

Study on the Spatial and Temporal Distribution of Fog and Haze

Hongtao Lin

College of Economics, Jinan University, Guangzhou 510632, China

865629968@qq.com

Abstract

The spatial and temporal distribution of haze is one of the most urgent environmental problems in the world. This paper calculates the weight by subjective discriminant analysis, and uses the partial least squares of regression analysis, fuzzy subjective discriminant analysis, 0-1 integer programming model, air quality index analysis chart, Euclid proximity analysis, cubic exponential smoothing method to correlate AQI indicators, PM2.5 other five sub-indicators and their correlation with the corresponding pollutants (content), PM2.5 spatial and temporal distribution, PM2.5 spatial distribution, and its relationship. In this paper, the correlation and independence of six basic monitoring indicators are analyzed quantitatively, especially the correlation and relationship between them and other five sub-indicators and their corresponding pollutants. The correlation coefficient matrix is used to analyze the correlation between AQI indicators, PM2.5 other five sub-indicators and their corresponding pollutants. Based on the pretreatment of Annex I data, a partial least squares regression analysis model was established to analyze the relationship between the other five sub-indicators of PM2.5 and the corresponding pollutants (content), and the regression equation was obtained. The spatial and temporal distribution of PM2.5 in Hangzhou and its related laws is studied. Firstly, the spatial and temporal distribution of PM2.5 is analyzed by visual icon method. Secondly, combined with the fuzzy subjective discriminant analysis model, 0-1 integer programming, air quality index analysis map model, Euclid proximity analysis model, regional pollution assessment is conducted, and the most polluted areas are obtained by mutual checking between the models. Among them, through the model of fuzzy subjective evaluation and analysis, we can get the weight of pollutants and establish the grade of air quality index. Using the size of air quality subordination degree, we will evaluate the regional pollution and select Shaoxing as the most polluted area. Besides, through 0-1 integer programming, we establish the constraint equation and select the most polluted area as Shaoxing. What's more, through the analysis map of air quality index, we can select Shaoxing as the most polluted area. Based on the model, the analysis map of air quality index is constructed, and the most polluted area is Shaoxing through observation and calculation; and the most polluted area is Shaoxing through Euclid proximity analysis model. It is concluded that Shaoxing is the most polluted area in Zhejiang Province by checking each other among the models. According to the meteorological data in Hangzhou area, the mathematical model of occurrence and the evolution law is established by reasonably considering the influence of weather and seasonal factors such as wind and humidity. Firstly, considering the effect of time on PM2.5 concentration, the cubic exponential smoothing prediction model is established to analyze the effect of time on PM2.5 concentration. Then we consider the influence of wind speed, temperature and humidity on PM2.5 concentration, and compare the predicted value with the actual monitoring value, so as to quantitatively and qualitatively analyze the occurrence and evolution of PM2.5, and conclude that the wind speed, temperature and humidity in the external factors will affect the PM2.5 pollutant diffusion to a certain extent.

Keywords

PM2.5; Fuzzy subjective discriminant analysis; 0-1 integer programming; Euclid proximity analysis; Air Quality Index Analysis Chart.

1. Introduction

The atmosphere provides an ideal environment for the multiplication of life on the earth and the development of human beings. Its state and change have a direct impact on human production, life and survival. Air quality has always been a hot issue of concern to the government, environmental protection departments and the people of the whole country.

The Ministry of Environmental Protection has published the newly revised Environmental Air Quality Standard^[1]. The main content of this revision are as follows: adjusting the classification of the functional zones of ambient air, merging the three zones into the second zone; adding the concentration limits of particulate matter (particle size is less than or equal) and the average concentration limits of ozone for 8 hours; adjusting the concentration limits of particulate matter (particle size is less than or equal), nitrogen dioxide, lead and benzo (a) pyrene, etc. The validity of data statistics has been adjusted. In line with the new standards, the Technical Regulations for Environmental Air Quality Index (AQI) have been implemented^[2].

In the above regulations, the air quality index is used as the air quality monitoring index to replace the original air quality monitoring index - air pollution index. The original monitoring index is the dimensionless index, and its sub-monitoring index is three basic indicators (sulfur dioxide, nitrogen dioxide and inhalable particulate matter). It is also a dimensionless index. Its sub-monitoring indicators are six basic monitoring indicators (sulfur dioxide, nitrogen dioxide, inhalable particulate matter, fine particulate matter, ozone and carbon monoxide, etc.). In the new standard, for the first time, the concentration index of fine particulate matter, the main factor causing haze, which is harmful to human health, is used as the air quality monitoring index^[4]. The issuance and implementation of new monitoring standards will play an important role in monitoring air quality and improving living environment. We must make every effort use the existing data to carry out the research on the spatial and temporal distribution of haze.

The formation mechanism and process are relatively complex, mainly from natural sources (plant pollen and spores, soil dust, sea salt, forest fires, volcanic eruptions, etc.) and man-made sources (fuel combustion, industrial production process emissions, transportation emissions, etc.), which can be divided into primary particulate matter (i.e., particulate matter directly discharged into the atmosphere from emission sources) and secondary particulate matter (i.e., through and through). Particulate matter produced by chemical reactions of atmospheric components. The composition is mainly composed of water-soluble ions, particulate organic matter and trace elements. One study suggested that sulfur dioxide (SO₂), nitrogen dioxide (NO₂) and carbon monoxide (CO) in monitoring indicators was the main gaseous substances before they formed under certain environmental conditions.

Suppose we know the data of 11 monitoring points in a province, such as Zhejiang Province, how to establish an appropriate mathematical model to quantitatively analyze the correlation and the independence of the six basic monitoring indicators, especially the correlation between them (content) and other five sub-indicators and their corresponding pollutants (content)?

At the same time, how to describe the spatial and temporal distribution of Hangzhou and its related laws from the data of observation points, and how to explain which areas are most polluted in Zhejiang Province according to the analysis?

When we know the meteorological data of a certain area, we should reasonably consider the influence of weather and seasonal factors such as wind and humidity, and how to establish a mathematical model of occurrence and evolution law, and use the data of this area to carry out quantitative and qualitative analysis.

The traditional methods for haze monitoring point data are too simple, using a single algorithm or establishing a single model for analysis. In this paper, through the use of a variety of algorithms, and innovative establishment of new model algorithms such as air quality analysis chart model,

comprehensive comparative analysis to enhance the feasibility of the results, reduce the degree of error in the analysis.

2. Problem Analysis

2.1 Analysis of Problems

Aiming at the distribution, occurrence and evolution of air quality index and PM_{2.5} in haze area, this paper applies partial least squares regression analysis model, fuzzy subjective discriminant analysis model, 0-1 integer programming model, air quality index analysis graph model, Euclid closeness analysis model and cubic exponential smoothing model to quantitatively analyze and rationally apply data. The correlation of air quality indicators, temporal and spatial distribution, occurrence and evolution of PM_{2.5} were calculated by using statistical analysis software such as the R language.

Using the data of 11 monitoring points in Zhejiang Province, firstly, we use correlation coefficient matrix to quantitatively analyze the correlation and independence of six basic monitoring indicators in AQI, especially the correlation between PM_{2.5} (content) and other five sub-indicators and corresponding pollutants (content). Secondly, we use partial least squares regression analysis model to study PM_{2.5}. The relationship between (content) and other five sub-indicators and corresponding pollutants (content) was obtained, and the corresponding regression equation was obtained.

The spatial and temporal distribution of Hangzhou and its related laws is studied. In the first part, the spatial and temporal distribution of PM_{2.5} and its related laws are obtained by using R statistical software. In the second part, firstly, we will establish objective analysis model (0-1 integer programming, Euclid proximity analysis model) and subjective analysis model (fuzzy subjective discriminant evaluation analysis model, air quality analysis map model) for pollution assessment, combining the characteristics of each model. The most polluted areas are selected by mutual amendment, which reduces the deviation caused by objective and subjective factors.

The problem needs to combine the meteorological data of Hangzhou area, reasonably consider the influence of weather and seasonal factors such as wind, humidity and so on, establish the prediction model of three exponential smoothing method in time series to analyze the occurrence and evolution of PM_{2.5}, get the prediction curve of PM_{2.5}, and analyze which external factors have greater impact on PM_{2.5}.

3. Hypothesis of the model

1. Assuming that the data given by the topic are true and credible, without considering human factors, it has statistical, monitoring and predictive significance.
2. Assuming that no biological and chemical reactions take place in the atmosphere;
3. The change of air quality caused by unexpected events is not considered.
4. There are no other similar pollution sources in all monitoring ranges.
5. Assume the same level of air pollution.

4. Establishment and Solution of Model

Starting from the hypothesis of the required problem and on the basis of the general analysis of the problem, three problems of the problem are analyzed and solved in detail, so that model I is established for problem 1, model II, model III, model IV and model V are established for problem 2, and model VI is established for problem 3.

Model I: Partial Least Squares Regression Analysis Model

In the pretreatment of data, this model establishes partial least squares regression analysis model, uses multiple regression analysis method and R language software program (see appendix), obtains the correlation and independence of six basic monitoring indicators in AQI, and obtains the correlation and regression equation between PM_{2.5} and other five sub-indicators and their corresponding pollutants .

Model II: Fuzzy Subjective Discriminant Evaluation and Analysis Model

By using the principle and method of fuzzy mathematics, 11 cities in Zhejiang Province were evaluated comprehensively in order to obtain comprehensive, objective and reasonable evaluation results and select the most polluted cities.

Model III: 0-1 Integer Programming

By processing and optimizing the data, this model uses 0-1 integer programming to select five cities with serious air quality pollution, and then select the cities with the most serious air pollution from five cities.

Model IV: Analytical Chart Model of Air Quality Index

Firstly, according to the weight of pollutants determined by model II, the analysis value of air quality index is calculated, and the comprehensive index of air quality index of each city is obtained by calculating the area of air quality grade analysis chart.

Model V: Euclid Closeness Analysis Model

This model solves the membership degree of each item in each level of air quality standard and the corresponding fuzzy subset of each monitoring point, and uses Euclid closeness degree to analyze and evaluate which area of Zhejiang Province has the most serious pollution.

Model VI: Cubic exponential smoothing model

In this model, the influence of time on PM2.5 concentration is analyzed by cubic exponential smoothing method, and the predicted values are compared with the actual monitoring values in combination with external factors, so as to quantitatively and qualitatively analyze the occurrence and evolution of PM2.5.

4.1 Model Analysis and Solution of Problem one**4.1.1 Model preparation**

Partial least squares regression analysis model provides a multi-pair multi-linear regression analysis modeling method, especially when the number of two groups of variables is large and there are multiple correlations, and the number of observation data (sample size) is small, the model established by partial least squares regression analysis has the advantages that traditional classical regression analysis methods do not have.

4.1.2 Data Processing

Because the data is too large, when we study the correlation and independence of the six basic monitoring indicators in AQI, we choose the urban area as the representative of the regional table for the relevant research; and because of the lack of a small part of the data, we ignore it and do not do the correlation research.

4.1.3 Establishment and Solution of Model

x_1, x_2, x_3, x_4, x_5 represent independent variables such as sulfur dioxide, nitrogen dioxide, particulate matter, carbon monoxide and ozone, respectively, and Y represents dependent variables such as fine particulate matter (PM2.5). The partial least squares regression analysis is used to establish the regression equation of Y and x_1, x_2, x_3, x_4, x_5 to determine the influence of five independent variables x_1, x_2, x_3, x_4, x_5 on y. The observation data matrix of independent variable is

$A = (a_{ij})_{457 \times 5}$, and the observation data matrix of dependent variable is $B = [b_1, \dots, b_{457}]^T$.

Step1: Data standardization

Converting each index value a_{ij} into a standardized index value \tilde{a}_{ij} ,

$$\tilde{a}_{ij} = \frac{a_{ij} - \mu_j^{(1)}}{s_j^{(1)}}, i = 1, 2, 3, \dots, 457, j = 1, 2, \dots, 5 \tag{1}$$

Among them, $\mu_j^{(1)} = \frac{1}{457} \sum_{i=1}^{457} a_{ij}, s_j^{(1)} = \sqrt{\frac{1}{457-1} \sum_{i=1}^{457} (a_{ij} - \mu_j^{(1)})^2}, j = 1, 2, \dots, 5$ are the sample mean and standard deviation of the second independent variable x_j . Correspondingly,

$$\tilde{x}_j = \frac{x_j - \mu_j^{(1)}}{s_j^{(1)}}, j = 1, 2, \dots, 5 \tag{2}$$

is a standardized indicator variable.

Similarly, b_i is converted into a standardized index value \tilde{b}_i , which has

$$\tilde{b}_i = \frac{b_i - \mu^{(2)}}{s^{(2)}}, i = 1, 2, \dots, 457 \tag{3}$$

where $\mu^{(2)} = \frac{1}{457} \sum_{i=1}^{457} b_i, s^{(2)} = \sqrt{\frac{1}{457-1} \sum_{i=1}^{457} (b_i - \mu^{(2)})^2}$ are the sample mean and standard deviation of the dependent variable y . Correspondingly,

$$\tilde{y} = \frac{y - \mu^{(2)}}{s^{(2)}} \tag{4}$$

is the corresponding standardized variable.

Step2: Computing the correlation coefficient matrix

Table 1 Coefficient Matrix of 6 Basic Monitoring Indicators in AQI

	SO2	NO2	PM10	CO	O3	PM2.5
SO2	1.0000	0.8085	0.7164	0.6766	-0.1764	0.6749
NO2	0.8085	1.0000	0.7799	0.7111	-0.1162	0.6915
PM10	0.7164	0.7799	1.0000	0.7370	-0.0750	0.8291
CO	0.6766	0.7111	0.7370	1.0000	-0.3455	0.8806
O3	-0.1764	-0.1162	-0.0750	-0.3455	1.0000	-0.3457
PM2.5	0.6749	0.6915	0.8291	0.8806	-0.3457	1.0000

Table 1 gives the correlation coefficient matrix of these six variables. From the correlation coefficient matrix, it can be seen that the correlation coefficient between O3 and other five variables is small, indicating that the linear correlation between O3 and NO2, CO, PM10, PM2.5 and SO2 is weak; the correlation between NO2, CO, PM10, PM2.5 and SO2 is large, indicating that they have a greater impact on SO2; besides the weak correlation between O3 and CO, the other four variables have a strong linear relationship with CO; while for PM2.5, CO, PM10 and SO2, there is a strong linear relationship between O3 and CO. The correlation coefficient between CO and PM10 is larger than other variables, indicating that CO and PM10 have the greatest impact on PM2.5.

Step3: The regression equation between dependent variable group and independent variable group.

The partial least squares regression analysis shows that the equation of y and x_1, x_2, x_3, x_4, x_5 is $y = -0.12x_1 - 0.12x_2 + 1.04x_3 + 3.34x_4 - 0.32x_5 - 29.04$ (5)

4.2 Model Analysis and Solution of Problem 2

The spatial and temporal distribution of PM2.5 and its related laws were obtained by using MATLAB and EXECL analysis software.

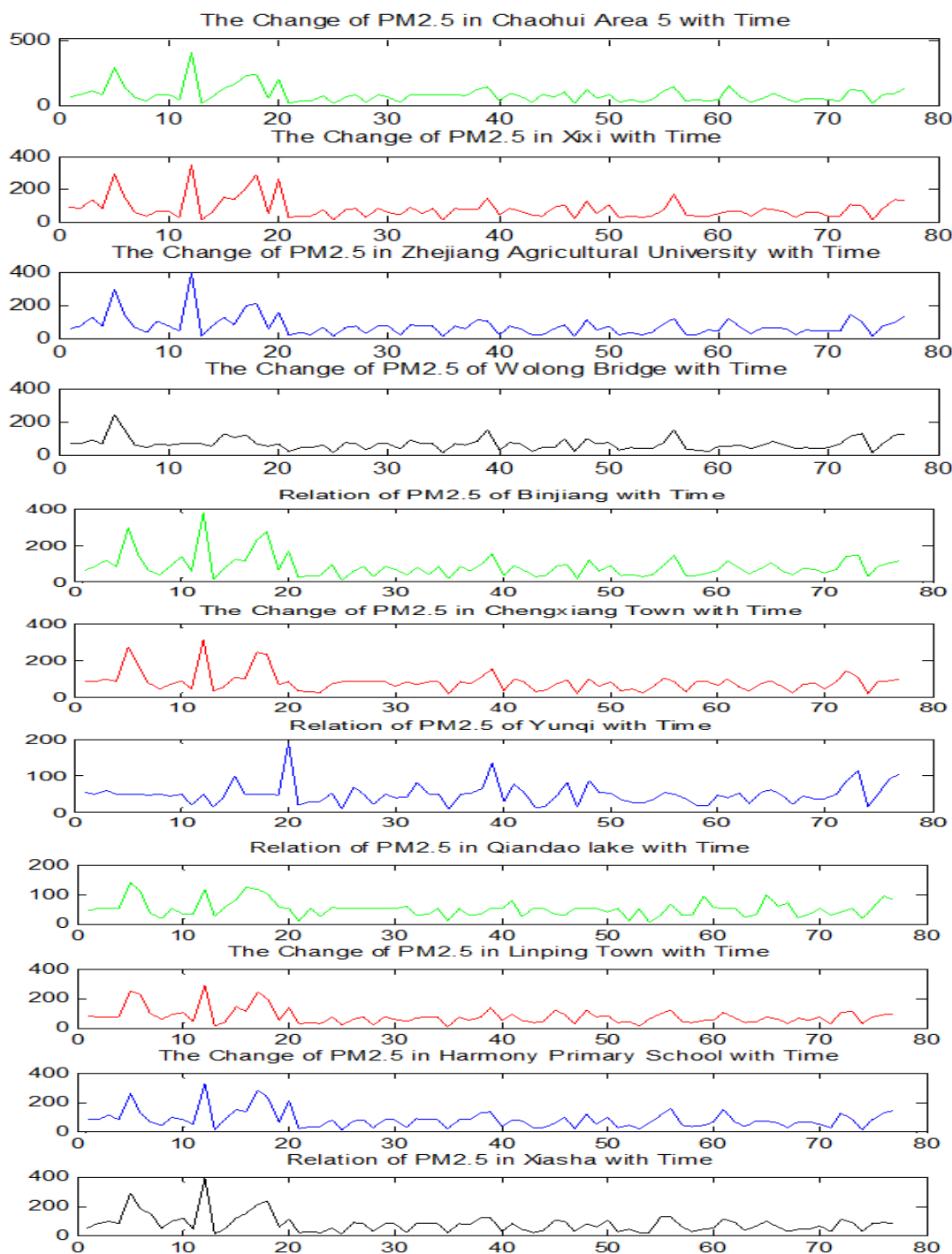


Fig.1 Spatial and Temporal Distribution of PM2.5 in Hangzhou

As can be seen from Figure 1, the PM2.5 values of Qiandao Lake, Linping Town, Harmony Primary School, Xiasha, Chaohui District, Xixi, Zhejiang Nongda, Binjiang and Chengxiang Town all show a trend of rising first and then declining with time, and all of them have a continuous peak from January 14 to February 2. After 2 days, the PM2.5 values gradually tend to flatten. This indicates that the cause of the sudden rise of PM2.5 index value at the beginning is probably caused by human factors. Because PM2.5 is very harmful to human health, the government attaches great importance to PM2.5 after its sharp rise, and takes correct measures to eventually make the value of PM2.5 in this area tend to be flat, and gradually approach 0. PM2.5 in Wolongqiao area remained stable after reaching the extreme value at the beginning, indicating that environmental control is extremely fast. The change trend of PM2.5 value in Yunqi area is unstable. There is a peak around February 2, and

then there are two successive peaks. This shows that the degree of governance in Yunqi area is not enough and there is a rebound phenomenon.

4.2.1 Model analysis

(1) Data preparation

Table2 Averages of pollutants from 1-15 to 3-31 at 11 monitoring points

	SO2	NO2	PM10	CO	O3	PM2.5
Huzhou	25.74	53.04	1.00	54.78	103.54	78.52
Jiaxing	26.12	52.23	1.65	65.88	101.57	68.73
Jinhua	30.21	46.93	1.02	44.82	111.64	79.50
Lishui	18.97	40.46	1.09	61.08	78.11	49.90
Ningbo	20.01	49.09	1.14	60.26	81.63	56.69
Quzhou	28.91	41.17	0.79	43.60	100.12	74.64
Shaoxing	35.44	62.89	1.04	50.79	109.98	73.21
Taizhou	15.86	36.52	1.08	60.20	79.28	53.01
Wenzhou	25.05	47.42	2.55	53.71	92.75	64.39
Zhoushan	8.78	24.52	0.92	73.81	57.84	41.39
Hangzhou	23.51	53.02	0.93	46.20	108.72	72.65

Table 3 Number of air quality categories in each monitoring site (units: days)

	excellent	good	Mild pollution	Moderate pollution	Severe pollution	Serious pollution
Hangzhou	6	34	24	8	4	0
Ningbo	13	42	18	1	2	0
Wenzhou	12	51	10	1	2	0
Jiaxing	9	44	15	4	4	0
Huzhou	4	27	32	5	8	0
Shaoxing	4	36	24	7	5	0
Jinhua	4	37	26	2	6	1
Quzhou	3	36	26	4	6	1
Zhoushan	30	38	6	2	0	0
Taizhou	14	46	13	1	1	1
Lishui	16	51	6	1	2	0

Graphical processing of the data in Table3 to get Figure 2.

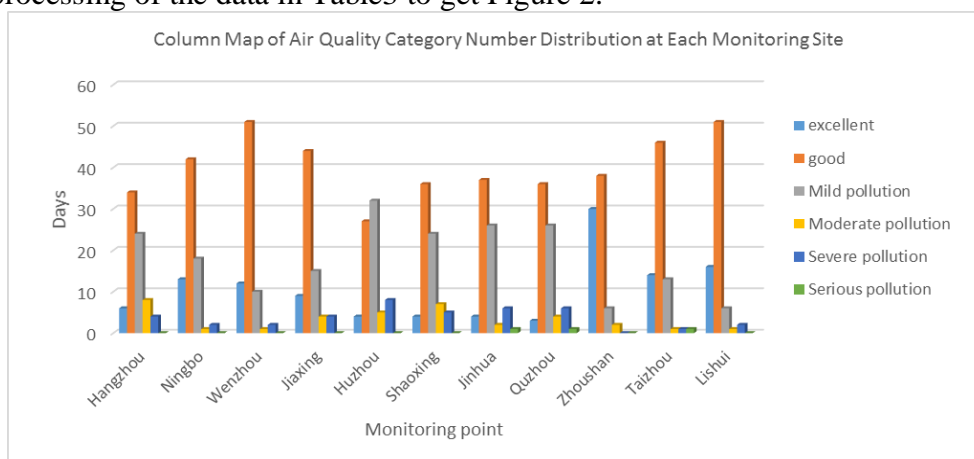


Fig.2 Histogram of air quality category number distribution at each monitoring site

Statistical analysis shows that there are fewer days with good air quality in Zhejiang Province, but there are more days with good air quality in Zhoushan area during 77 days of monitoring. Figure 2 shows that the air quality of most days at each monitoring point is good. At the same time, it also shows that the air quality of each monitoring point is slightly polluted days, and reflects that the air quality of each monitoring point is moderately, severely and severely polluted days are relatively few, but there are also several monitoring points with a larger proportion, such as Huzhou, Shaoxing, Jinhua and Quzhou, where the air pollution situation is relatively serious.

4.2.2 Establishment and Solution of Model

A. Model II: Fuzzy Subjective Discriminant Evaluation and Analysis Model

Step1: Building Subordinate Degree of Air Quality Index

In this paper, the grade of urban air quality index is divided into six categories, namely {serious pollution, severe pollution, moderate pollution, light pollution, good, excellent}, i.e. {6, 5, 4, 3, 2, 1} respectively.

Table4 Air Quality Classification

air quality	Serious pollution	Severe pollution	Moderate pollution	Mild pollution	good	excellent
Air quality grade (\leq)	6	5	4	3	2	1

According to the actual situation, this paper chooses the membership function of the large Cauchy distribution as the degree of membership function.

$$f(x) = \begin{cases} [1 + \alpha(x - \beta)^{-2}]^{-1} & 1 \leq x \leq 3.5 \\ (a \ln x + b) & 3.5 \leq x \leq 6 \end{cases} \tag{6}$$

α 、 β 、 a 、 b is the undetermined coefficient.

When the air quality is good, the membership degree is 0 and $f(1) = 0.01$ is chosen.

When the air quality is seriously polluted, the membership degree is 1 and $f(6) = 1$ is chosen.

Take $f(3.5) = 0.8$;

The calculated coefficients α 、 β 、 a 、 b are $\alpha = 1.7328$ 、 $\beta = 0.8677$ 、 $a = 0.3711$ 、 $b = 0.3351$.

The membership function of air quality is obtained as follows:

$$f(x) = \begin{cases} [1 + 1.7328(x - 0.8677)^{-2}]^{-1} & 1 \leq x \leq 3.5 \\ (0.3711 \ln x + 0.3351) & 3.5 \leq x \leq 6 \end{cases} \tag{7}$$

The membership degree of air quality is calculated by using the membership function of air quality as shown in Table 6.

Table 5 Subordinate Degree of Air Quality Grade

air quality	Serious pollution	Severe pollution	Moderate pollution	Mild pollution	good	excellent
Air quality grade (\leq)	1	0.9323	0.8495	0.7241	0.4253	0.001

Step2: Establishing a Matrix of Matrix of Optimum Selection for Paired Comparative Judgment of Objects

Firstly, the standard of scoring system is established, and 11 regions in Zhejiang Province are scored by scoring system, so as to get the matrix of pairwise comparison judgment and selection.

Establishing the Standard of Scoring System

Table 6 Scoring System Table

Relative importance	Equally important	A little more important	More important	very important	Absolutely important
Comparison score	1	2	3	4	5

Establishing the Matrix of Optimum Selection for Paired Comparative Judgment

Table 7 Comparisons in pairs

Contaminants	SO2	NO2	CO	O3	PM10	PM2.5
SO2	1.0000	0.2500	0.5000	0.3333	0.2000	0.2500
NO2	4.0000	1.0000	0.5000	0.5000	0.2500	0.3333
CO	2.0000	2.0000	1.0000	0.5000	0.5000	0.5000
O3	3.0000	2.0000	2.0000	1.0000	0.3333	0.2500
PM10	5.0000	4.0000	2.0000	3.0000	1.0000	0.2500
PM2.5	4.0000	3.0000	2.0000	4.0000	5.0000	1.0000

Consistency test

(i) Computing consistency index CI

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{8}$$

(ii) Identity formula

$$RI = \frac{\lambda'_{max} - n}{n - 1} \tag{9}$$

(iii) Computational Consistency Ratio CR

$$CR = \frac{CI}{RI} \tag{10}$$

CR = 0.088 < 0.1 is calculated, Therefore, the consistency of matrices is acceptable.

Weight calculation

Using Weight Formula

$$\omega_i = \sqrt[6]{m_{i1} \cdot m_{i2} \cdot \dots \cdot m_{i6}} \tag{11}$$

Among them, $m_1, m_2, m_3, m_4, m_5, m_6$ denotes SO2, NO2, CO, O3, PM10, PM2.5, respectively.

Calculate the initial weight

Table8 Initial Weight Table

Contaminants	SO2	NO2	CO	O3	PM10	PM2.5
Initial weight	0.3574	0.6609	0.8909	1.0000	1.7627	2.7982

(5) Obtain the weight proportion of each pollutant.

Formulas for calculating normalized weight coefficients

$$\omega_i = \omega_i' / \sum_{i=1}^6 \omega_i' \tag{12}$$

The normalized weight coefficients of each pollutant are calculated.

Table 9 Normalized Weight Table

Contaminants	SO2	NO2	CO	O3	PM10	PM2.5
Normalized weight	0.0478	0.0885	0.1193	0.1339	0.236	0.3746

Step3: Building an Air Quality Index Evaluation Function

The air quality index evaluation function of 11 regions in Zhejiang Province was obtained by substituting six indexes into the following formula.

$$f = \sum_{j=1}^r \omega_j m_j \tag{13}$$

Among them, for the weight of the first air quality index, the regional air quality index evaluation function is a comprehensive evaluation of the regional air quality index.

The evaluation function of air quality index is constructed as follows:

$$f = 0.0478m_1 + 0.0885m_2 + 0.1193m_3 + 0.1339m_4 + 0.2360m_5 + 0.3746m_6 \tag{14}$$

Step4: Obtain the membership degree of air quality index in 11 regions of Zhejiang Province.

The air quality sub-index data of 11 regions are substituted into formula (14) and table 10 is obtained.

Table 10 Urban Air Quality Index Assessment Value Table

region	Air Quality Index Assessment Value	region	Air Quality Index Assessment Value	region	Air Quality Index Assessment Value
Huzhou	0.5660	Ningbo	-0.2822	Wenzhou	0.2577
Jiaxing	0.5214	Quzhou	0.0264	Zhoushan	-1.2986
Jinhua	0.5545	Shaoxing	0.6171	Hangzhou	0.2699
Lishui	-0.6332	Taizhou	-0.5990		

From Table 5, we can see that the air quality index evaluation values of 11 regions are obtained, and the regional air quality evaluation values are normalized, and the membership degree of the air quality index of each region is obtained as shown in Table 11.

Table 11 Subordinate Degree of Urban Air Quality Index

region	Subordinate Degree of Air Quality Index	region	Subordinate Degree of Air Quality Index	region	Subordinate Degree of Air Quality Index
Huzhou	0.9733	Ningbo	0.5306	Wenzhou	0.8124
Jiaxing	0.9500	Quzhou	0.6916	Zhoushan	0.0000
Jinhua	0.9673	Shaoxing	1.0000	Hangzhou	0.8188
Lishui	0.3474	Taizhou	0.3652		

According to Table 6, the air quality index grades of 11 regions in Table 11 are classified into Table 12.

Table 12 Regional Air Quality Grades

region	Air quality grade	region	Air quality grade	region	Air quality grade
Huzhou	Severe pollution	Ningbo	Mild pollution	Wenzhou	Moderate pollution
Jiaxing	Severe pollution	Quzhou	Mild pollution	Zhoushan	excellent
Jinhua	Severe pollution	Shaoxing	Serious pollution	Hangzhou	Moderate pollution
Lishui	good	Taizhou	good		

According to Tables 11 and 12, the most polluted area is Shaoxing.

Model III: 0-1 Integer Programming Model

Step1: Establishing standardized processing tables

Because different variables have different units and varying degrees of variation, the data are optimized and standardized.

According to formula (1), the standardized data are calculated as shown in table 13.

Table 13 6 standardized air quality sub-index data

region	SO ₂	NO ₂	CO	O ₃	PM10	PM2.5
Huzhou	0.3039	0.6787	2.9955	-0.1213	0.6080	1.0779
Jiaxing	0.3557	0.5993	-0.3110	1.0586	0.4922	0.3096
Jinhua	0.9128	0.0797	-0.3497	-1.1800	1.0842	1.1548
Lishui	-0.6183	-0.5546	-0.3454	0.5484	-0.8870	-1.1680
Ningbo	-0.4767	0.2914	-0.3423	0.4612	-0.6801	-0.6352
Quzhou	0.7357	-0.4850	-0.3638	-1.3097	0.4069	0.7734
Shaoxing	1.6253	1.6443	0.0296	-0.5454	0.9866	0.6612
Taizhou	-1.042	-0.9409	-0.346	0.4549	-0.8183	-0.9240
Wenzhou	0.2099	0.1277	-0.2558	-0.235	-0.0263	-0.0310
Zhoushan	-2.0064	-2.1173	-0.3558	1.9016	-2.0787	-1.8358
Hangzhou	0.0001	0.6767	-0.3552	-1.0333	0.9125	0.6172

Step2: Establishing 0-1 Integer Programming Equation

$S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}, S_{11}$ means Huzhou, Jiaxing, Jinhua, Lishui, Ningbo, Quzhou, Shaoxing, Taizhou, Wenzhou, Zhoushan and Hangzhou.

The 0-1 variable is used to control whether to choose the city or not, but not when $S_i = 0$, and when $S_i = 1$, the city can be selected, then the 0-1 integer programming equation can be listed as follows:

Objective function:

$$\max Z = \sum_{i=1}^{11} \sum_{j=1}^6 S_i m_{ij} \tag{15}$$

$$s.t. \quad S_1, S_2, \dots, S_{11} \in \{0,1\} \tag{16}$$

$$S_1 + S_2 + \dots + S_{11} = 5 \tag{17}$$

Statistical analysis and image analysis were carried out by using R language and Lingo respectively. The results were as follows: $\{S_1, S_2, S_3, S_7, S_9\}$.

A screenshot of the Lingo image analysis results is shown in Figure 3.

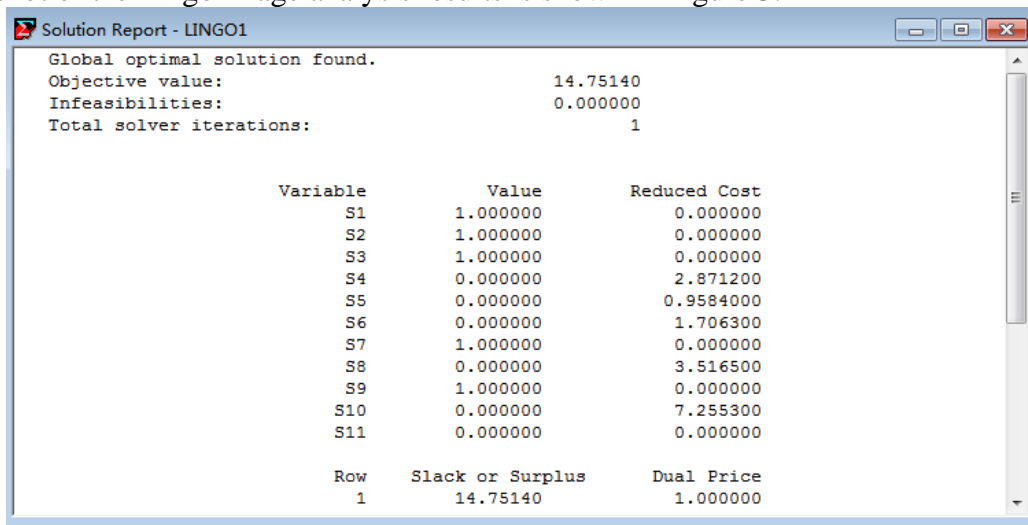


Fig.3 0-1 Integer Programming Result Diagram

Step3: Ranking of Air Quality Index in Five Cities

For the sorting problem, the cumulative total score is obtained in Table 14.

Table 14 Cumulative Total Scoresheet

Air quality index	SO2	NO2	CO	O3	PM10	PM2.5	Total score
Huzhou	0.3039	0.6787	-0.3321	-0.1213	0.6080	1.0779	2.2150
Jiaxing	0.3557	0.5993	0.6494	1.0586	0.4922	0.3096	3.4648
Jinhua	1.6253	1.6443	-0.1140	-0.5454	0.9866	0.6612	4.2579
Shaoxing	1.6253	1.6443	-0.0961	-0.5454	0.9866	0.6612	4.2759
Wenzhou	0.2099	0.1277	2.3944	-0.2350	-0.0263	-0.0310	2.4397

As can be seen from the table above, the results are as follows:

$$\{S_7 > S_3 > S_2 > S_9 > S_1\} \tag{18}$$

From the results, it can be seen that the pollution degree of these five cities is from large to small: Shaoxing, Jinhua, Jiaxing, Wenzhou and Huzhou. The most polluted city is Shaoxing.

Model IV: Air Quality Index Analysis Chart Model

Step1: Explain the Air Quality Index Analysis Chart

The air quality index analysis graph method is suitable for evaluating the objects described by multi-attribute architecture globally and wholly. Air Quality Analysis Graph Method is a multi-variable comparative analysis technique based on a graph similar to that on the display screen of navigation radar. It consists of several concentric regular polygons. Concentric regular polygons draw several rays outward, which are equidistant. Each regular polygon represents a certain score, which is increased from the inside to the outside, and one is put at the end of each ray. Indicators for investigation. The basic form of this model's ability and quality analysis graph is shown in Fig. 4.



Fig. 4 Basic form chart of air quality index analysis chart method

The air quality index analysis chart can be used to analyze the changes of the same pollutant in different cities at the same time. The qualitative evaluation results can be obtained by observation and calculation.

Step2: Establishing an Air Quality Index Analysis Table

The air quality index analysis chart method is used for comprehensive evaluation, that is, the evaluation index status of the evaluation object system is expressed by two-dimensional plane graph.

Table 15 Initial Value Table for Air Quality Index Analysis

City	SO2	NO2	CO	O3	PM10	PM2.5	Total score
Huzhou	0.3039	0.6787	-0.3321	-0.1213	0.6080	1.0779	2.2150
Jiaxing	0.3557	0.5993	0.6494	1.0586	0.4922	0.3096	3.4648
Jinhua	0.9128	0.0797	-0.2231	-1.1800	1.0842	1.1548	1.8284
Lishui	-0.6183	-0.5546	0.1586	0.5484	-0.8870	-1.1680	-2.5210
Ningbo	-0.4767	0.2914	0.4313	0.4612	-0.6801	-0.6352	-0.6080
Quzhou	0.7357	-0.4850	-1.4773	-1.3097	0.4069	0.7734	-1.3559
Shaoxing	1.6253	1.6443	-0.1140	-0.5454	0.9866	0.6612	4.2579
Taizhou	-1.0420	-0.9409	0.1041	0.4549	-0.8183	-0.9240	-3.1661
Wenzhou	0.2099	0.1277	2.3944	-0.2350	-0.0263	-0.0310	2.4397
Zhoushan	-2.0064	-2.1173	-0.7684	1.9016	-2.0787	-1.8358	-6.9051
Hangzhou	0.0001	0.6767	-0.8229	-1.0333	0.9125	0.6172	0.3504

In order to facilitate the analysis, we normalized the initial value of air quality index analysis in Table 13 to [0,10], and obtained the air quality index analysis values of 11 cities in Zhejiang Province, as shown in Table 16.

Table 16 Analytical Value of Air Quality Index

City	SO2	NO2	CO	O3	PM10	PM2.5
Huzhou	6.2685	6.6747	5.7851	5.9938	7.4194	9.7428
Jiaxing	6.2906	6.612	6.8301	7.4038	7.1754	7.174
Jinhua	6.5285	6.2016	5.9012	4.7286	8.4225	10
Lishui	5.8746	5.7007	6.3076	6.794	4.2702	2.233
Ningbo	5.9351	6.3689	6.5978	6.6899	4.7061	4.0147
Quzhou	6.4529	5.7556	4.5659	4.5736	6.9959	8.7247
Shaoxing	6.8328	7.4373	6.0173	5.4869	8.2169	8.3495
Taizhou	5.6937	5.3956	6.2495	6.6822	4.4151	3.0491
Wenzhou	6.2283	6.2396	8.6878	5.8578	6.0832	6.0352
Zhoushan	5.2818	4.4665	5.3206	8.4111	1.76	0
Hangzhou	6.1387	6.6731	5.2626	4.9039	8.0609	8.2026

Step3: Get the air quality index analysis chart

According to table 16, air quality index analysis chart method was used to obtain 11 regions of Zhejiang Province air quality index analysis figure 5.

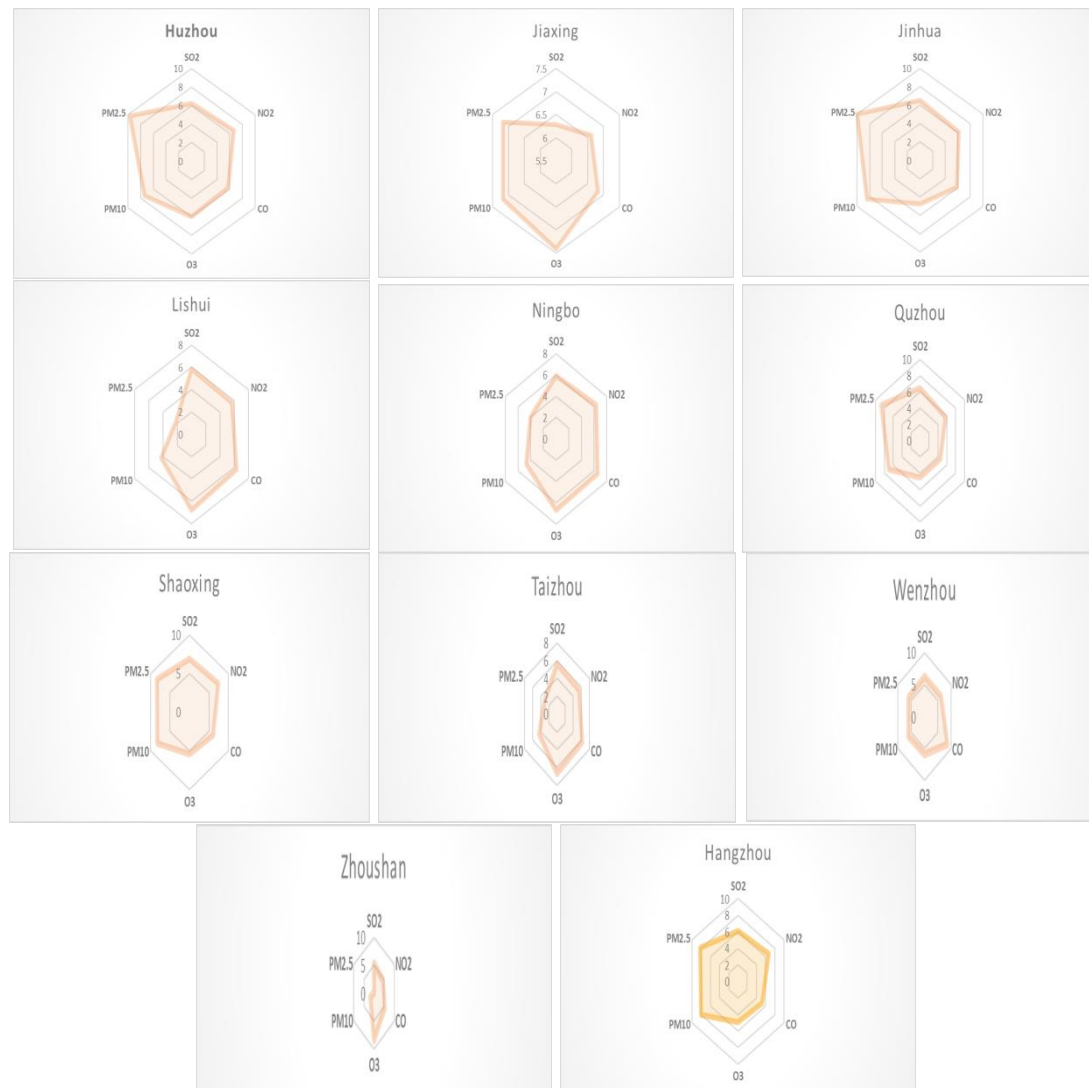


Fig. 5 Regional Air Quality Index Analysis Chart

According to the area formula of trigonometric function, the area of shadow is calculated separately, as shown in Table 17.

Table 17 Air Quality Index Area Table

region	Air quality index area	region	Air quality index area	region	Air quality index area
Huzhou	126.3161	Ningbo	85.0295	Wenzhou	110.1752
Jiaxing	124.3228	Quzhou	97.7794	Zhoushan	42.7062
Jinhua	125.3797	Shaoxing	128.4339	Hangzhou	110.3144
Lishui	69.6784	Taizhou	70.9219		

Table 17 shows that the order of air quality index area from large to small is: Shaoxing, Huzhou, Jinhua, Jiaxing, Hangzhou, Wenzhou, Quzhou, Ningbo, Taizhou, Lishui and Zhoushan. The most polluted area is Shaoxing.

Model V: Euclid Closeness Analysis Model

Step1: Model preparation

Let A and B be two fuzzy subsets on the universe U, then the Euclid closeness between A and B (a measure of the closeness between two fuzzy subsets) is defined as:

$$e(A, B) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\mu_A(X_i) - \mu_B(X_i))^2} \tag{19}$$

In the formula, A and B are the membership degrees of the fuzzy subsets $\mu_A(X_i)$ and $\mu_B(X_i)$. A and B can be expressed as vectors:

$$A = \{\mu_A(X_1), \mu_A(X_2), \dots, \mu_A(X_n)\} \tag{20}$$

$$B = \{\mu_B(X_1), \mu_B(X_2), \dots, \mu_B(X_n)\} \tag{21}$$

If B, A_1, A_2, \dots, A_m is a known M+1 fuzzy subset, and has:

$$e(B, A_k) = \min_{1 \leq j \leq m} \{e(B, A_j)\} \tag{22}$$

It is said that the fuzzy subset B and A are closest to each other.

Step2: Establishing Subordinate Degree of Pollutants

In this paper, the following membership function formulas are used to determine the membership degree of each single factor. It is assumed that the category of regional air quality index is divided into M levels.

$$\mu_{A_k}(X_i) = \frac{V_{ik}}{V_{im}} \quad k = 1, 2, \dots, m \tag{23}$$

$$\mu_B(X_i) = \frac{C_i}{V_{im}} \tag{24}$$

(note): When $C_i > V_{im}$, $\mu_B(X_i) = 1$

In Formula $\mu_{A_k}(X_i)$: Subjection Degree of Standard Limit at Level K of Item A

$\mu_B(X_i)$ - Subordinate degree of actual monitoring value of item I;

V_{ik} - The standard limit of level K of Item I;

V_{im} - The maximum pollution level standard limit of Item I;

C_i - The measured value of item I.

Step3: Establishing Air Quality Grades

From the observation data, we divide the air quality level into six grades, as shown in Table 17.

Table 18 Six Air Quality Standards

project	level					
	1	2	3	4	5	6
	Concentration limits (mg/m ³)					
SO2	0.02	0.06	0.10	0.20	0.30	0.40
NO2	0.04	0.04	0.08	0.12	0.16	0.24
PM10	0.05	0.15	0.25	0.35	0.42	0.50
CO	2.00	4.00	4.00	6.00	14.00	20.00
O3	0.16	0.20	0.30	0.40	0.80	1.00
PM2.5	0.04	0.08	0.12	0.15	0.25	0.35

The data units in tables 18 and 3 are unified as mg/m³ and then substituted into (23) formula. The membership degree of each item of air quality standard and the corresponding fuzzy subset of each monitoring point can be obtained. The fuzzy subset A can be expressed in vector form as follows:

$$\begin{aligned}
 A_1 &= \{0.05, 0.17, 0.10, 0.10, 0.16, 0.11\} \\
 A_2 &= \{0.15, 0.17, 0.30, 0.20, 0.20, 0.23\} \\
 A_3 &= \{0.25, 0.33, 0.50, 0.30, 0.30, 0.34\} \\
 A_4 &= \{0.50, 0.50, 0.70, 0.40, 0.40, 0.43\} \\
 A_5 &= \{0.75, 0.67, 0.84, 0.80, 0.80, 0.71\} \\
 A_6 &= \{1.00, 1.00, 1.00, 1.00, 1.00, 1.00\}
 \end{aligned}$$

Similarly, by substituting formula (24), the vector form is expressed as the following fuzzy subset B.

$$\begin{aligned}
 B_1 &= \{0.06, 0.22, 0.21, 0.05, 0.05, 0.22\} \\
 B_2 &= \{0.07, 0.22, 0.20, 0.06, 0.07, 0.20\} \\
 B_3 &= \{0.08, 0.20, 0.22, 0.05, 0.04, 0.23\} \\
 B_4 &= \{0.05, 0.17, 0.16, 0.05, 0.06, 0.14\} \\
 B_5 &= \{0.05, 0.20, 0.16, 0.06, 0.06, 0.16\} \\
 B_6 &= \{0.07, 0.17, 0.20, 0.04, 0.04, 0.21\} \\
 B_7 &= \{0.09, 0.26, 0.22, 0.05, 0.05, 0.21\} \\
 B_8 &= \{0.04, 0.15, 0.16, 0.05, 0.06, 0.15\} \\
 B_9 &= \{0.06, 0.20, 0.19, 0.07, 0.05, 0.18\} \\
 B_{10} &= \{0.02, 0.10, 0.12, 0.05, 0.07, 0.12\} \\
 B_{11} &= \{0.06, 0.22, 0.22, 0.05, 0.05, 0.21\}
 \end{aligned}$$

The Euclid closeness degree between the fuzzy subset $B_i (i=1,2,\dots,11)$ and $A_j (j=1,2,\dots,6)$

is calculated by formula (22). Then, according to the principle of closeness degree, the closest A_k to B_i is found and the air quality evaluation is obtained, such as table 19.

Table 19 Closeness of monitoring points to air quality standards at all levels

City Proper	Euclid proximity						Minimum closeness
	e(B1,A1)	e(B1,A2)	e(B1,A3)	e(B1,A4)	e(B1,A5)	e(B1,A6)	
Huzhou(B1)	0.09	0.12	0.23	0.39	0.65	0.91	0.09
Jiaying(B2)	0.09	0.12	0.23	0.39	0.65	0.91	0.09
Jinhua(B3)	0.10	0.12	0.23	0.39	0.65	0.91	0.10
Lishui(B4)	0.08	0.14	0.25	0.41	0.67	0.93	0.08
Ningbo(B5)	0.08	0.13	0.25	0.40	0.67	0.92	0.08
Quzhou(B6)	0.10	0.13	0.24	0.40	0.66	0.91	0.10
Shaoxing(B7)	0.12	0.23	0.11	0.38	0.65	0.90	0.12
Taizhou(B8)	0.08	0.14	0.25	0.41	0.67	0.93	0.08
Wenzhou(B9)	0.08	0.13	0.24	0.40	0.66	0.91	0.08
Zhoushan(B10)	0.08	0.15	0.27	0.43	0.69	0.94	0.08

Hangzhou(B11)	0.10	0.12	0.23	0.39	0.66	0.91	0.10
---------------	------	------	------	------	------	------	------

Table 19 shows that Shaoxing is the most polluted area.

4.2.3 Summary and Analysis

According to the analysis of fuzzy subjective discriminant evaluation, the areas with higher pollution degree are Shaoxing, Jinhua, Quzhou and Huzhou.

The 0-1 integer programming results show that Shaoxing, Jinhua, Jiaying, Wenzhou and Huzhou are the most polluted areas, and Shaoxing is the most polluted area.

According to the analysis of air quality index, the order of pollution degree is: Shaoxing, Huzhou, Jinhua, Jiaying, Hangzhou, Wenzhou, Quzhou, Ningbo, Taizhou, Lishui and Zhoushan. The most polluted areas are Shaoxing.

According to Euclid proximity analysis, Shaoxing is the most polluted area.

To sum up, the most polluted city is Shaoxing.

4.3 Model Analysis and Solution of Problem 3

4.3.1 Analysis of the problem

Firstly, we consider the influence of time on PM2.5 concentration, and then establish a three-time exponential smoothing prediction model to analyze the influence of time on PM2.5 concentration. Then we consider the influence of external conditions such as wind speed, temperature and humidity on PM2.5 concentration, and compare the predicted value with the actual monitoring value, so as to quantitatively and qualitatively analyze the occurrence and evolution of PM2.5.

4.3.2 Establishment and Solution of Model

Step1: Data Processing

Because there are too many data, we replace the daily PM2.5 value of Hangzhou with the daily average value of PM2.5 of 11 monitoring points in Hangzhou, as shown in Appendix 5.

Step2: Modeling

When time series y_1, y_2, \dots , α is a weighted coefficient, $0 < \alpha < 1$, cubic exponential smoothing is performed on the basis of quadratic exponential smoothing, and its calculation formula is as follows:

$$\begin{cases} S_t^{(1)} = \alpha y_t + (1 - \alpha) S_{t-1}^{(1)} \\ S_t^{(2)} = \alpha S_t^{(1)} + (1 - \alpha) S_{t-1}^{(2)} \\ S_t^{(3)} = \alpha S_t^{(2)} + (1 - \alpha) S_{t-1}^{(3)} \end{cases} \quad (25)$$

Among them, $S_t^{(3)}$ is the cubic exponential smoothing value.

Prediction model of cubic exponential smoothing method

$$\hat{y}_{t+m} = \alpha_t + b_t m + C_t m^2, m = 1, 2, \dots, \quad (26)$$

Among

$$\left\{ \begin{aligned} \alpha_t &= 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)} \\ b_t &= \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)S_t^{(1)} - 2(5-4\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}] \\ c_t &= \frac{\alpha^2}{2(1-\alpha)^2} [S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}] \end{aligned} \right. \quad (27)$$

Step3: Solution of the Model

Using MATLAB^[4] programming to solve knows $\alpha_{77} = 108.64, b_{77} = 12.74, c_t = 0.63$;

The prediction curve is D:

$$y_{77+m} = 108.64 + 12.74m + 0.63m^2, m = 1, 2, \dots \quad (28)$$

The evolution of PM2.5 is Figure 7:

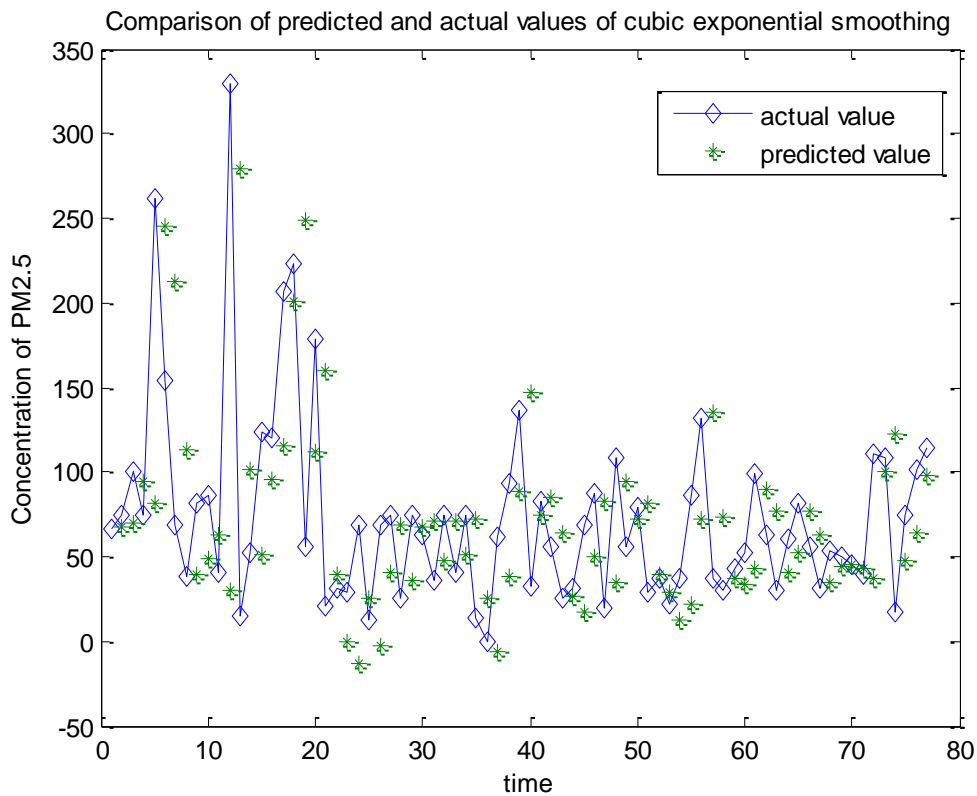


Fig.6 Comparison of the predicted and actual values of the third exponential smoothing

Figure 7 shows that most of the predicted values of PM2.5 predicted by the cubic exponential method are basically consistent with the actual values, and the evolution law of PM2.5 is obtained. It also shows that PM2.5 has a great influence on the change of time. Based on the evolution law of PM2.5 in this chart and the data and information in the corresponding time between Annex III and the above figure, we conclude that with the increase of wind speed and temperature, the PM2.5 diffusion range will gradually expand, and along the same wind direction, the PM2.5 variation will be larger; with the gradual increase of humidity, the PM2.5 diffusion distribution will not change very much, relatively gentle. From the above analysis, it can be seen that wind speed, temperature and humidity in external factors will affect the PM2.5 pollutant diffusion to a certain extent.

5. Conclusion

We can see from the correlation coefficient matrix that the correlation coefficients of O₃ and other five variables are small, indicating that the linear correlation degree of O₃ with NO₂, CO, PM₁₀, PM_{2.5} and SO₂ is weak; the correlation degree of NO₂, CO, PM₁₀ and PM_{2.5} with SO₂ is large, indicating that they have a greater impact on SO₂; except for the weak correlation between O₃ and CO, the other four variables have a strong linear relationship with CO; but for PM_{2.5}, CO, PM_{2.5} and SO₂. The correlation coefficient between PM₁₀ and PM_{2.5} is higher, which shows that CO and PM₁₀ have the greatest influence on PM_{2.5} compared with other variables. At the same time, we can get the equation by partial least squares regression analysis.

Secondly, we draw the conclusion that Shaoxing, Jinhua, Quzhou and Huzhou are the most polluted areas from the analysis of Fuzzy Subjective Discrimination Evaluation; Shaoxing, Jinhua, Jiaying, Wenzhou and Huzhou are the most polluted areas from 0-1 integer programming, and Shaoxing is the most polluted area; the order of pollution degree from large to small is shown by the analysis of air quality index: Shaoxing, Huzhou, Huzhou. Jinhua, Jiaying, Hangzhou, Wenzhou, Quzhou, Ningbo, Taizhou, Lishui and Zhoushan are the most polluted areas, and Shaoxing is the most polluted area. To sum up, the most polluted city is Shaoxing.

In addition, from Figure 7, it can be seen that most of the predicted PM_{2.5} points predicted by the cubic exponential method are basically consistent with the actual values, and the evolution law of PM_{2.5} is obtained. It also shows that PM_{2.5} has a great influence on the change of time. Based on the evolution law of PM_{2.5} in this chart and the data and information in the corresponding time between Annex III and the above figure, we conclude that with the increase of wind speed and temperature, the PM_{2.5} diffusion range will gradually expand, and along the same wind direction, the PM_{2.5} variation will be larger; with the gradual increase of humidity, the PM_{2.5} diffusion distribution will not change very much, relatively gentle. From the above analysis, it can be seen that wind speed, temperature and humidity in external factors will affect the PM_{2.5} pollutant diffusion to a certain extent.

References

- [1] Environmental Air Quality Standard, National Standard of the People's Republic of China, GB 3095-2012.
- [2] Technical Provisions for Environmental Air Quality Index (Trial Implementation): National Environmental Protection Standard of the People's Republic of China, HJ 633-2012.
- [3] Xue Wei. Statistical Analysis and Application of SPSS (Third Edition). Beijing: Renmin University Press, 2013.
- [4] Solving Applied Mathematics Problem in MATLAB (Third Edition). Beijing: Tsinghua University Press, 2013.
- [5] Jiang Qiyuan, Xie Jinxing, Ye Jun. Mathematical model. Beijing: Higher Education Press, 2003.
- [6] Si Shoukui, Sun Xijing. Mathematical Modeling Algorithms and Applications. Beijing: National Defense Industry Press, 2014.