Study on Optimal Control Strategy of PM2.5 in Wuhan, China

Qiuling Xie, Shanliang Zhu, Zongdi Sun

***Abstract*—**In this paper, we analyzed the correlation relationship among PM2.5 from other five Air Quality Indices (AQIs) based on the grey relational degree, and built a multivariate nonlinear regression equation model of PM2.5 and the five monitoring indexes. For the optimal control problem of PM2.5, we took the partial large Cauchy distribution of membership equation as satisfaction function. We established a nonlinear programming model with the goal of maximum performance to price ratio. And the optimal control scheme is given.

***Keywords*—**Grey relational degree, multiple linear regression, membership function, nonlinear programming.

I. Introduction

FOR the problem of air quality monitoring, forecasting and control, many countries and local governments have formulated the corresponding policies, regulations and management measures. In 2012, China Environmental Protection Department revised and implemented the "ambient air quality standards", and the new standards will get full implementation in 2016. The new standard takes AQI as air quality monitoring index instead of Air Pollution Index (API) which contains six basic monitoring indices, that are SO_2, NO_2, PM10, PM2.5, O_3 and CO [1]. The publication and implementation of new monitoring standards will play an important role in the monitoring of air quality and improving the living environment.

In the new standard, for the first time the PM2.5 concentrations are taken as indicators of air quality monitoring. As we know, PM2.5 is the main factor of producing haze, and it will cause great harm to human health. Therefore, PM2.5 has attracted more and more people's attention [1]-[3], and gradually become a hot issue of scientific research [4]-[8]. In this paper, the PM2.5 and other five monitoring indicators were analyzed quantitatively, and the multiple regression model was established. Then considering the pollution reduction feasible planning of PM2.5, the regional multiple objective optimization model is established under certain assumptions, and the optimal comprehensive PM2.5 control scheme of in the next five years is given.

Qiuling Xie is with Financial Office, Qingdao University of Science & Technology, Shandong, China. (e-mail: qustmaths@126.com).
Shanliang Zhu is with Department of Mathematics & Physics, Qingdao University of Science & Technology, Shandong, China. (e-mail: 2424488018@qq.com).
Zongdi Sun is with Department of Economic & Management, Qingdao University of Science & Technology, Shandong, China. (e-mail: 15054213717@163.com).

II. Correlation Factor Analysis of PM2.5

The data of this paper come from the real time monitoring data of Wuhan City, China [9]. First of all, the data were analyzed and sorted out. By drawing PM2.5 and other five monitoring indicators of the scatter plot, we found that PM2.5 and the other five indicators have a strong correlation.

A. Correlation Analysis of Monitoring Index Based on Grey Correlation Degree

According to the grey incidence theory [10], we built the six factors, namely, PM2.5, sulfur dioxide, nitrogen dioxide, respirable particulate matter, carbon monoxide, ozone of the original data matrix $x_i = \left(x_i(1), x_i(2), \cdots, x_i(k) \right)$, in which $x_i(k)$ means the original data of the $i(i = 1, 2, \cdots, 6)$ factor in the $k(k = 1, 2, \cdots, 238)$ day.

Seeking the initial image matrix $x_i^{'}$:

$$x_i^{'} = \left(\frac{x_i(1)}{x_i(1)}, \frac{x_i(2)}{x_i(1)}, \cdots, \frac{x_i(k)}{x_i(1)} \right) = \left(x_i^{'}(1), x_i^{'}(2), \cdots, x_i^{'}(k) \right) \quad (1)$$

Seeking the differencing sequence $\Delta_{0i}(k)$:

$$\Delta_{0i}(k) = \left| x_0^{i}(k) - x_i^{'}(k) \right| \quad (2)$$

Calculation of correlation coefficient $\xi_{0i}(k)$ and grey correlation degree $\Delta_{0i}(k)$:

$$\xi_{0i}(k) = \frac{\min_i \min_k \Delta_{0i}(k) + \varphi \max_i \max_k \Delta_{0i}(k)}{\Delta_{0i}(k) + \varphi \max_i \max_k \Delta_{0i}(k)} \quad (3)$$

Among them, φ is the significant difference used to increase the difference of correlation coefficient, $\varphi \in (0,1)$, generally $\varphi = 0.5$. Then get the grey correlation degree is:

$$\gamma_{0i} = \frac{1}{n-1} \sum_{k=1}^{n} \xi_{0i}(k), (\mathrm{n} = 1, 2, \cdots, 238) \quad (4)$$

MATLAB software calculation shows that the grey correlation degree of each other from small to large are as in Table I. According to the analysis of the table, we can know the correlation among carbon, particulate matter from PM2.5 are more than 0.9, that means carbon, monoxide and particulate matter from PM2.5 has a strong correlation; Nitrogen dioxide, sulfur dioxide, the correlation among these two factors from PM2.5 were more than 0.8, this means they have strong correlation; However, the correlation between ozone and PM2.5 is only 0.6809, which shows that ozone has a weak correlation with PM2.5.

TABLE I
GREY RELATIONAL DEGREE RANKING OF MONITORING INDEXES

Sort	1	2	3	4	5
Index	CO	PM10	NO_2	SO_2	O_3
Correlation degree	0.9770	0.9724	0.8810	0.8600	0.6809

TABLE II
CORRELATION COEFFICIENT AMONG MONITORING INDICATORS

	CO	NO_2	O_3	SO_2	PM10	PM2.5
CO	1	0.6245	-0.383	0.6570	0.5859	0.8216
NO2	0.6245	1	-0.064	0.8051	0.7274	0.7323
O3	-0.383	-0.064	1	-0.1815	-0.0685	-0.3548
SO2	0.6570	0.8051	-0.181	1	0.6778	0.7241
PM10	0.5859	0.7274	-0.0685	0.6778	1	0.7787
PM2.5	0.8216	0.7323	-0.3548	0.7241	0.7787	1

In Table II, data are correlation coefficients between PM2.5 and other five monitoring indexes. We can see that PM2.5 is negatively correlated with O_3, and the other four items are positively correlated. PM2.5 and CO are highly correlated, and with NO_2, SO_2, PM10 are strongly correlated.

B. Mutlivariate Linear Regression Model between PM2.5 and Five Other Indicators

Step 1: The Establishment of Multiple Linear Regression Model

For the quantitative analysis of the correlation between PM2.5 and the other five indicators, this paper establishes a multiple regression function. The model of multiple linear regression analysis is:

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \tag{5}$$

y means PM2.5 concentrations, $x_1, x_2, ..., x_m (m = 1, 2, ..., 5)$ respectively indicating five AQI index: the sultur dioxide, nitrogen dioxide, respirable particulate matter, carbon monoxide and ozone, $\beta = (\beta_0, \ \beta_1, ..., \beta_m)$ are the regression coefficients.

Now there existing $n(n = 238)$ independent observation data, and by (3), we can get:

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} + \varepsilon_i \\ \varepsilon_i \sim N(0, \sigma^2), \quad i = 1, \cdots, 238 \end{cases} \tag{6}$$

Therefore, we can know:

$$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{15} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & \cdots & x_{n5} \end{bmatrix} \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_{239} \end{bmatrix} \tag{7}$$

Chose an estimated value $\hat{\beta}$ from β, which minimized the square sum of the random error, that is:

$$\min \varepsilon^t \varepsilon = \min(Y - X\beta)^T (Y - X\beta) \tag{8}$$

Bring $\hat{\beta}$ into the model, we can get the estimated equation of the model:

$$\hat{Y} = X^T \hat{\beta}$$

Step 2: Model Solution

Using MATLAB software to calculate the data, get the coefficient of multiple regression equation and its confidence interval, as shown in Table III.

TABLE III
THE COEFFICIENTS OF REGRESSION EQUATIONS AND THEIR CONFIDENCE INTERVALS

The coefficients of Regression equations	Numerical value	Confidence intervals
β_0	-33.025	[-47.032,-19.006]
β_1	0.177	[-0.225,0.579]
β_2	0.444	[0.140,0.748]
β_3	0.878	[0.697,1.058]
β_4	2.208	[1.789,2.627]
β_5	-0.332	[-0.462,-0.202]

From Table III, we can see that PM2.5 with sulfur dioxide, nitrogen dioxide, respirable particulate matter, carbon monoxide and ozone have close relation. Among them, carbon monoxide and ozone has the closest relationship, only ozone and PM2.5 has negative correlation, the other indexes with PM2.5 have positive correlation.

Step 3: Test for Regression Models

Whether there is a linear relationship between the concentration of PM2.5 and the five AQI indices as shown in the regression model is to be tested. By calculating we can get the complex correlation $R = 0.9149 > 0.8$ and the probability $P = 0.000 < 0.05$. This shows that the multivariate regression model is more reliable.

III. PM2.5 OPTIMAL CONTROL MODEL

In this section, we consider the control strategy of PM2.5. In the next five years, we will implement comprehensive management and special treatment planning, and gradually achieve the goal of governance. This paper will address the following issues:

(1) Assume the average annual concentrations of PM2.5 in the

region was estimated to be 280 ($\mu g / m^3$), requirements within the next five years gradually reduce PM2.5 annual average concentration, and ultimately achieve statistical index of average concentration of 35 ($\mu g / m^3$). Please give a reasonable treatment plan, namely given the annual year-end annual average governance indicators.

(2) According to the estimates, for cost of comprehensive control, when reduce a PM2.5 concentration unit, needs to invest a unit cost ($million), management of special input costs is 0.005 times the square of the PM2.5 concentrations reduced ($million). Design effective special treatment plan, not only to achieve the scheduled PM2.5 emission reduction plan, but also make the funds more reasonable, to develop a total funding of five years and year by year budget plan.

In view of the double restriction of the emission reduction and the fund's investment, in this paper, we use the quantitative processing. From the practical significance of the problem, there is reason to believe that in five years, the sooner and the more complete the part of the emission reduction targets evaluated, the better. For example, before the completion of the first two years is better than third years to complete, that is, the evaluation is high. This kind of evaluation has certain fuzziness, and the different evaluation form a fuzzy set, assume the mark set is: {Excellent, Very good, Good, Ordinary, Medium, Bad, Very bad}, and the corresponding set is {7, 6, 5, 4, 3, 2, 1}. According to the actual situation, taking the partial large Cauchy membership function:

$$f(x)=\begin{cases}\left[1+\alpha\left(x-\beta\right)^{-2}\right]^{-1}, & 1\le x\le 4\\ a\ln x+b, & 4\le x\le 7\end{cases} \tag{9}$$

α,β,a,b are parameters in this function. This paper provides the followings: The appraisal is "excellent", the membership degree is 1, namely, $f(7)=1$; The evaluation is "ordinary", the membership degree is 0.8, namely, $f(4)=0.8$. The evaluation is "bad", the membership degree is 0.01, namely, $f(1)=0.01$. After calculation, we can get $\alpha=2.4944,\beta=0.8413,a=0.3574,b=0.3046$. The quantitative values {Excellent, Very good, Good, Ordinary, Medium, Bad, Very bad} of emission reduction at different times are 0.94, 0.88, 0.8, 0.65, 0.35, 0.01. In this paper, the first year of emission reduction as "good", second year as "ordinary", fourth year as "medium", fifth year as "bad". Given the time weight of emission reduction, that is, the first year of emission reduction weight is $\omega_1=0.9$, the second year of emission reduction weight is $\omega_2=0.88$, the third year of emission reduction is $\omega_3=0.8$, the fourth year of emission reduction is $\omega_4=0.65$, the fifth year of emission reduction is $\omega_5=0.35$. By taking the annual emission reduction y_i as the independent variable, the governing excellent degree S as the dependent variable, we can establish the function equation:

$$S=\omega_1 y_1+\omega_2 y_2+\omega_3 y_3+\omega_4 y_4+\omega_5 y_5 \tag{10}$$

Total capital investment model is:

$$f=\sum_{i=1}^{5} y_i+0.005\sum_{i=1}^{5} y_i^2 \tag{11}$$

Define cost performance:

$$E=\frac{S}{f}=\frac{\sum_{i=1}^{5}\omega_i y_i}{\sum_{i=1}^{5} y_i+0.005\sum_{i=1}^{5} y_i^2}$$

So the governance strategy is transformed into the nonlinear programming model:

$$\max\ E \qquad \text{s.t.}\begin{cases}0<y_1+y_2+y_3+y_4+y_5\le 245\\ y_i>0, i=1,\cdots,5\end{cases} \tag{12}$$

In the model,

$$\omega_1=0.94,\omega_2=0.88,\omega_3=0.8,\ \omega_4=0.65,\omega_5=0.35.$$

We solve the model (12) to obtain the maximum value $E=0.6374$. When E takes the maximum value, we get

$$S=207.7041,\ f=325.8489.$$

Then we can obtain the optimal control scheme.

TABLE IV
PM2.5 OPTIMAL CONTROL SCHEME

Year	Reduction of PM2.5	Comprehensive management fee	Special treatment cost	Total cost
The first year	80.4680	80.4680	32.3755	112.8435
The second year	71.0551	71.0551	25.2241	96.2792
The third year	58.5046	58.5046	17.1139	102.6185
The fourth year	34.9724	34.9724	6.1153	41.0877
The fifth year	0	0	0	0
summary	245.0001	245.0001	80.8288	325.8489

From Table IV, we can see that the annual concentration value decreases are inconsistent. In the first year, the concentration reduces of about 80 $\mu g / m^3$, and it would reduce of about 71 $\mu g / m^3$ in the second year. The total investment cost of up to 245 million yuan for five years. Special projects of investment are decreasing. This environmental governance programs can meet the needs of urgency and feasibility, and also enables the environmental management department to see the results while reducing funding pressures. The program can improve people's degree of satisfaction.

IV Summary and Suggestion

In this paper, firstly we quantitatively analyze the correlation relationship among PM2.5 from SO_2, NO_2, PM10, O_3 and CO, and build a multivariate nonlinear regression equation model of PM2.5 and the five monitoring indexes. Based on statistical data, we get the relationship between PM2.5 and other five monitoring indexes. Then considering the pollution reduction feasible planning of PM2.5, the regional multiple objective optimization model is established under certain assumptions, and the optimal comprehensive PM2.5 control scheme of in the next five years is given.

Because the formation mechanism and process of PM2.5 is more complex [1], [11], this paper only gives five main monitoring indexes that affect PM2.5. Although it can be more obvious that the five factors contribute to the main effect of PM2.5, we need to further explore the impact factors of PM2.5. So we will be able to get more accurate decomposition results of the impact factors of PM2.5.

References

[1] D. H. Xia, B. F. Jiang, Y. L. Xie. Modeling and analysis of PM2.5 generation for key factors identification in China (J). Atmospheric Environment, 2016, (134): 208-216.

[2] W. Q. Wang, Z. D. Niu. VAR model of PM2.5, weather and traffic in Los Angeles long beach area (C). 2009 International Conference on Environmental Science and Information Application Technology, 2009: 66–69.

[3] Y. Zhao, Y.A. Hasan, Machine learning algorithms for predicting roadside fine particulate matter concentration level in Hong Kong Central (J). Comput. Ecol. Softw. 2013, 3 (3): 61–73.

[4] Y. M. Hong, G. F. Zhou. Asymmetries in stock returns: statistical test and economic evaluation (J). Rev. Financ. Stud. 2007, (20): 1547-1581.

[5] S. L. Feng, D. Gao, et al. The healtheffectsofambientPM2.5 and potential mechanisms (J). Ecotoxicology and Environmental Safety. 2016, (128): 67-74.

[6] Eric Cokera, Jokay Ghoshb, et al. Modeling spatial effects of PM2.5 on term low birth weight in Los Angeles County (J). Environmental Research. 2015, (142): 354–364.

[7] Daisuke Goto, Kayo Ueda, et al. Estimation of excess mortality due to long-term exposure to PM2.5 in Japan using a high-resolution model for present and future scenarios (J). Atmospheric Environment. 2016, (140): 320-332.

[8] Harikishan Perugua, Heng Weia, Zhuo Yaoa. Integrated data-driven modeling to estimate PM2.5 pollution from heavy-duty truck transportation activity over metropolitan area (J). Transportation Research Part D: Transport and Environment. 2016, (46): 114–127.

[9] Environmental Protection Bureau of Whuhan city (EB). (2015-01-08). http://www.whepb.gov.cn.

[10] Han Zhonggeng. Mathematical modeling method and its application (M). 2005, Beijing: Beijing Higher Education.

[11] Weekly cycle of magnetic characteristics of the daily PM2.5 and PM2.5–10 in Beijing, China (J). Atmospheric Environment. 2014, (98): 357–367.