Research on Indoor Air Quality Evaluation Based on Improved BP Neural Network

Yong Wang, Yuanhui Cui

Dalian Polytechnic University, Dalian, China

Abstract: In order to make a scientific evaluation of indoor air quality, according to GB/T18883-2002, five representative indoor environmental pollutants CO_2, CO, TVOC, formaldehyde and particulate matter are selected to build a standard evaluation table, and a scientific evaluation model was established based on BP neural network and variable learning rate momentum method. Through MATLAB simulation, and compared with the three BP neural network models mentioned in the literature, the convergence speed of the momentum BP neural network model with variable learning rate is increased by about 51%, and the evaluation accuracy is up to 100%, which shows that the model can evaluate indoor air quality accurately and reliably.

Keywords: indoor air quality, evaluation model, momentum method with variable learning rate, BP neural network

I. Introduction

With significant changes in contemporary lifestyle and working conditions, people spend 90% time indoors on average. Recent years see increased indoor pollutant contents due to indoor decoration or poor ventilation conditions, which will definitely cause persistent damage to the human body [1-2]. Study on indoor air quality has a long history. As early as 2002, the World Health Organization (WHO) proposed five major environmental factors that harm humanity, one of which is quality of indoor air environment [3]. In India, indoor air quality assessment has been made in and around urban slums based on carbon dioxide, carbon monoxide, sulfur dioxide, nitrogen dioxide and suspended particulate matter [4]. Another indoor air quality assessment in a Brazilian primary school was made based on xylene, nitrogen dioxide, sulfur dioxide, ozone, acetic acid, and formic acid isomers [5]. At the same time, some domestic scholars evaluate indoor air quality based on concentration values of indoor carbon dioxide, carbon monoxide, hydrogen sulfide, and air humidity [6]. Moreover, there are many books [7], reports [8] and standards [9] about indoor air quality, such as China's GB/T18883-2002 "Indoor Air Quality Standards".

Indoor air quality evaluation methods are roughly divided into subjective and objective ones. Where, subjective evaluation method is to judge the quality of indoor air environment via subjective feelings of the human body, but this method is greatly susceptible to subjective consciousness. Moreover, some colorless and odorless harmful gases in the air are beyond perception of subjective consciousness. Hence, this method has great limitations. Objective evaluation method mainly takes pollutants with relatively great impact on the human body as evaluation factors. So far, such methods include fuzzy comprehensive evaluation method [10], analytic hierarchy process [11], and gray correlation method [12]. The method can fairly reflect indoor air quality, but lacks subjective feelings. As a result, both evaluation methods have respective advantages and unavoidable disadvantages. As technology progresses continuously, many complex problems are handed over to neural network. Based on BP neural network, this paper proposes to optimize the network by adding momentum and adaptively adjusting the network learning rate, so that more accurate and reliable comprehensive evaluation result is possible.

II. Determination of Indoor Air Quality Parameters

CO_2 content in the air accounts for 0.03% ~ 0.04%, and its concentration can well reflect the effect of indoor ventilation, so it can be used as one pollutant for indoor air quality assessment. Indoor CO derived from combustion products of fuel gas has strong irritation and will affect human metabolism, growth and development, even causing death in the case of excessive content. TVOC is a gas causing the most serious impact on indoor air quality. Mainly from furniture and decoration materials, it will produce irritating odors when the content exceeds 0.3mg/m^3, making the inhaler develop discomforts like headaches, loss of appetite, etc. Indoor formaldehyde mainly from wall coatings and paints will seriously damage people's respiratory system and immunity, thus playing a very important role in deciding indoor air quality. At the same time, inhalable particulate matter also greatly impacts indoor air quality. Mainly from indoor smoking, cooking fume and dust at cleaning, it can cause headaches, fatigue, etc. in the case of excessive concentration. Accordingly, five representative pollutants: CO_2, CO, TVOC, formaldehyde and particulate matter are selected as evaluation factors for indoor air quality.

III. Design of BP Neural Network Model

A. BP Neural Network Structure

In this paper, indoor air quality is assessed based on three-layer BP neural network model. The network topology is shown in Fig. 1. The node number of network input layer, hidden layer, and output layer is set as *n*, *p* and *q*, respectively. Where, the network input layer has *n*=5 nodes, corresponding to 5 indoor air environmental pollutants; the output layer has *q*=1 node, corresponding to the output result of indoor air quality evaluation; w_{ih}, w_{ho} are the weights from the *i*-th input node of the input layer to the *h*-th node of the hidden layer, and that from the *h*-th node of the hidden layer to the *o*-th node of the output layer, respectively.

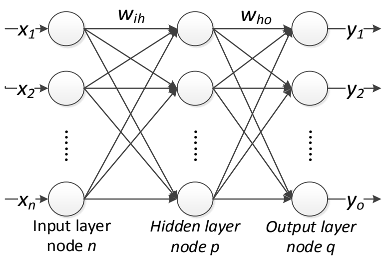

Figure 1. Topological structure of three-layer BP neural network.

B. Momentum BP Neural Network With Variable Learning Rate

The traditional BP neural network modifies the weights of the network output layer and the hidden layer using the steepest descent algorithm based on the principle of error back propagation. The modification is as follows:

Calculate error function *E* of the network based on the system's training output *yo* and the expected output *do*:

$$E = \frac{1}{2}\sum_{o=1}^{q}(d_o(k) - yo_o(k))^2 \tag{1}$$

Where: *E* is the network training error; *o* is the number of output layer nodes; *k* is the number of input samples; *do(k)* is the expected output of the *k*-th sample; *yo(k)* is the actual output of the *k*-th sample.

According to the error function *E*, sequentially modify the output and hidden layer weight variations:

$$\Delta w_{ho}(k) = -\alpha\frac{\partial E}{\partial w_{ho}} = -\alpha\frac{\partial E}{\partial yi_o}\frac{\partial yi_o}{\partial w_{ho}} = -\alpha(d_o(k) - yo_o(k))yo_o^{'}(k)ho_h(k) \tag{2}$$

$$\Delta w_{ih}(k) = -\alpha\frac{\partial E}{\partial w_{ih}} = -\alpha\frac{\partial E}{\partial hi_h(k)}\frac{\partial hi_h(k)}{\partial w_{ih}} = -\alpha\sum_{o=1}^{q}((d_o(k) - yo_o(k))yo_o^{'}(k)w_{ho})ho_h^{'}(k)x_i(k) \tag{3}$$

Where: $\Delta w_{ho}(k)$, $\Delta w_{ih}(k)$ are the partial derivatives of the error function against the output and the hidden layer weight, respectively; $yo_o(k)$, $ho_o(k)$ are the output and hidden layer outputs, respectively; $x_i(k)$ is the network input; α is the learning rate, $0<\alpha<1$.

The traditional BP neural network based on steepest descent algorithm has a slow convergence rate and shows "greedy search" so that it is prone to local extremum, making optimal result impossible [13]. To effectively avoid these shortcomings, this paper proposes momentum BP algorithm with variable learning rate for the network improvement.

Momentum BP algorithm with variable learning rate is a combination of additional momentum method and adaptive learning rate method based on steepest descent algorithm. Its formula can be expressed as:

$$\Delta w(k+1) = \eta\Delta w(k) + \alpha(k)(1-\eta)\frac{\partial E(k)}{\partial w(k)} \tag{4}$$

$$w(k+1) = w(k) + \Delta w(k+1) \tag{5}$$

Where, η is a momentum factor, $0<\eta<1$, generally at about 0.95; α is the learning rate, $0<\alpha<1$; *k* is training times, $\frac{\partial E(k)}{\partial w(k)}$ is the partial derivative of error function against the weight.

As can be seen from (4) and (5), after momentum term is added to the BP neural network, the network correction concerns both the previous gradient and the previous correction result. At this time, the network weight will move toward low and flat error surface, which helps the network get rid of the local minimum of the error surface. However, the adaptive learning rate can improve the learning rate of BP neural network and speed up the network convergence. Where, the momentum factor η and the adaptive learning rate α are set as follows [14-15].

The modifier formula for momentum factor η is:

$$\eta = \begin{cases} 0, E(k) > 1.04E(k-1) \\ 0.95, E(k) < E(k-1) \\ \eta, otherwise \end{cases} \tag{6}$$

Where: η is a momentum factor, $0<\eta<1$; *k* is training times; *E*(*k*) is error function.

The modifier formula for adaptive learning rate α is:

$$\alpha(k+1) = \begin{cases} 1.05\alpha(k), E(k+1) < E(k) \\ 0.7\alpha(k), E(k+1) > 1.04E(k) \\ \alpha(k), otherwise \end{cases} \tag{7}$$

Where, α is the learning rate, $0<\alpha<1$; *k* is training times; *E*(*k*) is error function.

The flow chart of the momentum BP neural network with variable learning rate is shown in Fig. 2.

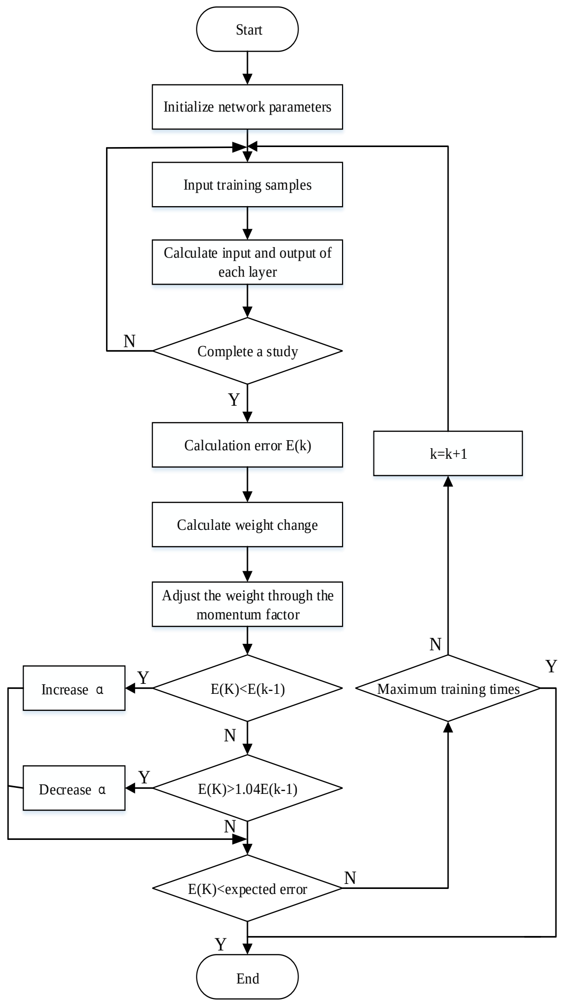

Figure 2. Workflow of improved BP neural network algorithm

IV. Set Up the Standard Evaluation

This study is based on GB/T18883-2002 "Indoor Air Quality Standards" and other relevant national grading standards, etc. Considering the characteristics of IAQ influencing factors, the indoor air quality is assessed according to four levels by judging the five typical environmental pollutants CO_2, CO, TVOC, formaldehyde, and particulate matter, namely no pollution (level I), light pollution (level II), moderate pollution (level III), and severe pollution (level IV), as shown in Table 1.

Table 1. Classification Standard of Indoor Air Quality

Level	CO_2 /(ppm)	CO /(mg·m^{-3})	TVOC /(mg·m^{-3})	Formaldehyde /(mg·m^{-3})	Particulate matter /(ug·m^{-3})
Ⅰ	≤550	≤2.8	≤0.2	≤0.035	≤45
Ⅱ	775	3.85	0.3	0.057	98
Ⅲ	1225	7.25	0.6	0.95	195
Ⅳ	2125	14.5	0.8	1.7	360

V. Simulation of Indoor Air Quality Evaluation Model

A. Establishment of Training Samples and Models

To fully train the BP neural network and achieve good adaptability in the trained evaluation model, the 4 levels in indoor air quality classification standard in Tab. 1 are expanded, followed by even data insertion between each two levels to obtain 500 sets of training samples required by the network. There is no unified reference standard for network input and output, which will greatly slow down the network convergence speed, so the concentration values of the five gases should be normalized to the unified reference standard as the model input, while the evaluation level should be normalized as the model output. Where, the normalization formula is [16]:

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{8}$$

Where, y is the normalized value, x is the original value, x_{max} and x_{min} are the maximum and minimum values in the original data x. To enable good convergence in the neural network, the data is limited to the interval [0,1].

The network output also needs renormalization according to the renormalization formula to obtain the evaluation grade value:

$$x = y*(x_{\max} - x_{\min}) + x_{\min} \tag{9}$$

Where, y is the transformed value, x is the original data, x_{max}, x_{min} are the maximum and minimum values in the original data, respectively.

The number of hidden layer nodes p can be derived according to empirical formula [17]:

$$p = \sqrt{n+q} + a \tag{10}$$

Where, n=5, q=1,a is a constant between [0，10], so p has a value range of 3~13. For the sake of ideal result, first set the minimum number of hidden layer nodes. By gradually increasing the number of hidden layer nodes, it is finally determined that the simulation effect is optimal when the number of hidden layer nodes is 6. Thus, 5-6-1 three-layer BP neural network model structure is built. According to I, II, III, IV grade evaluation standards, the network output is set as follows: When the output range is (0.5, 1.5), the corresponding evaluation result is I, which is indicated by figure 1; at this time, at the critical value 1.5, "false alarms" is a better choice compared to "underreporting" according to the principle of "false alarms and underreporting". Therefore, when the output range is [1.5, 2.5), the corresponding evaluation result is II, which is indicated by figure 2. Similarly, when the output range is [2.5, 3.5), the corresponding evaluation result is III, which is indicated by figure 3; when the output range is [3.5, 4.5), the corresponding evaluation result is IV, which is indicated by figure 4. The maximum number of steps in the network is set to 1000; the training accuracy is set to 0.001; the initial momentum factor is set to 0.95. As the network adopts adaptive learning rate, the learning rate can be adjusted adaptively according to the network error during training. Given that setting of the initial learning rate can ensure stable learning of the network under any value, the initial learning rate is set to 0.01. The hidden layer excitation function takes S-type tansig; the output layer excitation function takes linear puelin.

This paper simulates 10 sets of pollutant concentration data, as shown in Table 2. After the model training is completed, the pollutant concentration data simulated in Tab. 4 is conveyed to the network for verification.

Table 2. Concentration Data of 10 Groups of Pollutants to Be Verified

No.	CO_2 /(ppm)	CO /($mg\cdot m^{-3}$)	TVOC /($mg\cdot m^{-3}$)	Formaldehyde /($mg\cdot m^{-3}$)	Particulate matter /($ug\cdot m^{-3}$)
1	683	1.95	0.17	0.027	41
2	1457	3.58	0.71	0.971	201
3	286	1.57	0.14	0.019	29
5	397	1.74	0.08	0.024	27
6	361	1.35	0.11	0.013	35
7	1018	2.16	0.57	0.032	135
8	519	2.62	0.09	0.029	78
9	816	2.70	0.41	0.034	43
10	417	0.87	0.14	0.007	28

B. Data Fusion Simulation Test and Result Analysis

In this paper, a network model is built using MATLAB2017a software platform, and a model is established for indoor air quality evaluation by momentum BP algorithm with variable learning rate. To prove evaluation performance reliability of the momentum BP model with variable learning rate, comparison is made between momentum BP neural network model with variable learning rate and the traditional BP neural network model, the additional momentum BP neural network model and the BP neural network model with adaptive learning rate. To reflect fairness of the network test results, the same sample data is used for the four models. For the traditional BP neural network model, when the learning rate is excessively set, the network will oscillate, resulting in failure to converge, so the learning rate is set to 0.32 based on verification. For the additional momentum BP neural network model, the learning rate is also set to 0.32, while the initial momentum factor is set to 0.95. For the BP neural network model with adaptive learning rate, its essence is to set the constant learning rate to a variable learning rate based on the traditional BP neural network. At this time, the initial learning rate can be arbitrarily set, so the initial learning rate is set to 0.01.

In comparison of these four models, 500 sets of training data is first sent to the network model for training, and then the training steps of the four models are compared as shown in Table 3.

Table 3. Comparison of Training Steps of Four Models

Evaluation model	Training step
Traditional BP	68
Additional momentum BP	62
BP with Adaptive learning rate	37
Momentum BP with variable learning rate	33

After the model training is completed, the 10 sets of gas concentration data simulated in Tab. 2 are taken into the four models for testing, and the air quality evaluation results of the four models are shown in Fig. 3.

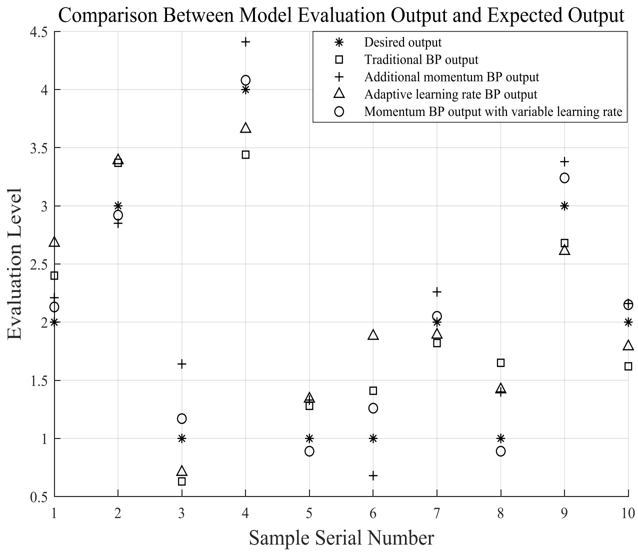

Figure 3. Model Evaluation Results

Based on the evaluation output of the four models shown in Fig. 3, comparison can be made in the actual output levels and errors between the four models as shown in Table 4.

Table 4. Actual Output Level and Error Comparison of Four Models

No.	Expected Output	Traditional BP		Additional Momentum BP		BP with Adaptive Learning Rate		Momentum BP with variable learning rate	
		Output level	Error	Output level	Error	Output level	Error	Output level	Error
1	2	2	0	2	0	3	1	2	0
2	3	3	0	3	0	3	0	3	0
3	1	1	0	2	1	1	0	1	0
4	4	3	-1	4	0	4	0	4	0
5	1	1	0	1	0	1	0	1	0
6	1	1	0	1	0	1	0	1	0
7	2	2	0	2	0	2	0	2	0
8	1	2	1	1	0	2	1	1	0
9	3	3	0	3	0	3	0	3	0
10	2	2	0	2	0	2	0	2	0

As can be known from Tabs. 3 and 4, although the four models can converge after a limited number of iterations, the traditional BP neural network model needs 68 iterations to reach convergence, and its evaluation accuracy rate is only 80%; the additional momentum BP neural network model needs 63 iterations to reach convergence. Despite the iteration times similar to the traditional BP neural network model, it essentially adds momentum term to the traditional BP, which achieves better results by reducing the network oscillation, with evaluation accuracy rate up to 90%. BP neural network model with adaptive learning rate has the same evaluation accuracy rate as the traditional BP neural network model, but it reduces the network convergence time by about 45.6%. In contrast to the above three neural network models, momentum BP neural network model with variable learning rate takes only 33 iterations to achieve 100% evaluation accuracy rate, which not only speeds up the network convergence, but also improves network evaluation accuracy. Seen from the evaluation results, the model is superior to the other three BP models. Hence, momentum BP neural network model with variable learning rate can provide an accurate and reliable assessment of indoor air quality.

VI. Conclusion

In the indoor air quality assessment, by analyzing the five major environmental pollutant factors CO_2, CO, TVOC, formaldehyde and particulate matter, this paper establishes an evaluation model based on BP neural network. The network weights are effectively adjusted via the momentum BP algorithm with variable learning rate. The simulation results indicate that the evaluation model can enable effective integration and output after training to achieve accurate and reliable indoor air quality assessment, so that people can understand their environmental conditions and make effective response, which demonstrates certain application value of the model. Compared with the traditional BP neural network model, additional momentum BP neural network model and BP neural network model with adaptive learning rate, momentum BP neural network model with variable learning rate has faster convergence speed, smaller error and higher classification accuracy. Therefore, the momentum BP algorithm model with variable learning rate proposed herein is effective for indoor air quality assessment.

Acknowledgment

This research was supported in part by the Innovate UK (Grant 104013) and the institutional strategic grant - Global Challenges Research Fund (GCRF), that City, University of London, receives from Research England, UK Research and Innovation (UKRI) and the China Scholarship Council Fund under Grant No. 2018 082 10359. There is no any conflict of interest.

References

[1] Ayesha. A, Zeeshan. M, Hashmi. I, Zahid. U, Bhatti. M, "Microbial quality assessment of indoor air in a large hospital building during winter and spring seasons," *Building and Environment*, vol. 135, pp. 68-73, 2018.

[2] Kelly. F, Fussell. J, "Improving indoor air quality, health and performance within environments where people live, travel, learn and work," *Atmosphere Environment*, vol. 200, pp. 90-109, 2019.

[3] Shujun. H, Kunquan. L, Guangyi. Z, "Evaluation and analysis on the indoor air quality of colleges based on fuzzy analytic hierarchy process," *Environmental Engineering*, vol.32, pp. 90-94, 2014.

[4] Kulshreshtha. P, Khare. M, Seetharaman. P, "Indoor air quality assessment in and around urban slums of Delhi city, India," *Indoor Air*, vol. 18, pp. 488-498, 2018.

[5] Godio. R, Avigi. D, Campos. V, Tavares. T, Marchi. M, et al, "Indoor air quality assessment of elementary schools in Curitiba, Brazil," *Water Air and Soil Pollution: Focus*, vol. 9, pp. 171-177, 2009.

[6] Bochen. L, Xibo. D, Qinyao. C, "Research on air quality monitoring system and evaluation method in closed environment," *Journal of Harbin University of Science and Technology*, vol. 24, pp. 60-65, 2019.

[7] Röbbel. N, "Health co-benefits of climate change mitigation housing sector," Malta: WHO Library Cataloguing-in-Publication Data, 2011, pp. 129-41.

[8] Cheney, Washington, "Indoor air quality investigation review," Washington: Eastern Washington University, 2018, pp. 73-91.

[9] Nielsen. G, Larsen. S, Wolfoff. P, "Re-evaluation of the WHO (2010) formaldehyde indoor air quality guideline for cancer risk assessment," *Archivesof Toxicology*, vol. 91, pp. 35-61, 2017.

[10] Wei. G, Dong. W, "Application of fuzzy comprehensive evaluation method in evaluation of air quality in tunnel construction," *Shanxi Architecture*, vol. 44, pp. 148-150, 2018.

[11] Limin. T, Dongsheng. Y, Mengmeng. W, "Research on factors influencing indoor air quality based on analytic hierarchy process," *China Energy and Environmental Protection*, vol. 40, pp. 58-61, 2018.

[12] Zhu. C, Li. N, "Study on Grey Clustering Model of Indoor Air Quality Indicators," *Procedia Engineering*, vol. 205, pp. 2815-2822, 2017.

[13] Yu. T, Hanmin. Y, Weijuan. K, "Medical image registration of mutual information based on Powell and SA hybrid algorithm," *Microcomputer Information*, vol. 25, pp. 125-127, 2009.

[14] Weijun. H, Meng. H, Chenhui. W, "Design of automobile exhaust detection system based on improved BP neural network," *Transducer and Microsystem Technologies*, vol. 38, pp. 95-97, 2018.

[15] Jiajia. F, Yunsong. Q, Nilin. G, "Ship trajectory identification method based on improved BP neural network," *Computer Engineering and Design*, vol. 40, pp. 3649-3644, 2019.

[16] Jianfang. C, Yaojun. H, "Evaluation of large-scale online learning behavior based on parallel Adaboost-BP network," *Computer Applications and Software*, vol. 34, pp. 267-272, 2017.

[17] Hengde. Z, Tingyu. Z, Tao. L, Tianhang. Z, "Forecast of air quality pollutants' concentrations based on BP neural network multi-model ensemble method," *China Environmental Science*, vol. 38, pp. 1243-1256, 2018.