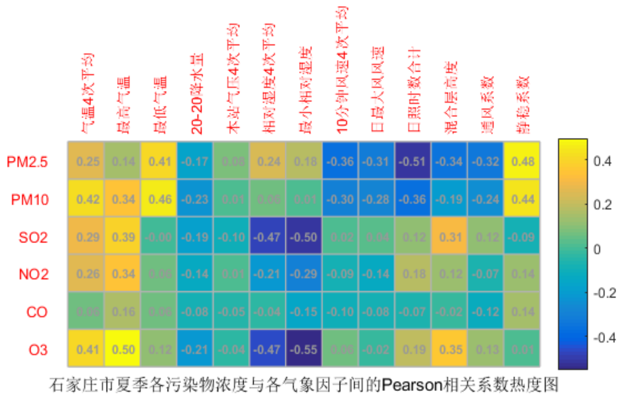

Fig.7 Heat map of correlation coefficient between pollutant concentration and meteorological factors in Shijiazhuang City in summer

Analysis of Figure 2-7 can obtain the relationship between various meteorological factors in Xingtai City in summer. For example, the linear correlation coefficient between SO2 and relative humidity is -0.47. This negative correlation is due to the chemical nature of SO2: It is the main component of acid rain. Usually, sulfur dioxide is further oxidized in the presence of a catalyst such as nitrogen dioxide. The formation of sulfuric acid; considering the relationship between PM2.5 and various meteorological factors, we can see that the linear correlation coefficient between PM2.5 and temperature is 0.41. As we discussed earlier, the temperature has a greater influence on the concentration of particulate matter. When the temperature is high, the air flow is slow and the particles are difficult to diffuse.

The Pearson correlation coefficient is applicable if the two variables conform to the bivariate normal distribution. If the data set has extreme values, the result will have a greater impact.

3.1.2 Spearman

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

Using the Spearman model, we modeled and programmed the concentration of various pollutants in summer and the meteorological factors in Xingtai City, and obtained the linear correlation coefficient of different meteorological factors.

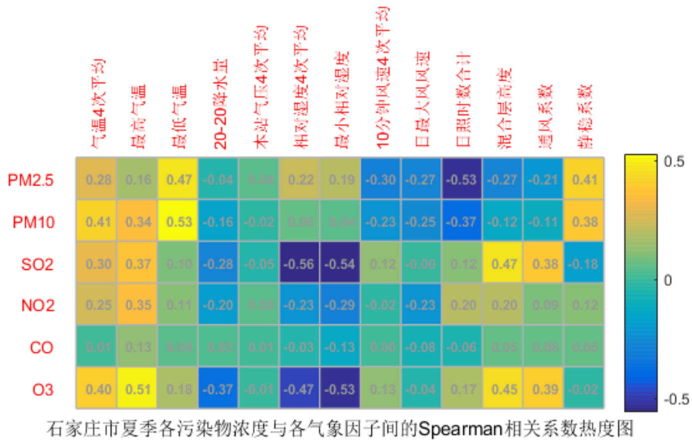

Fig.8 Correlation coefficient between the concentration of various pollutants in summer and the meteorological factors in Shijiazhuang City

By comparing Fig. 8 with Fig. 7, it can be seen that the results of the first linear analysis by the Spearman model and the Pearson model are similar, and there are only small differences in the numerical values. After analysis and discussion, the results obtained by the Spearman model are more It is in line with the empirical model, so when the hybrid model is finally designed, the Spearman model is used for the design.

3.1.3 MIC

Maximal Information Coefficient (MIC) 属 于 Maximal Information-based Nonparametric Exploration (MINE) the largest information-based nonparametric exploration used to measure the linear or nonlinear strength of two variables

Using the MIC model, we modeled and programmed the summer pollutant concentrations and meteorological factors in Xingtai City, and obtained the linear correlation coefficient of different meteorological factors.

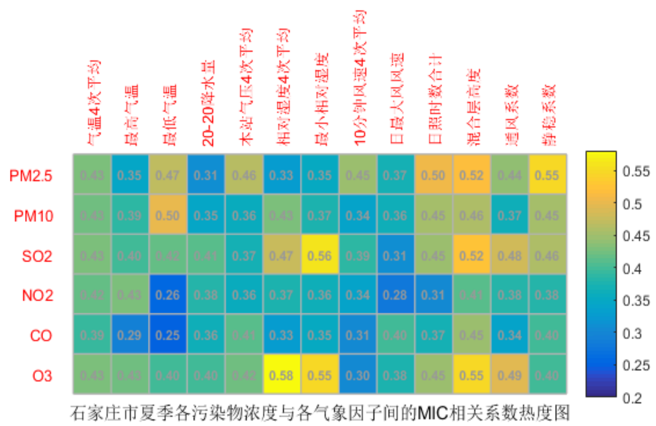

Fig.9 Correlation coefficient between the concentration of various pollutants in summer and the meteorological factors in Shijiazhuang City

By analyzing Fig. 9, it can be seen that the MIC model has a coefficient ranging from 0 to 1 compared to the Spearman model and the Pearson model. When the coefficient is 0, it is irrelevant. When 1 is used, it is related. The concept of negative correlation is only considered from the relevant magnitude. The MIC model can identify any functional relationship, but it needs to be improved in efficiency compared to the first two models.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

3.1.4 Kendall correlation coefficient

The correlation coefficient of the Kendall model ranges from -1 to 1, when τ is 1, it means that two random variables have a consistent rank correlation; when τ is -1, it means that two random variables have completely opposite rank correlations. When τ is 0, it means that the two random variables are independent of each other

Using the Kendall model, we modeled and programmed the summer pollutant concentrations and meteorological factors in Xingtai City, and obtained the linear correlation coefficient of different meteorological factors. The thermal analysis visualization relationship is shown in Figure 10:

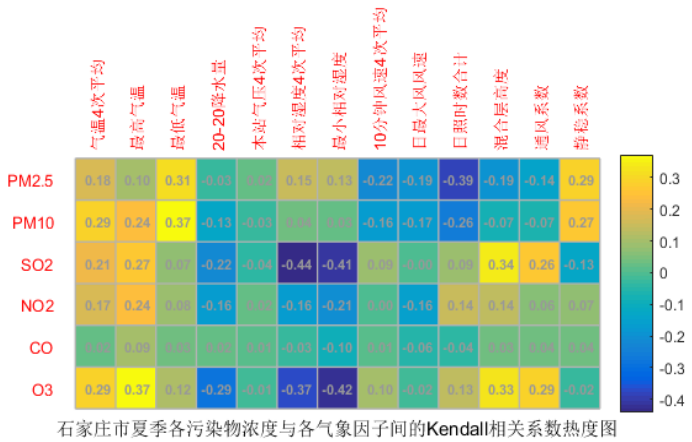

Figure 10 Heat transfer diagram of Kendall correlation coefficient between various pollution concentrations and meteorological factors in Shijiazhuang in summer

It can be seen from Fig. 10 that the difference between the Kendall model and the Spearman model is that the data is faster and more efficient than the Spearman model when the data needs to be ordered.

3.2 Establishment and Solution of Mixed Mathematical Model of Contaminant Related Meteorological Factors

Based on the above analysis, considering the three correlation coefficients of Pearson, Kendall and MIC, the final mixed correlation coefficient Rf is defined as follows:

$$R_f = \begin{cases} \widehat{R_f}, & \widehat{R_f} > \beta \\ 0, & \widehat{R_f} < \beta \end{cases}$$

In order to avoid making the variable screening too radical, this article takes the final data from Shijiazhuang in summer, the screening results of the six pollutants are shown in Table 1, and the visual display is shown in Figure 11:

Table 1 Screening results of various pollution concentrations and meteorological factors in Shijiazhuang in summer

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.4262	0.4261	0.4306	0.4191	0.3942	0.4262
Maximum temperature	0.3455	0.3899	0.3993	0.4281	0.2929	0.4986
Minimum temperature	0.4690	0.5005	0.4191	0.2558	0.2535	0.3996
20-20 precipitation	0.3125	0.3463	0.4118	0.3797	0.3645	0.3998
The site pressure 4 times average	0.4628	0.3581	0.3687	0.3561	0.4130	0.4211

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

Relative humidity 4 times average	0.3280	0.4263	0.4741	0.3749	0.3288	0.5810
Minimum relative humidity	0.3491	0.3748	0.5619	0.3636	0.3498	0.5501
10 minutes wind speed 4 times average	0.4494	0.3375	0.3868	0.3417	0.3060	0.3044
Maximum wind speed	0.3722	0.3640	0.3083	0.2775	0.4021	0.3760
Total sunshine hours	0.5113	0.4523	0.4456	0.3064	0.3710	0.4458
Mixed layer height	0.5162	0.4609	0.5245	0.4111	0.4458	0.5465
Ventilation coefficient	0.4388	0.3653	0.4800	0.3844	0.3415	0.4933
Static stability coefficient	0.5457	0.4535	0.4602	0.3841	0.3981	0.4012

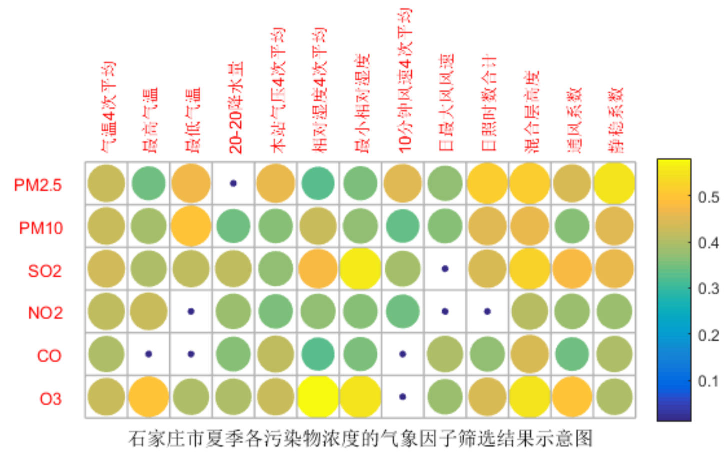

Figure 11 Screening results of various pollution concentrations and meteorological factors in Shijiazhuang in summer

By analyzing Figure 11, we can clearly see the correlation between meteorological factors and pollutant concentrations, the depth of the color and the intensity associated with the area of the circle. The screening data for spring, summer, autumn and winter in Shijiazhuang in spring, autumn, winter and Xingtai are shown in the following:

Table 2 Screening results of various pollution concentrations and meteorological factors in Shijiazhuang in spring

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.3735	0.4070	0.4271	0.3605	0.3899	0.8252
Maximum temperature	0.4024	0.3833	0.4893	0.3658	0.4465	0.7708
Minimum temperature	0.3809	0.3984	0.4543	0.4468	0.4036	0.7462
20-20 precipitation	0.3220	0.3772	0.3717	0.3607	0.3345	0.3064
The site pressure 4 times average	0.3480	0.3076	0.4107	0.3479	0.4052	0.5607
Relative humidity 4 times average	0.3725	0.4088	0.4005	0.4832	0.4355	0.3404
Minimum relative humidity	0.4208	0.3794	0.4781	0.4430	0.4305	0.3680
10 minutes wind speed 4 times average	0.4382	0.3562	0.5205	0.3841	0.4340	0.5155

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

Maximum wind speed	0.4419	0.3662	0.5255	0.3937	0.4279	0.6020
Total sunshine hours	0.4448	0.3120	0.4520	0.3820	0.3601	0.3563
Mixed layer height	0.4597	0.3528	0.4675	0.3574	0.4342	0.3970
Ventilation coefficient	0.4903	0.3376	0.4534	0.4064	0.3946	0.5263
Static stability coefficient	0.6297	0.5072	0.5458	0.5092	0.5602	0.3781

Table 3 Screening results of various pollution concentrations and meteorological factors in Shijiazhuang in autumn

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.5113	0.4311	0.4077	0.5017	0.4822	0.8325
Maximum temperature	0.5112	0.4563	0.5057	0.5033	0.4822	0.8325
Minimum temperature	0.4874	0.4175	0.4160	0.4497	0.4588	0.7751
20-20 precipitation	0.3632	0.4992	0.5216	0.6435	0.4023	0.4656
The site pressure 4 times average	0.4585	0.3922	0.3301	0.3608	0.3798	0.5908
Relative humidity 4 times average	0.3576	0.4023	0.3860	0.4021	0.4182	0.7560
Minimum relative humidity	0.3574	0.3853	0.3779	0.3961	0.4317	0.7265
10 minutes wind speed 4 times average	0.4303	0.3520	0.3642	0.3695	0.4440	0.5370
Maximum wind speed	0.4862	0.4061	0.4613	0.4142	0.4525	0.3720
Total sunshine hours	0.4344	0.3870	0.4195	0.3396	0.4131	0.6792
Mixed layer height	0.4287	0.4129	0.4238	0.4326	0.4983	0.6509
Ventilation coefficient	0.4854	0.4588	0.4680	0.3613	0.4551	0.4825
Static stability coefficient	0.6937	0.5967	0.5631	0.5014	0.6607	0.5712

Table 4 Screening results of various pollution concentrations and meteorological factors in Shijiazhuang in winter

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.3682	0.3794	0.3630	0.3522	0.3841	0.5581
Maximum temperature	0.4046	0.4296	0.3210	0.3202	0.3940	0.5461
Minimum temperature	0.3358	0.3094	0.3754	0.3107	0.3290	0.3668
20-20 precipitation	0.3333	0.3354	0.3181	0.3276	0.3460	0.3109
The site pressure 4 times average	0.4777	0.3797	0.3869	0.3572	0.3700	0.4084
Relative humidity 4 times average	0.6790	0.6033	0.4133	0.3600	0.5375	0.5418
Minimum relative humidity	0.6016	0.5496	0.4232	0.4204	0.5472	0.6268
10 minutes wind speed 4 times average	0.5044	0.5010	0.4431	0.4390	0.4964	0.3731
Maximum wind speed	0.3447	0.3851	0.4733	0.5301	0.3801	0.4071

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

Total sunshine hours	0.5762	0.5178	0.4267	0.4410	0.4970	0.6294
Mixed layer height	0.7060	0.6731	0.4822	0.4942	0.5939	0.6838
Ventilation coefficient	0.6769	0.6917	0.5766	0.6260	0.6012	0.5367
Static stability coefficient	0.3782	0.4190	0.3907	0.3753	0.3622	0.6036

Table 5 Screening results of various pollution concentrations and meteorological factors in spring in Xingtai City

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.3779	0.3720	0.4808	0.3643	0.4051	0.8159
Maximum temperature	0.3652	0.4791	0.4244	0.3541	0.4031	0.7963
Minimum temperature	0.4159	0.3267	0.4382	0.3786	0.4405	0.7451
20-20 precipitation	0.3139	0.3356	0.3489	0.2687	0.3870	0.3115
The site pressure 4 times average	0.2817	0.3047	0.4112	0.3889	0.3772	0.5530
Relative humidity 4 times average	0.3718	0.4380	0.3999	0.5018	0.4396	0.4307
Minimum relative humidity	0.4197	0.4056	0.4320	0.4138	0.3946	0.2996
10 minutes wind speed 4 times average	0.2541	0.2952	0.3316	0.2904	0.3293	0.3818
Maximum wind speed	0.3366	0.3282	0.3362	0.4131	0.2640	0.3408
Total sunshine hours	0.4198	0.2677	0.4021	0.3427	0.3001	0.4092
Mixed layer height	0.3653	0.3389	0.3871	0.3818	0.3905	0.3283
Ventilation coefficient	0.3458	0.3189	0.3774	0.3639	0.3827	0.3741
Static stability coefficient	0.5157	0.4142	0.6167	0.5535	0.5889	0.3202

Table 6 Screening results of various pollution concentrations and meteorological factors in summer in Xingtai City

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.3902	0.4286	0.4211	0.3037	0.2571	0.4082
Maximum temperature	0.4201	0.3685	0.4053	0.3829	0.3841	0.4466
Minimum temperature	0.5667	0.5670	0.4100	0.2807	0.2852	0.3327
20-20 precipitation	0.3787	0.3967	0.3984	0.3301	0.3798	0.3823
The site pressure 4 times average	0.4010	0.3657	0.4190	0.4034	0.4080	0.3607
Relative humidity 4 times average	0.3660	0.4239	0.4916	0.3526	0.3551	0.5148
Minimum relative humidity	0.4603	0.4267	0.4998	0.3893	0.3566	0.4684
10 minutes wind speed 4 times average	0.4210	0.3964	0.3587	0.3241	0.4056	0.3328
Maximum wind speed	0.3310	0.3205	0.3110	0.2801	0.3412	0.2920

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

Total sunshine hours	0.3898	0.3464	0.3485	0.3067	0.3054	0.3851
Mixed layer height	0.3852	0.3556	0.4058	0.4499	0.3708	0.4950
Ventilation coefficient	0.4176	0.3373	0.3968	0.3270	0.3507	0.3901
Static stability coefficient	0.4311	0.4185	0.3670	0.3730	0.3602	0.3941

Table 7 Screening results of various pollution concentrations and meteorological factors in autumn in Xingtai City

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.5124	0.4169	0.4477	0.4619	0.4813	0.7732
Maximum temperature	0.5685	0.4289	0.3842	0.4218	0.4754	0.7770
Minimum temperature	0.5112	0.4458	0.4594	0.4600	0.4741	0.7751
20-20 precipitation	0.4107	0.4345	0.4719	0.5270	0.3175	0.4507
The site pressure 4 times average	0.4838	0.4025	0.3544	0.3740	0.3858	0.5933
Relative humidity 4 times average	0.3883	0.3761	0.3374	0.4332	0.4274	0.7484
Minimum relative humidity	0.3841	0.3824	0.4028	0.4381	0.3801	0.7927
10 minutes wind speed 4 times average	0.4006	0.3342	0.3426	0.3785	0.3614	0.4176
Maximum wind speed	0.4985	0.4815	0.4811	0.4756	0.3771	0.4894
Total sunshine hours	0.4186	0.3755	0.3220	0.3786	0.4650	0.6266
Mixed layer height	0.4370	0.3969	0.3697	0.3485	0.4353	0.4797
Ventilation coefficient	0.4368	0.4386	0.4423	0.4030	0.3775	0.4654
Static stability coefficient	0.6853	0.6493	0.7843	0.5960	0.5837	0.4955

Table 8 Screening results of various pollution concentrations and meteorological factors in winter in Xingtai City

	PM2.5	PM10	SO2	NO2	CO	O3
Temperature average four times	0.3298	0.3340	0.3519	0.4162	0.4356	0.4867
Maximum temperature	0.3278	0.3320	0.2782	0.3940	0.3480	0.4847
Minimum temperature	0.2970	0.3484	0.3889	0.3882	0.3509	0.3640
20-20 precipitation	0.3534	0.3534	0.3534	0.2996	0.3132	0.2862
The site pressure 4 times average	0.4777	0.3797	0.3404	0.3332	0.3700	0.4084
Relative humidity 4 times average	0.5578	0.4927	0.3963	0.3472	0.4331	0.5051
Minimum relative humidity	0.6314	0.5491	0.4928	0.4078	0.5318	0.5647
10 minutes wind speed 4 times average	0.3738	0.4710	0.4345	0.3353	0.3930	0.4158
Maximum wind speed	0.4256	0.3822	0.3875	0.4078	0.4120	0.4750

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0043

Total sunshine hours	0.5080	0.5021	0.4343	0.4260	0.4466	0.5237
Mixed layer height	0.5081	0.4884	0.3875	0.4344	0.4133	0.5506
Ventilation coefficient	0.4816	0.4657	0.3793	0.4348	0.4311	0.5434
Static stability coefficient	0.4144	0.3943	0.4183	0.4201	0.4348	0.4948

4. Conclusion

This paper proposes an automatic and efficient screening method for multiple indicators of meteorological influence factors of pollutant concentration. Combined with a variety of data visualization methods, it is analyzed that the relationship between the same pollutant and its meteorological impact factors is different in different cities or different seasons. Subsequently, by plotting the linear correlation matrix heat maps of six pollutants (PM2.5, PM10, etc.) and 13 meteorological indicators (highest temperature, minimum temperature, etc.), it is found that the same meteorological impact factor has different effects on different pollutant concentrations. Large differences, using Pearson correlation coefficient, Kendall correlation coefficient and MIC correlation coefficient, which are three different measures of correlation between measurement variables, propose an automatic and efficient multi-index screening method for pollutant concentration meteorological impact factor. Finally, taking Shijiazhuang and Xingtai as examples, the screening results of major meteorological impact factors corresponding to 48 combinations of two different cities, four different seasons and six different pollutants were given.

References

[1] Hong Zhongxiang, Hu Fei, Research progress in theory and methods of air pollution prediction [J], Climate and Environmental Research, 4(3): 225-230, 1994.

[2] Jin Ronghua, Xu Zhifang, et al. Development and application of modern weather forecast in China in recent years, Journal of Meteorology, 72(6): 1069-1078, 2014.

[3] Li Xiaofei, Zhang Mingjun, Wang Shengjie, etc. Analysis of the characteristics and influencing factors of China's air pollution index, Environmental Science, 33(6): 1936-1943, 2012..

[4] Zheng Meiqin, Lu Zhenli, Characteristics of PM10 pollutants in Rizhao urban area and its relationship with meteorological elements [J], Journal of Nanjing Institute of Meteorology, 29(3): 413-417, 2006.

[5] Chen Lei, Ma Jinghui, Qi Fuhai, Analysis of a typical continuous particulate matter pollution process in Shanghai, Meteorology, 42(2): 203-212, 2016.

[6] Zhang Hengde, Zhang Bihui, Lu Mengyao, Preliminary construction of the static and stable weather comprehensive index in Beijing and its application in environmental meteorology, Meteorology, 43(8): 998-1004, 2017.

[7] He Xinhe, Ma Jianzhong, Xu Jing, Ma Zhiqiang, Xue Min, Qi Junli, Numerical simulation of a PM2.5 pollution process in the Beijing-Tianjin-Hebei region in October 2014, Meteorology, 42(7): 827-837, 2016.

[8] Zhang Ziyin, Zhao Xiujuan, Xiong Yajun, Ma Xiaohui, Mid-term forecasting experiment of Beijing-Tianjin-Hebei smog based on dynamic statistical forecasting method, Journal of Applied Meteorology, 29(1): 57-69, 2018.

[9] Xie Juanying, Gao Hongchao, Distinguishing Gene Subset Selection Algorithm Based on Statistical Correlation and K-means, Journal of Software, 25(9): 2050-2075, 2014.

[10] Huang Xiaoyan, Zhao Huasheng, Huang Ying, Lin Kaiping, He Li, Genetics _ Neural network ensemble prediction method in the tropical cyclone precipitation forecast in Guangxi, Journal of Natural Disasters, 26(6): 184-196, 2017.

[11] Xiong Yajun, Zhang Xiaoling, Ma Xiaohui, et al. Application of KNN data mining algorithm in the level forecasting in Beijing, Meteorology, 41(1): 98-104, 2015.