Diffusion Test and Prediction of PM$_{2.5}$ Concentration in Urban Street Traffic Microenvironment

Yanshu Ni$^{1, *}$, Lianglin Ma$^{2, b}$

$^1$Department of Electrical Engineering, Suihua University, Suihua, Heilongjiang, China
$^2$College of Intelligent Manufacturing and Automobile, Chongqing Vocational College of Electronic Engineering, Chongqing, China
$^*$email: nys601350111@163.com, $^b$email: 285729976@qq.com

Abstract. A detailed test plan was developed by analyzing the factors affecting PM$_{2.5}$ diffusion in the urban street traffic microenvironment. After field investigation, several test streets within the third Ring Road of Harbin city were selected to investigate and test the hourly traffic flow of the streets, and the geometric structure of the block (road width, building height) was measured. The research carried out tests on the PM$_{2.5}$ concentration, wind speed, temperature and relative humidity of actual street test points. Based on the actual test data, BP artificial neural network and the improved LMBP neural network were used to carry out simulation research on MATLAB platform respectively. A PM$_{2.5}$ concentration prediction model was established to compare and analyze the reliability of the model. Scientific and reasonable prediction of PM$_{2.5}$ pollution in the traffic microenvironment in a specific area.

1. Introduction

The actual monitoring in the field is the most direct and accurate method to study the PM$_{2.5}$ diffusion in automobile exhaust in the micro-environment of street traffic. It is of practical significance to be applied to the streets with a large amount of vehicle emissions on a micro scale. Accurately analyze the emission of strong pollution sources and various influencing factors that affect the diffusion of PM$_{2.5}$ through actual tests. The test site selected in this paper is within the third ring road of Harbin urban area, and many typical streets have been selected. Study the diffusion characteristics of PM$_{2.5}$ in automobile exhaust under the action of various influencing factors. It provides sufficient basic information for the analysis of the diffusion law of PM$_{2.5}$ concentration in automobile exhaust in the street environment. Since PM$_{2.5}$ is a short-term online forecast, the system has high requirements for the real-time prediction results. The BP neural network is used to establish the PM$_{2.5}$ prediction model. The LMBP neural network improves the convergence speed and realizes rapid training, which can meet the real-time requirements of short-term online prediction.

2. PM$_{2.5}$ diffusion test in street environment

2.1. Choice of test street

The basic form of a city street consists of one board, two boards and three boards. Taking the streets within the third Ring road of Harbin city as the research object, the basic information of main streets in the main urban area was statistically analyzed to find out the streets meeting the test requirements.
The information related to pollutant dispersion factors in street environment is mainly considered, including the width of the street, the height of buildings on both sides, the form of the street and the traffic flow, etc. The traffic flow of the street is controlled by traffic lights and traffic volume. Cars work under various working conditions, and the selected street can represent the average condition of various driving conditions in the city[1].

2.2. Test program

2.2.1. Layout of measuring points in street environment.
Each test street is equipped with four measuring points, and the height of the measuring points from the ground is 1.5m (the position of the nose and mouth of the residents of normal height). As many vehicles are parked near the sidewalk on both sides of many streets, the measurement point layout is based on the actual test environment, specifically: measurement point 1 is located at the edge of the road where motor vehicles are parked, measurement point 2 is 1m away from the road edge, and measurement point 3 Point 2 is located 2m away, and measuring point 4 is located 1m away from the building[2]. Measuring point 1 collects PM2.5 concentration data in the urban street environment, that is, the source of PM2.5 emission from motor vehicles; measuring point 2 and measuring point 3 collect PM2.5 concentration data in the sidewalk area, which is the main source of residents Outdoor activity area; measurement point 4 collects PM2.5 concentration data in the indoor ventilation entrance area of the building, which affects the indoor air quality of residents.

2.2.2. Collection of test data.
- Measurement of PM2.5 concentration.
  As PM2.5 concentration and meteorological data vary greatly in the atmosphere, each street is tested three times to reflect the objectivity and authenticity of the test data. Three identical tests were carried out in three time periods. PM2.5 concentrations in street environments are measured using the AEROCET 531 particle counter/dust meter. The instruments were calibrated before the test and the data were recorded manually. During the test, each instrument was fixed at the position of 1.5m to determine the exact position of the four measuring points. At the same time, it was opened and tested for three times to read the effective average value.
- Measurement of meteorological factors.
  Meteorological factors that influence the diffusion of PM2.5 concentration in street environment include atmospheric temperature, relative humidity and wind speed. The measurement of meteorological factors and PM2.5 concentration are carried out simultaneously. AS847 thermometer and AR866 thermosensitive anemometer were used in the experiment.

2.3. Analysis of test data
The test was carried out in accordance with the established test plan, and the cycle test method was adopted for PM2.5 concentration, wind speed, air temperature, air relative humidity and other influencing factors to analyze the test data.

2.3.1. Analysis of PM2.5 concentration test data.
As can be seen from figure 1, the measured PM2.5 concentration fluctuates within the range of 8.2 g/m3–67.4 g/m3. There were 21 variation trends in the decline and decline of PM2.5 concentration at the 4 measurement points k1, k2, k3 and k4. Many factors in urban street environment jointly affect the diffusion of PM2.5, and the diffusion process is very complicated. According to the statistics after the test, there are 38 groups of data with the highest PM2.5 concentration at the measuring point 1, and then it spreads to the other three measuring points in turn. Of these 21 trends, $k_4 < k_3 < k_2 < k_1$ is the largest, indicating that the concentration of PM2.5 in the street environment decreases from the emission source to the direction of buildings in the block in turn. In many street environments where
the tests were conducted, actual wind speeds were slightly lower and there were more parked cars around the roads, making it difficult for fine particles to spread quickly.

\[\text{Figure 1. Statistics of PM}_{2.5}\text{ data in street environment}\]

2.3.2. Estimation of emission intensity and analysis of its influence on PM$_{2.5}$ concentration diffusion.

The relationship between the traffic flow of 26 test streets and the PM$_{2.5}$ concentration at test point 1 during the survey was actually calculated. In general, streets with heavy traffic have relatively high concentrations of PM$_{2.5}$. However, some streets with small traffic still have a high concentration of PM$_{2.5}$. This indicates that the traffic flow has no obvious influence on the PM$_{2.5}$ concentration on the street. Mainly because the average speed of vehicles in the street is not high, traffic jams often occur. There is not much traffic flow per unit time. When the vehicle is driving at low speed, the PM$_{2.5}$ concentration in the exhaust will be higher than that at high speed, and the PM$_{2.5}$ concentration will also be higher in road conditions with low traffic flow.

According to the Measurement Method of Air Pollution Discharged by Urban Motor Vehicles (HJ/T180-2005), the vehicles are classified into small cars represented by cars, medium cars represented by minibuses, and heavy and large cars represented by public buses[3]. According to the limits of each model in The National standard "Fuel Consumption Limits of Light Commercial Vehicles" in China, this paper considers that large and medium-sized vehicles all use diesel oil, and cars are regarded as using gasoline[4]. The traffic flow per hour and the emission intensity of PM$_{2.5}$ were calculated for gasoline and diesel vehicles. Statistics of vehicleflow per hour and PM$_{2.5}$ emission intensity of 26 test streets are shown in figure 2 and figure 3.

\[\text{Figure 2. Vehicle statistics of gasoline and diesel vehicles}\]
\[\text{Figure 3. Measurement of PM}_{2.5}\text{ emission intensity in street}\]

As can be seen from figure 4, the emission intensity of PM$_{2.5}$ fluctuates between 0.03 and 0.30mg/kmꞏh. Its fluctuation range matches the traffic flow per hour. When the traffic volume is large, the emission intensity of PM$_{2.5}$ increases. When the traffic flow is small, the emission intensity of PM$_{2.5}$ decreases.
2.3.3. Analysis of the influence of street width $W$ and building height $H$ on PM$_{2.5}$ concentration diffusion.

The aspect ratio of street canyons is an important factor affecting pollutant dispersion in urban streets. When the street width is fixed, the higher the building is, the more difficult it is to diffuse. The statistics of 26 street widths $W$ and building height $H$ are shown in figure 4. As can be seen from the figure, the number of building floors in the test street varies within the range of 2 to 32 floors, the height of the building fluctuates within the range of 8 to 101m, the width of the street fluctuates within the range of 22 to 65m, and the ratio of height to width of the street valley varies within the range of 0.31 to 2.97, including these three types of block structures.

2.3.4. Analysis of the influence of wind speed, temperature and relative humidity on PM$_{2.5}$ concentration diffusion.

In order to ensure the effectiveness and scientific nature of the test, the wind speed, temperature, relative humidity and PM$_{2.5}$ concentration are measured simultaneously. Each measurement point of each street is tested simultaneously, and statistics are shown in figure 5, figure 6 and figure 7. With the increase of wind speed value in street environment, PM$_{2.5}$ test value decreases slightly, and wind speed is negatively correlated with PM$_{2.5}$ concentration. With the decrease of temperature, the corresponding PM$_{2.5}$ concentration increases, which is roughly in a negative correlation. There is no obvious change rule between relative humidity and PM$_{2.5}$ concentration.
3. **PM$_{2.5}$ concentration prediction in traffic microenvironment**

3.1. **BP neural network model and L-M algorithm**

BP neural network (error back propagation multi-layer feedforward neural network) is a multi-layer network that performs weight training on nonlinear differentiable functions. An algorithm with strong nonlinear mapping capabilities. It consists of forward transmission of information and backward propagation of error. In the forward pass, the input information is passed from the input layer to the output layer through the hidden layer calculation layer by layer. The state of each layer of neurons only affects the state of the next layer of neurons. If the output layer does not get the desired output, the error change value of the output layer is calculated. Then it propagates back, and the error signal is transmitted back along the original connection path through the network, and the weights of the neurons in each layer are modified until the desired goal is reached. Figure 8 shows a three-layer BP neural network with a hidden layer. The learning rule is error correction, back propagation[5].

![Figure 8. Three-layer BP Neural Network](image)

The Levenberg-Marquardt algorithm (L-M algorithm for short) is a Gauss-Newton method improved by standard numerical optimization techniques. Suppose the column vector $x = (x_1, x_2, \cdots, x_n)$.

Using Newton's method as the learning rule of the network model, then:

$$
H_k^{-1}x_{k+1} = x_k - H_k^{-1}g_k
$$

(1)

$H_k^{-1}$ is the inverse of the Hessian matrix, and $g_k$ is the gradient.

Suppose the error function is $e(x)$, representing the column vector, and the length is $N$; the performance function is the error sum of squares function $F(x)$, then:

$$
F(x) = \sum_{i=1}^{N} e_i^2(x) = e^T(x)e(x)
$$

(2)

The gradient of $F(x)$ is:

$$
\nabla F(x) = 2J^T(x)e(x)
$$

(3)

Among them, the matrix $J(x)$ of $n \times n$ is the Jacobian matrix. Each row of the matrix is the $i$-th component of vector $e(x)$, the partial derivative of all elements of $x_i$, where $1 \leq i \leq N$:

$$
J(x) = \begin{bmatrix}
\frac{\partial e_1(x)}{\partial x_1} & \frac{\partial e_1(x)}{\partial x_2} & \cdots & \frac{\partial e_1(x)}{\partial x_n} \\
\frac{\partial e_2(x)}{\partial x_1} & \frac{\partial e_2(x)}{\partial x_2} & \cdots & \frac{\partial e_2(x)}{\partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial e_n(x)}{\partial x_1} & \frac{\partial e_n(x)}{\partial x_2} & \cdots & \frac{\partial e_n(x)}{\partial x_n}
\end{bmatrix}
$$

(4)

Based on this, the approximate expression of the Hessian matrix can be further obtained:

$$
\nabla^2 F(x) = 2J^T(x)J(x)
$$

(5)

It can be seen from the formula that $\nabla^2 F(x)$ may be positive semi-definite, which means that matrix $H = 2J^T(x)J(x)$ may be irreversible. Therefore, the Gauss-Newton method is slightly changed. Add a small increment $\mu I$ to the $n \times n$ matrix $H$ to make it positive definite, where $\mu > 0$ and $I$ are the $n \times n$ identity matrix, namely:
\[ M = H + \mu l \]  

The calculation expression of the deformed Gauss-Newton method can be obtained:

\[ x_{k+1} = x_k - \left[ J^T(x_k)J(x_k) + \mu I \right]^{-1} J^T(x_k)e(x_k) \]  

During the training process, as the iteration continues, \( \mu \) also keeps decreasing. When \( \lim_{\mu \to 0} \Delta x = -\left[ J^T(x)J(x) \right]^{-1} J^T(x)e(x) \) is the Gauss-Newton method, because of its second-order convergence speed, the closer the error is to the minimum, the faster the calculation speed and the higher the accuracy\[6\].

3.2. Data preprocessing

According to the measured typical streets in Harbin urban area, four measuring points are arranged in each street. The data obtained after three cycles of testing is used as the training sample of the network. Due to the large and scattered training data, in order to improve the training efficiency of the neural network, a good weight matrix is obtained and the data is normalized. To make it fall between \([-1,1]\), the method is as follows:

\[ x = 2 \left( \frac{x_0 - x_{\min}}{x_{\max} - x_{\min}} \right) - 1 \]  

In the formula, \( x \) is the normalized variable value; \( x_0 \) is the original variable value; \( x_{\max} \) and \( x_{\min} \) represent the maximum and minimum values of the original variable.

3.3. Network structure

Five influencing factors including emission intensity, H/W, wind speed, temperature, and relative humidity are selected as the five input variables of the training sample. The PM2.5 concentration is used as the output variable of the training sample. The number of neurons in the input layer is 5, and the number of neurons in the output layer is 1. The neuron transfer function of the middle layer of the grid adopts the tansig function, and the neuron transfer function of the output layer adopts the pure linear pureline function. The final selected mesh model is shown in figure 9.

3.4. Result analysis

After training, the neural network can reflect the PM2.5 concentration under five factors of emission intensity, H/W, wind speed, temperature, and relative humidity in the actual test traffic microenvironment. In order to observe the effect of grid prediction, simulations were carried out using the standard BP algorithm and the neural network improved by the LM algorithm. Select 50 sets of data as samples, and compare the predicted value and actual value of PM2.5 concentration under two network prediction models. It can be seen from the PM2.5 prediction data values in table 1 that the LMBP neural network model has higher prediction accuracy than the BP network. It can also be seen from figure 10 that although the BP network forecast results reflect a certain development trend of the forecast data, its forecast fluctuates greatly and large errors occur. It can be seen from figure 11 that the LMBP network model simulates the predicted data very well, except for individual actual data will produce measurement errors. The error percentage of the predicted PM2.5 concentration is shown in figure 12, and most of them are within 5%. The percentage of errors in the few sets of data in the figure exceeds 10%, which is caused by errors in the actual test itself. It can be seen that the LMBP
neural network can accurately realize the PM2.5 concentration prediction in the traffic microenvironment under the action of various influencing factors.

Table 1. The Prediction Results of BP Algorithm and LMBP Algorithm for Partial Data

| Serial number | Actual PM2.5 concentration | BP forecast | BP relative error | LMBP forecast | LMBP relative error |
|---------------|----------------------------|-------------|-------------------|---------------|-------------------|
| 1             | 64.1                       | 39.076      | -39.03%           | 64.088        | -0.018%           |
| 2             | 45.1                       | 40.450      | -10.31%           | 44.976        | -0.273%           |
| 3             | 36.8                       | 38.815      | 5.48%             | 36.717        | -0.222%           |
| 4             | 36.7                       | 28.311      | -22.85%           | 36.750        | 0.138%            |
| 5             | 28.2                       | 20.097      | -28.73%           | 28.286        | 0.305%            |
| 6             | 30.5                       | 23.921      | -21.56%           | 30.484        | -0.051%           |
| 7             | 30.8                       | 26.659      | -13.44%           | 30.675        | -0.403%           |
| 8             | 19.4                       | 22.777      | -10.91%           | 19.499        | -0.581%           |
| 9             | 29.5                       | 25.872      | 17.41%            | 29.489        | 0.051%            |
| 10            | 19.8                       | 25.131      | -12.29%           | 19.903        | -0.037%           |

4. Conclusions
The PM2.5 concentration test results in the traffic micro-environment of urban blocks show that factors such as traffic flow, block geometry, wind speed, temperature and relative humidity have an impact on the diffusion of PM2.5 concentration in the traffic micro-environment. When the PM2.5
concentration diffuses, wind speed has the greatest impact on it; in areas where pedestrian activities are concentrated, H/W and emission intensity have the strongest impact on PM2.5. Combining the measured data, the reasonable flow of traffic and the improvement of the traffic micro-environment will help reduce the pollution of automobile exhaust PM2.5. Use BP neural network and LMBP neural network to predict the PM2.5 concentration respectively, select five influencing factors including emission intensity, H/W, wind speed, temperature, and relative humidity as the five input variables of the training sample, and PM2.5 concentration as the output variable of the training sample. After comparative analysis, the LMBP neural network can accurately and quickly realize the PM2.5 concentration prediction under the action of multiple influencing factors.

The experiments and analyses conducted in this paper have limitations in both depth and breadth, and a lot of experiments and in-depth research are needed. Research on the correlation between automobile exhaust PM2.5 concentration and vehicle operating conditions and other performance and use factors, and correlation analysis between CO, NO2, SO2 and other gaseous pollutants and the formation of secondary particles. It can provide a better basis for traffic pollution control measures and atmospheric environmental governance.

References
[1] Cai, J. (2005) Research on the law of traffic accidents and the planning and design of urban road systems. Chinese Safety Science Journal, 15(4): 16–20
[2] Xing H., Xu W.J., Cai M. (2011) Three-dimensional numerical simulation of pollutant diffusion in street canyons with signal-controlled pedestrian crossings. Journal of Environmental Science, 31(10): 2102–2108
[3] Measurement method of air pollutants emitted by urban motor vehicles (HJ/T180-2005)
[4] Fuel consumption limit of light commercial vehicles (GB.20997-2007)
[5] Pei Y.L., Wang X.N. (2004) Traffic impact prediction model based on BP neural network. Journal of Harbin Institute of Technology, 36(8): 1034–1037
[6] Shi B.H., Zhu X.F. (2008) Research on improved algorithm of LMBP neural network. Control Engineering, 15(2): 164–167