PREDICTION OF PM$_{2.5}$ CONCENTRATION OF BP NEURAL NETWORK BASED ON IMPROVED PARTICLE GROUP ALGORITHM

Wei Zhang [1, 2], Shaofeng Xie [1, 2, *], Zhihao Zhou [1, 2], Yun Zhao [1, 2], Guohong Li [1, 2]

[1] School of Surveying and Mapping Geography, Guilin University of Technology, China - (463879463, 1312163074, 3269717468, 417249652, 935331685)@qq.com
[2] Guangxi Key Laboratory of Spatial Information and Surveying, Gui Lin, China

Commission VI, WG VI/4

KEY WORDS: Improved Particle Swarm Optimization Algorithm; BP Neural Network; PM$_{2.5}$ Concentration Prediction; Algorithm Optimization

ABSTRACT:

There exists the shortage of low accuracy when using BP neural network model to predict PM$_{2.5}$ concentration in air. An improved particle swarm optimization (IPSO) algorithm combined with BP neural network was proposed. Using the advantages of improved PSO algorithm global optimization ability, the weight and threshold of BP neural network are optimized, pollutant data and meteorological data are used as input data, PM$_{2.5}$ concentration is used as output data, and IPSO-BP model is established for simulation prediction. Comparing and analyzing the IPSO-BP model, PSO-BP model and BP model, the results show that the MAE and RMSE of the IPSO-BP model are 6.94 and 8.47, respectively, and the R^2 is 0.77. The accuracy test indicators are better than the PSO-BP model, and the BP model, PM$_{2.5}$ concentration has the highest prediction accuracy, validating the validity of the model's prediction of PM$_{2.5}$ concentration.

1. INTRODUCTION

Massive smog attacks have seriously affected people's traffic and harmed people's health. PM$_{2.5}$, a particle having a diameter of less than or equal to 2.5 μm, is the main component of bismuth. According to SCHWARTLANDER, representative of the World Health Organization in China, the daily mortality rate of cardiopulmonary related diseases increased by 1.5% for every 10$\mu g \cdot m^{-3}$ increase in daily mean concentration of PM$_{2.5}$; studies have shown that when PM$_{2.5}$ concentration is greater than 115$\mu g \cdot m^{-3}$ Sensitive people such as the elderly and children will feel severe discomfort (Sorensen, 2003). Therefore, the efficient and accurate PM$_{2.5}$ concentration prediction has important guiding significance for air pollution prevention and control.

At present, air quality prediction methods are divided into physical chemical mechanism model and machine learning (Hou, 2018). The mechanism model simulates the physical and chemical processes of pollutants. Machine learning models can simplify the forecasting process by using specific algorithms to find hidden transformation rules from the data according to statistical principles. Machine learning-based predictive models include multiple linear regression (Elbayoumi, 2014), time series analysis (Chelani, 2007), gray model (Mao Lei, 2014), support vector machine (SVM) (Zuo, 2018), Bayesian (Balachandran, 2013) and other traditional methods and neural networks. Algorithm (Grivas, 2006) is the dominant artificial intelligence method.

BP neural networks are widely used in many fields, but there are some limitations that cannot be overcome, such as the disadvantages of relying on initial weights and thresholds, and slow convergence. When BP neural network is used for nonlinear mapping, it will often fall into a local minimum, resulting in low prediction accuracy. In view of the above problems, Zhang (2019) used the particle swarm optimization idea to introduce the crossover and mutation operations of the genetic algorithm in the optimization process, and designed an improved PSO-GA hybrid algorithm to set the BP initial weight and threshold. Effectively avoid falling into local minima and improve convergence speed. The simulation results show that both the improved PSO-GA-BP prediction model and the PSO-BP prediction model can obtain good prediction results. Shi (2017) proposed the gray cloud PM$_{2.5}$ concentration prediction of gray wolf intelligence optimization. The research shows that the neural network optimized by gray wolf algorithm is suitable for PM$_{2.5}$ concentration prediction and air quality prediction. Yang (2016) proposed a genetic algorithm optimization BP neural network PM$_{2.5}$ concentration prediction method, which has good learning and generalization ability, but did not take into account the influence of meteorological factors on PM$_{2.5}$ concentration, so the prediction accuracy Lower. Zhang (2017) proposed a PM$_{2.5}$ prediction model based on attribute reduction and BP neural network, which can be used to predict particle concentration and pollution situation, optimize BP neural network structure, and improve the prediction accuracy of prediction model. Ma (2014) used the improved PSO to optimize the fuzzy neural network. The algorithm improved the slow convergence speed of the conventional neural network algorithm, and also overcome the shortcomings of the conventional PSO algorithm which is easy to fall into the local optimum, and the prediction accuracy is improved.

---

* Corresponding author

Therefore, in order to further improve the prediction accuracy of $PM_{2.5}$ concentration, this paper proposes a method based on improved particle swarm optimization algorithm to optimize BP neural network. The air quality data and meteorological data from January to October 2018 in Wuhan are selected as research objects to construct $PM_{2.5}$ concentration prediction model was experimentally verified.

2. IPSO-BP NEURAL NETWORK PREDICTION MODEL

2.1 Principle of Standard Particle Swarm Algorithm

The PSO algorithm is derived from the study of bird predation behavior. When birds prey, the easiest and most effective way for each bird to find food is to search for the area around the bird that is closest to the food. The PSO algorithm is inspired by this biological population behavioral feature and used to solve the optimization problem. Each particle in the algorithm represents a potential solution to the problem, and each particle corresponds to a fitness value determined by the fitness function. The velocity of the particle determines the direction and distance of the particle's movement. The velocity is dynamically adjusted with the movement experience of itself and other particles, thus achieving the individual's optimization in the solvable space.

The PSO algorithm can optimize the global and has strong parameter global search ability, which is widely used in the field of function optimization (Zhou, 2013). The specific principle is to assume that a population consists of n particles, denoted as $X=（X_1, X_2, …, Xn）$, each particle is in a D-dimensional vector space, and the position of the i-th particle in space can be Expressed as $X_i=（X_{i1}, Xi2, …, X_{iD}）$, the optimal position of the i-th particle is represented as $Pbest_i=（P_{i1}, P_{i2}, …, P_{im}）$, and the corresponding particle velocity is $V_i=[Vi1, Vi2, …, ViD]$, the optimal position in the population is expressed as $Pg=[Pg1, Pg2, …, PgD]$.

After finding the optimal position and global optimal position of the individual, the particle updates its speed and position accordingly, as shown in equations (1) and (2).

$$V_{id}^{k+1}=\omega V_{id}^{k}+c1r1(P_{id}^{k}-X_{id}^{k})+c2r2(P_{gd}^{k}-X_{id}^{k}) \tag{1}$$

$$X_{id}^{k+1}=X_{id}^{k}+V_{id}^{k+1} \tag{2}$$

where
$l = 1, 2,…, n$
$d = 1, 2,…, D$
k = the evolution number of the current algorithm
$c1, c2$ = acceleration factors ($c1$ is used to adjust the particles to obtain the optimal position, $c2$ is used to adjust the coefficient of the particle to obtain the optimal position of the population, $c1$ and $c2$ are generally set to the same value)
$r1, r2$ = random numbers of the interval [0,1]
ω = the inertia weight

In order to avoid blind search of particles or to detach from the population space during evolution, the speed and position of the particles are generally limited to a certain range.

2.2 Improved Particle Swarm Optimization

In this paper, the improvement of particle swarm optimization algorithm has two parts, one is the optimization of inertia weight, and the other part is to optimize the learning factor.

The larger the inertia weight ω, the better the global search, and the smaller the ω, the better the accurate local search and the optimal solution. Therefore, the use of varying inertia weights can effectively avoid falling into local optimal solutions. In order to better balance the global and local search ability of the algorithm and avoid the oscillation near the premature and global optimal solutions, the weighted linear decreasing PSO algorithm is used in this paper. The choice of ω is as shown in equation (3).

$$\omega = \omega_{max} - k\frac{\omega_{max} - \omega_{min}}{k_{max}} \tag{3}$$

where
ω_{max} = aximum values of the inertia weight
ω_{min} = minimum values of the inertia weight
k_{max} = maximum number of iterations

Research shows that (Zhou et al., 2013) can greatly improve the performance of PSO optimization algorithm when the inertia weight decreases linearly from 0.9 to 0.4. Therefore, the inertia weight in the algorithm can be expressed as equation (4).

$$\omega = 0.9 - \frac{0.5 \times k}{k_{max}} \tag{4}$$

The learning factors $c1$, $c2$ are closely related to the cognitive state of the particles. From equation (1), it can be seen that $c1$ determines the individual cognitive level contribution rate of the particles, and $c2$ determines the horizontal contribution rate of the particle population. Therefore, in the early iteration period, the particle fitness is large, and the larger $c2$ can be used to control the particle to develop in the optimal direction of the group; in the later iteration, the particle fitness is gradually reduced, and the individual $c1$ can be released through the larger $c1$ until find the best location. In view of this, this paper updates the learning factor according to equations (5) and (6):

$$c1 = c1_{max} - (c1_{max} - c1_{min})(cos(\frac{\pi k}{2k_{max}}))^2 \tag{5}$$

$$c2 = c2_{min} + (c2_{max} - c2_{min})(cos(\frac{\pi k}{2k_{max}}))^2 \tag{6}$$

where
$c1_{max} = c2_{max}$ =maximum values of the learning factor
$c1_{min} = c2_{min}$ =minimum values of the learning factor

The improvement of the particle swarm optimization algorithm by adjusting the inertia weight coefficient and the learning factor not only reduces the number of iterations of the algorithm, but also reduces the computation time. Moreover, the improved model can effectively avoid the problem that the particle swarm

falls into the local optimum and improve the prediction accuracy of the particle swarm optimization algorithm.

The specific steps of using the IPSO algorithm to optimize BP neural network to predict PM$_{2.5}$ concentration are as follows:

1. Determine the network topology based on the input and output values of the data, and initialize the weights and thresholds of the BP neural network.
2. Set the number of iterations, the learning factor, and the size of the population. The velocity and position of the initial particles are assigned to random values, and the range of values is within a defined interval.
3. Initialize the particle swarm to construct the mapping between particle swarms and neural network weights and thresholds.
4. The network model is obtained by training the training sample network, and the reciprocal of the sum of squared errors between the actual output of the network and the expected output is used as the fitness function. The calculated particle fitness value is compared with the

optimal fitness value, and the individual and global optimal fitness values are continuously updated.
5. According to formula (1) and formula (2), after updating the particle velocity and position, the particles are reinitialized with a set probability, and new individual and group extremum are set compared with the fitness value of the new particle.
6. After iterating to the maximum number of times, the optimal particle obtained by the algorithm initializes the BP neural network connection weight and threshold.
7. After training the network prediction model, the optimal solution for PM$_{2.5}$ concentration prediction is output.

2.3 Technical Process

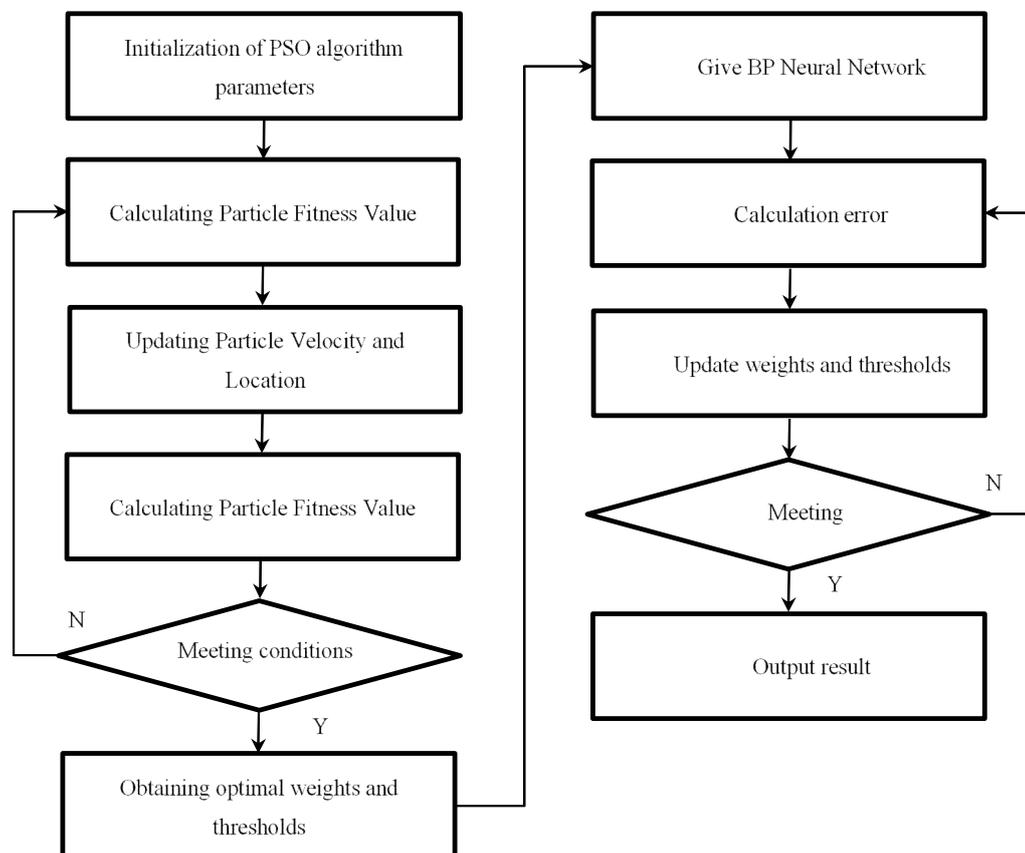The flow chart of the IPSO-BP model is shown in Figure (1):

Figure 1. Flow chart of IPSO-BP model

2.4 Accuracy Assessment

The evaluation criteria of the three models use mean average error (MAE), root mean square error (RMSE) and correlation coefficient (R^2) as shown in equations (7), (8) and (9) Show:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - x_i^{'}\right| \qquad (7)$$

$$RMSE = [\frac{1}{n}\sum_{i=1}^{n}(x_i - x_i^{'})]^{1/2} \qquad (8)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - x_i^{'})^2}{\sum_{i=1}^{n}(\overline{x} - x_i^{'})^2} \qquad (9)$$

where n = the total number of samples
          x_i = the true value of PM$_{2.5}$ concentration at time i

$x_i{'}$ = the model prediction value at the same time

$\overline{x}$ = the average value of the model prediction output

MAE reflects the actual situation of the predicted value error. The smaller the value, the smaller the error. The RMSE reflects the stability of the model's predicted output value. The smaller the value, the higher the stability. The R^2 reflects the actual PM$_{2.5}$ concentration and the model predictive output. The degree of value association, the closer its value is to 1, the better the performance.

3. EXPERIMENTAL RESULTS AND ANALYSIS

In order to ensure the authenticity of the experiment, the data of this paper is downloaded from the China Air Quality Online Testing and Analysis Platform. The data includes air quality data from January 1 to October 31, 2018 in Wuhan: PM$_{10}$、SO$_2$、NO$_2$、CO、O$_3$_8h，CO concentration unit is mg/m^3, and PM$_{10}$、SO$_2$、NO$_2$、O$_3$ concentration units are $\mu g/m^3$. Meteorological data: the highest temperature, the lowest temperature and the rainfall every day, the temperature is $°C$, and the unit of rainfall is mm. The data from January 1 to September 30 is selected as the training data, and the data from October 1 to October 31 is output. Experiments in the MATLAB 2016a environment, three neural network prediction models were constructed using the neural network toolbox.

3.1 Model Parameter Determination

The number of hidden layer nodes in BP neural network has a great influence on the prediction accuracy of BP neural network. The number of nodes is too small, the network cannot learn well, the training times need to be increased, the training precision is also affected, the number of nodes is too much, the training time Increase, the network is easy to overfit. The determination of the number of hidden layer nodes in this paper uses the trial and error method. First, the empirical formula is as shown in equation (8).

$$l = \sqrt{m+n} + a \qquad (8)$$

where
l = the number of hidden layers
m = the number of output layer nodes
n = the number of input layer nodes
a = a constant between 0 and 10

After determining the approximate range, the forecasting result is compared with the actual value by continuously adjusting to determine the optimal number of nodes. Since the weight and threshold of each operation of the BP neural network are random, in order to determine the optimal number of hidden layer nodes, the experiment uses MAE as the evaluation standard, and each node number runs independently 10 times, and the average error of 10 times is taken average. It can be seen from Table 1 that n=3 is the optimal number of nodes of the model hidden layer. Therefore, the network structure of the BP neural network in this paper is 8-3-1.
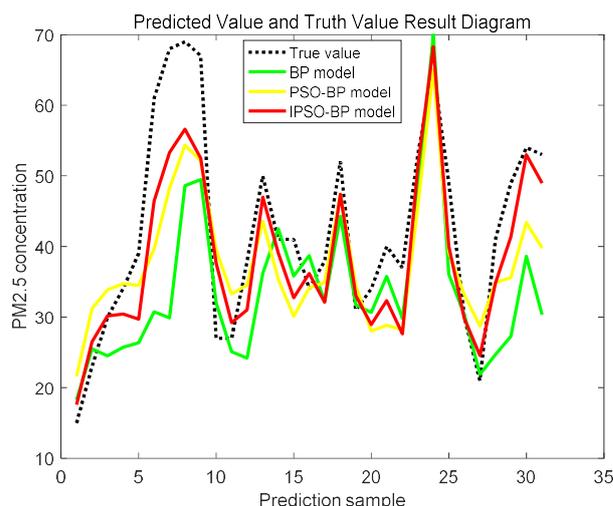
n	3	4	5	6	7	8	9	10	11	12	13
MAE	8.03	8.30	8.73	8.50	8.71	9.49	10.00	8.70	8.40	9.72	10.08

Table 1 The relation between the number of hidden layer nodes and MAE

3.2 Experimental Result

The PM$_{2.5}$ concentration values predicted by the three models are consistent with the changes of the true values as shown in Figure 2, but the predicted trend of the IPSO-BP model is more in line with the true value. The prediction trend of the PSO-BP model is slightly worse than the IPSO-BP model, but better than the BP model. The prediction trend of the BP model is consistent with the change of the true value, but the effect is not as good as the IPSO-BP model and the PSO-BP model.

The residual diagrams of the three models are shown in Fig. 3. It can be seen from the figure that the residual of the IPSO-BP model swings up and down at 0 and the amplitude is small and stable. The residual of the PSO-BP model is not much different from the residual of the PSO-BP model, but significantly smaller than the BP model. The residual amplitude of the BP model is the largest and significantly larger than the other two models.

Figure 2. Predicted value and truth value result diagram
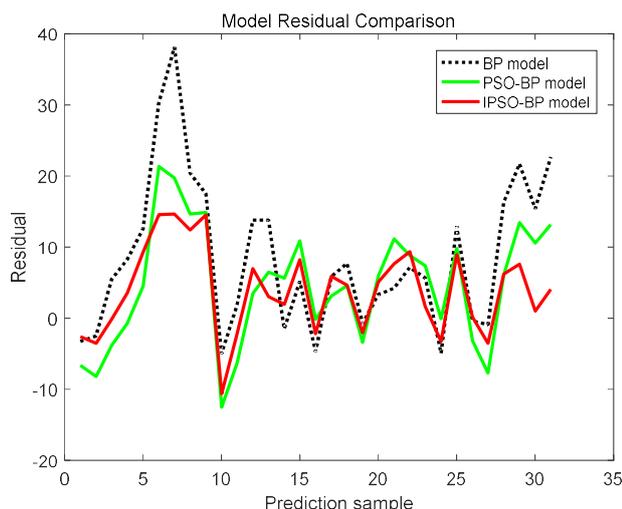
Figure 3. Model residual comparison

3.3 Accuracy Evaluation

The accuracy evaluation data of the three models is shown in Table 2. It can be seen from the table that the BP model can be used to predict the PM$_{2.5}$ concentration, but the prediction accuracy is not high. The prediction accuracy has been improved by the particle swarm optimization algorithm. The MAE and RMSE of the model are 6.94 and 8.47, respectively, which are smaller than the PSO-BP model and the BP model. The R^2 of the model is 0.77, which is larger than the PSO-BP model and the BP model. It is proved that IPSO-BP can improve the prediction accuracy of PM$_{2.5}$ concentration.

MODELS	MAE	RMSE	R^2
BP	8.65	10.26	0.61
PSO-BP	7.69	8.89	0.64
IPSO-BP	6.94	8.47	0.77

Table 2. Comparison of performance indicators of three models

4. CONCLUSION

In this study, the improved PSO algorithm is combined with BP neural network, and 8 factors including PM$_{10}$、SO$_2$、NO$_2$、CO 、O$_3$_8h, daily maximum temperature, minimum temperature and rainfall are established as input parameters, and PM$_{2.5}$ concentration is used as output. Parameter-based particle swarm optimization algorithm for BP neural network model. Taking the Wuhan City from January to October 2018 as the research object, the prediction accuracy of the model is verified. The experimental results show that the MAE of the IPSO-BP model is 6.94, the RMSE is 8.47, and the R^2 value is 0.77. The evaluation indexes are better than the PSO-BP model and the BP model, which proves that the proposed IPSO-BP model can effectively improve PM$_{2.5}$ prediction accuracy.

REFERENCES

Balachandran, S., Chang, H., Pachon, J., Holmes, H. A., Mulholland J., Russell, A., 2013. Bayesian-Based Ensemble Source Apportionment of PM$_{2.5}$. *Environmental science & technology*, 47(23), 13511-13518.

Chelani, A. B., Devotta, S., 2007. Prediction of ambient carbon monoxide concentration using nonlinear time series analysis technique. *Transportation Research Part D*, 12(8).

Elbayoumi, M., Ramli N. A., Yusof, N. F. F. N., Yahaya, A. S. B., Madhoun, W. A., Ul-Saufie, A.Z., 2014. Multivariate methods for indoor PM$_{10}$ and PM$_{2.5}$ modelling in naturally ventilated schools' buildings. *Atmospheric Environment*, 94.

Grivas, A. C., 2005. Artificial neural network models for prediction of PM$_{10}$ hourly concentrations, in the Greater Area of Athens, Greece. *Atmospheric Environment*, 40(7).

Hou, J. X., Li, Q., Zhu, Y. J., Feng, W., Lin, S. F., 2018. The PM$_{2.5}$ forecasting method of machine learning and WRF atmospheric model. *Science and Mapping Science*, 43(02), 114-120+141.

Ma, T. C., Liu, D., Li, X. J., 2014. Improved particle swarm optimization based fuzzy neural network for PM$_{2.5}$ concentration prediction. *Computer Engineering and Design*，35（9），3258-3262.

Mao, W., Sun, Y., Feng, W., Feng, S. X., 2014. Grey Prediction and Correlation Factors Analysis of PM$_{2.5}$ Concentration in Air. *Journal of Ningxia University (Natural Science Edition)*, 35(03), 283- 288.

Shi, F., Lou, W. G., Zhang B., 2017. Predicting the concentration of PM$_{2.5}$ in neural network based on intelligent optimization of grey wolves. *Journal of Computer Applications*, 37(10), 2854-2860.

Sorensen, M., Daneshvar, B., Hansen, M., Dragsted, L.O., Hertel, O., Knudsen L., Loft S., 2003. Personal PM$_{2.5}$ exposure and markers of oxidative stress in blood. *Environmental Health Perspectives*, 111(2).

Yang, Y., Fu, Y. L., 2016. Research on prediction mass concentration of PM$_{2.5}$. *Computer Simulation*, 33（3），413-418

Zhang, X., Du, J. L., 2019. Improved $PM_{2.5}$ Concentration Prediction of PSO-GA-BP. *Computer Engineering and Design*, 40(06), 1718-1723

Zhang, Y., Li, Y., Quan, Q. M., 2017. Prediction model of $PM_{2.5}$ based on attribute reduction and BP neural net-work. *Environmental Science&.Technology*，40（S1）, 341-346.

Zhou, J., Chen, J. H., Liu, G. X., 2013. Summary on inertia weight in particle swarm optimization algorithm. *Guangdong Electric Power*，26（7）, 6-12.

Zuo, J. J., Ni, Z. W., Zhu, X. H., Li J. M., Wu Z. J., 2018. Fusion of Synergistic Artificial Fish School Algorithm and SVM for Haze Prediction Method. *Pattern Recognition and Artificial Intelligence*, 31(08), 725- 739.