Demand Prediction of Emergency Supplies in Campus Public Health Emergencies Based on GM-GA-BP Neural Network

Yuhan Yue*

School of Computer Science & Engineering, Sichuan University of Science & Engineering
Yibin, China, 644002

*Corresponding author: 18383961008@163.com

Abstract. Medical alcohol is the most widely used and highly demanded emergency supply in both campus public environments and medical settings. It has a sharp contradiction between supply and demand in the event of a public health event. For this reason, forecasting the demand for medical alcohol is crucial for averting and managing unexpected public health issues on campus. An enhanced GM-GA-BP neural network prediction model was put out in this article. The benefits of the equal dimension and new information complement GM(1,1) model's new and old information replacement and small sample data prediction are combined. Following data fitting and modification, a BP neural network based on conjugate gradient algorithm is built to carry out training and prediction. Meanwhile, the connection weight and threshold were globally optimized through genetic algorithm.To forecast the demand for medical alcohol at a certain university in Yibin City from 2022 to 2023, a thorough comparison of a combination model vs the conventional BP neural network model was made. The results showed that the MAPE had been decreased by roughly 3% and the absolute percentage error was as low as 0.12686%. These findings have provided a scientific reference for the university to improve its emergency material reserve strategy represented by medical alcohol.

Keywords: Campus Emergency Supplies, Medical Alcohol Demand, BP Neural Network, Genetic Algorithm, GM(1,1) Model, Combined Model Prediction.

1. Introduction

In recent years, public health emergencies such as H1N1 influenza, novel coronavirus pneumonia have broken out one after another, posing a serious threat to human life and health. Alcohol with a concentration of 75% can be used to disinfect the skin and surrounding area, reducing the number of bacteria and lowering the virus’s survival rate.The contradiction between supply and demand of alcohol above is acute. On the one hand, ensuring timely and sufficient supply of medical alcohol can maximize the protection of the life safety of campus teachers and students. On the other hand, ensuring the rational procurement and transportation of emergency materials such as medical alcohol and disinfectant can avoid the safety hazards of storing flammable and explosive materials under the situation of excessively centralized procurement [1]. Therefore, the analysis of alcohol supply and demand scenarios by constructing an improved GM-GA-BP neural network combination model is of great reference significance for universities to further formulate procurement plans and modify strategies for preventing and controlling public health emergencies.

Currently, there is a lack of research on the application of big data technology and machine learning methods in the specific scenario of prevention and outbreak management of sudden public health emergencies in universities. Previous research has mainly focused on the analysis and summary of emergency management mechanisms [2]. However, given the uncertainty and unpredictability of public health emergencies, there is a lack of discussion on the process of emergency medical supplies stockpiling in university hospitals and how to dynamically predict changes in demand. Numerous accomplishments have been made by scholars in the study of the BP. For instance, Zou Lin [3] employed a BP neural network model to partition the severity of household difficulties caused by sudden public health incidents. Xie Yiyun [4] constructed a recognition model of weak links in the donation process in public health crises based on BP. Today's mainstream forecasting models also include the random forest RF model [5], the grey prediction GM(1,1) model

[6], the ARIMA model [7], and so on. Due to the limitations of a single predictive model in covering diverse features, the perspective of information presentation is constrained to some extent. Many scholars have established combined predictive models based on certain mathematical backgrounds to improve the accuracy of predictions. Wu Lei et al. [8] combined the ARIMA model with the BP using certain weight coefficients to obtain a more accurate prediction of the sales volume of a certain drug. Jiao Aonan et al. [9] proposed a grey-GA-BP combination model, which took the training data of the grey model as the input of the neural network to optimize the prediction accuracy. The results showed that the combination model can better predict the effects of community human resources.

There is no study on the scenario analysis of campus emergency material demand using the combination of the GM(1,1), GA, and the BP in the existing research results. Therefore, this paper establishes a GM-GA-BP neural network combination forecasting model, corrects the prediction results of every single step of the GM(1,1) with the actual data according to the empirical weight value, and brings the original data into the BP neural network model. After genetic algorithm optimization, the prediction research is carried out. The marginal contribution that may provide is that a single model has the limitations of large residuals and low accuracy. In contrast, this paper uses the GM(1,1) to require a small number of samples, which is suitable for short-term linear sequence prediction [10]. The advantages make up for the shortcomings of BP that require a large amount of data as training samples; the prediction results modified according to the weight value can take into account the influence of the actual value in each step of the prediction process of the GM(1,1). The genetic algorithm effectively avoids the BP falling into local optimization and gives full play to the advantages of dealing with nonlinear data.

2. Data pre-processing

2.1. Data profil

Based on the statistical data provided by the logistics service department of a certain university in a city, this paper selected medical alcohol in the hygiene department and the school infirmary as representative emergency supplies, and took the consumption of medical alcohol from January 2022 to February 2023 as the main research object. The experimental data during this period conforms to the basic stability of local epidemics in the post-epidemic background. It is a non-linear time series data and the significant fluctuations in the demand for emergency campus supplies during the epidemic unsealing period coincides with the irregular and explosive nature of public health emergencies. Therefore, it has relatively new and high practical reference value.

2.2. Data preprocessing operation based on equal dimension and new information complement GM(1,1)

Considering that the demand for emergency materials on campus is affected by basic environmental needs, seasonal epidemic disease cycle fluctuations and uncertainties in the context of public health emergencies, the smaller the prediction dimension is, the better it can simulate the real situation of short-term prediction. Thus, this paper innovatively divided the forecast dimension into the first and second half by month. The grey prediction based on GM(1,1) is suitable for solving the prediction problem of small sample data in the case of uncertain, poor information and time-related. The original sequence is tested by SPSS tool, and the value of the translation transformation is satisfied in the interval $(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+1}})$, which meets the modeling conditions.

Let the original series be $\mathbf{X}^{(0)}=(\mathbf{x}^{(0)}(1),\mathbf{x}^{(0)}(2),\dots,\mathbf{x}^{(0)}(n))$, and the primary cumulative sequence be $\mathbf{X}^{(1)}=(\mathbf{x}^{(1)}(1),\mathbf{x}^{(1)}(2),\dots,\mathbf{x}^{(1)}(n))$, where:

$$\mathbf{x}^{(1)}(k)=\sum_{i=1}^{k}\mathbf{x}^{(0)}(i)(k=1,2,\dots,n) \tag{1}$$

In the generated sequence of X's adjacent mean values (k) is:

$$\mathbf{z}^{(1)}(k)=0.5\mathbf{x}^{(1)}(k)+0.5\mathbf{x}^{(1)}(k-1),k=2,3,\dots,n \tag{2}$$

The grey differential equation is constructed:

$$\mathbf{x^{(0)}(k)+az^{(1)}(k)=b} \tag{3}$$

The whitening differential equation is:

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=b\,x^{(0)}(k)+a\,z^{(1)}(k)=b \tag{4}$$

A is the development coefficient, and b is the grey action. The least square method is used to estimate the parameters by $\mathbf{[a,b]^T=(B^TB)^{-1}B^TY}$, and the following is obtained:

$$B=\begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & n \end{bmatrix},\ Y=\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \tag{5}$$

The time-corresponding sequence of the GM(1,1) differential equation is obtained:

$$\hat{x}^{(1)}(k+1)=[x^{(0)}(1)-\frac{b}{a}]e^{-ak}+\frac{b}{a} \tag{6}$$

Then the final fitting value $\hat{x}^{(0)}$ is obtained by cumulative reduction.

In the alcohol demand data, due to the impact of the unsealing policy of the new crown pneumonia epidemic in late December 2022, the data of this period time hid the sample's own pattern, so a 5-dimensional traditional GM(1,1) was established for this anomaly, and the first five original data predicted values were corrected by averaging with the original data of this point to improve the prediction accuracy of the combined model. The equal dimension and new information complement GM(1,1) introduces new known information over time to improve the effect of grey segment whitening and remove the oldest data in the sequence. After multiple weight combination tests, the predicted value $\hat{x}^{(0)}(k+1)$ of the subsequence is multiplied by 0.3 and the actual value is multiplied by 0.7. After summing them up, they are used as corrected input values to supplement. The posterior error ratio of the model is 0.278 < 0.35, which is in line with the target accuracy and is used for metabolism and dynamic prediction.

3. Prediction model establishment

3.1. Model introduction

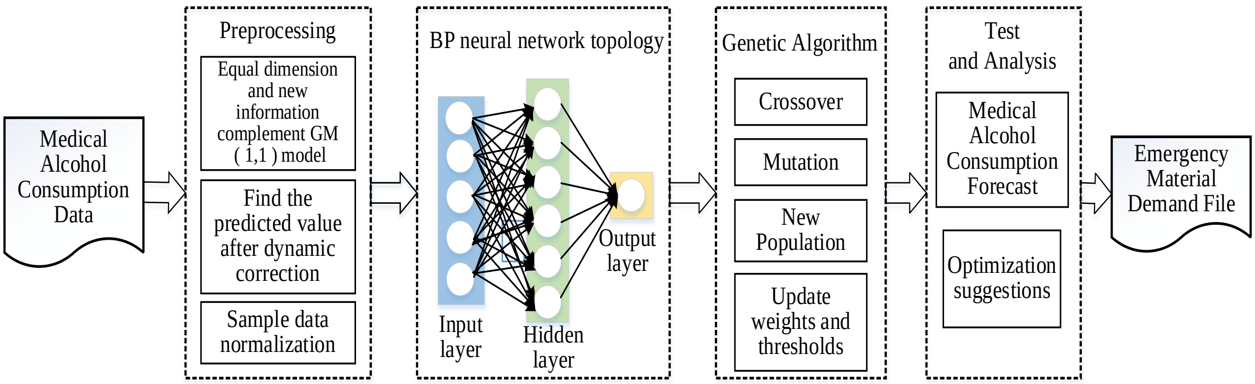

Figure 1. Scenario prediction flow chart of medical alcohol demand in universities

The scenario prediction method of medical alcohol demand in universities carried out in this paper is shown in Figure 1, and the specific operation steps are as follows.

Step 1: The discrete data on medical alcohol consumption in a university in the past year were collected, and the data preprocessing was carried out according to the above description. The improved GM(1,1) fitting value will become the input sequence of BP, that is $\boldsymbol{Z_t=0.3Z_{1t}+0.7Z_{2t}}$. $\boldsymbol{Z_{1t}}$

is the predicted value of GM(1,1), and Z_{2t} is the actual value. At the same time, data normalization processing can further improve the prediction accuracy. The data with large distribution changes are mapped into the interval [-1,1], and the formula is :

$$y_i=(y_{max}-y_{min})\times\frac{x_i-min(x)}{max(x)-min(x)}+y_{min} \tag{7}$$

Step 2: Constructing BP. BP is a multi-layer feedforward network based on error back propagation. If the error between the output value of the forward process and the expected value exceeds the target range, it is necessary to continuously adjust the connection weight and bias weight in the back propagation process to the hidden layer and the input layer until the error and other parameters reach the preset range, and then stop learning. The formula is expressed as:

$$y_t=f\{\sum_{i=1}^{n}\omega_i[f(\sum_{j=1}^{m}\omega_j\, y_{t-j})+\xi_{1t}]+\xi_{2t}\} \tag{8}$$

Where y_t is the input vector, $y_{t-j}(j=1,2,...,m)$ is the output vector, $\omega_i(i=1,2,...,n)$ is the connection weight vector between the input layer and the hidden layer, $\omega_j(j=1,2,...,m)$ is the connection weight vector between the hidden layer and the output layer, f is the transfer function, ξ_{1t} and ξ_{2t} are the bias weights of the hidden layer and the output layer, respectively. The BP with strong adaptability and mapping ability and high fault tolerance is applied to this study to establish a network model with 5 input nodes, 6 single hidden layer nodes and 1 output node. The specific modeling process is described as follows.

Step 3: Optimization of BP using genetic algorithm global search capability. By setting the coding method, the initial population is generated. After continuous selection, crossover and mutation, the optimal connection weight and bias value are obtained, and the training speed of the model is improved. The specific optimization process is described below.

Step 4: Prediction and analysis of combined model. The data set is randomly sorted and divided into 70 % for model training, and the remaining 30 % data is used for model testing. The GM-GA-BP model fitting value can better approach the real value, and the accuracy of the model is high. Thus, the scenario file of medical alcohol demand in universities is generated. The model is applied to provide an objective reference for the relevant functional departments of universities to prevent and control public health emergencies.

3.2. Establishment of GM-GA-BP neural network model

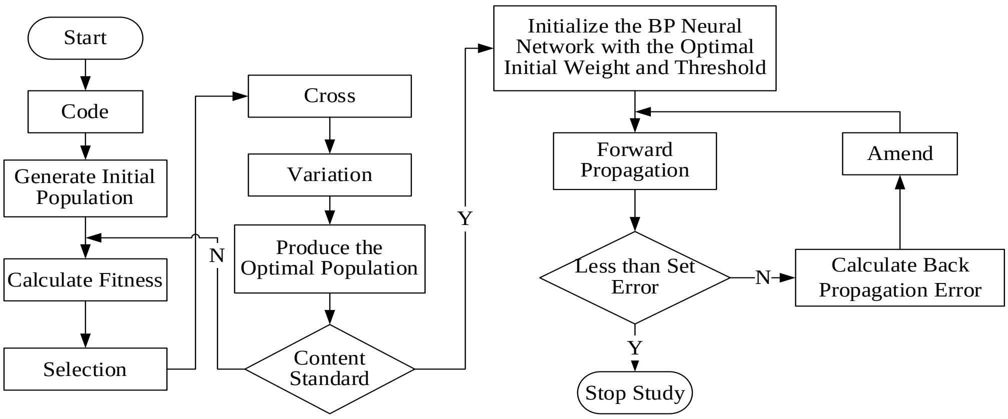

Figure 2. The specific prediction flow chart of the combined model

The specific prediction process of the combined model is shown in Figure 2. The newff function is used to construct the BP. The fitting value of the improved GM(1,1) with a step size of 5 is taken as the input, and the actual data of the next month is taken as the output. The training samples and test samples are constructed according to the ratio of 7:3. The number of hidden layer nodes is an important parameter of the prediction model. The empirical formula is $L=\sqrt{u+v}+n$, where L is the number of hidden layer neurons, u is the number of input neurons, v is the number of output neurons, and n is a constant of [0,10]. According to the results of the trial and error method, the number of hidden layer nodes is determined to be 6. According to past experience, the hyperbolic tangent S-type transfer function tansig is selected as the hidden layer in this study, and the linear transfer function purelin is selected as the output layer. By comparing the training function traingd of the fastest gradient descent algorithm, the trainlm of the Levenberg-Marquardt algorithm and the traincgp of the conjugate gradient algorithm, the fitting results of alcohol demand in university hospitals from January 2022 to October 2022 are shown in Table 1.Experiments show that the neural network prediction residual optimized by the conjugate gradient algorithm is the smallest.

Table 1. Three kinds of training function fitting results table

Training function	Training frequency	Learning rate	Minimum target error	Minimum performance gradient	R	MAPE	Iteration times	MSE
traingd	10000	0.001	1e-8	1e-16	0.90678	9.7618%	10000	0.00417
trainlm	1000	0.001	0.001	1e-16	0.97272	6.7188%	3	4.63e-05
traincgp	1000	0.001	0.001	1e-16	0.98193	4.7595%	8	0.000893

The conjugate gradient algorithm improves the search direction of the traditional BP, the convergence speed is faster, and the over-fitting phenomenon is avoided. Based on this algorithm, the model is constructed with 1000 training times. After 8 training iterations, the R value is close to 1, and the MAPE and MSE are relatively minimum. Many experimental results show that the model construction effect is ideal, so the traincgp function is selected for the next step of genetic algorithm optimization. The operation is as follows:

(1) The global search ability of genetic algorithm is affected by the change of population quantity. Firstly, it is determined that the binary coding method is adopted in this paper, and the initial population number is 20.

(2) Select the fitness function, the overall fitness calculation formula is :

$$F=\sum_{i=0}^{n} \boldsymbol{f_t}，\ \boldsymbol{f_t}=\frac{\mathbf{1}}{|\boldsymbol{M_p}-\boldsymbol{M_q}|^{\mathbf{2}}} \tag{9}$$

The individual fitness value is the reciprocal of the square of the actual value and the predicted value error, $\boldsymbol{M_p}$ is the actual value, and $\boldsymbol{M_q}$ is the predicted value. The larger the function value, the better the fitness.

(3) Compared with the survival rules of the fittest in biological evolution, the roulette selection method is used to find out the excellent individuals. The coding crossover method is the two-point crossover method, and the mutation is performed by the Gaussian mutation method. According to the actual situation of solving the problem in this paper, the maximum evolution algebra of model parameters is set to 30, the crossover probability is 0.8, and the mutation probability is 0.2.At the same time, the gradient descent momentum function learngdm is added to learn the weight and threshold change rate, so that the BP based on conjugate gradient algorithm can quickly achieve global optimization.

4. Solution of GM-GA-BP neural network model

4.1. Output in brief

The GM-GA-BP model was trained for alcohol demand in university hospitals from January 2022 to October 2022, and the network output results from November 2022 to February 2023 were compared with the actual values to carry out model prediction and evaluation. Mean absolute percentage error (MAPE), R, and mean square error (MSE) were selected as model evaluation indicators. After setting the training parameters as in Table 1, the training error trend of the improved combined model is obtained as shown in Figure 3. The error decreased continuously with the number of iterations more smoothly and broke through the target error to 0.000991 at 13 iterations. The regression fitting diagram is shown in Figure 4. The R value is around 1, and most of the data points are gathered around the fitting line. The fitting effect is good, and the MAPE value is 1.888 %.

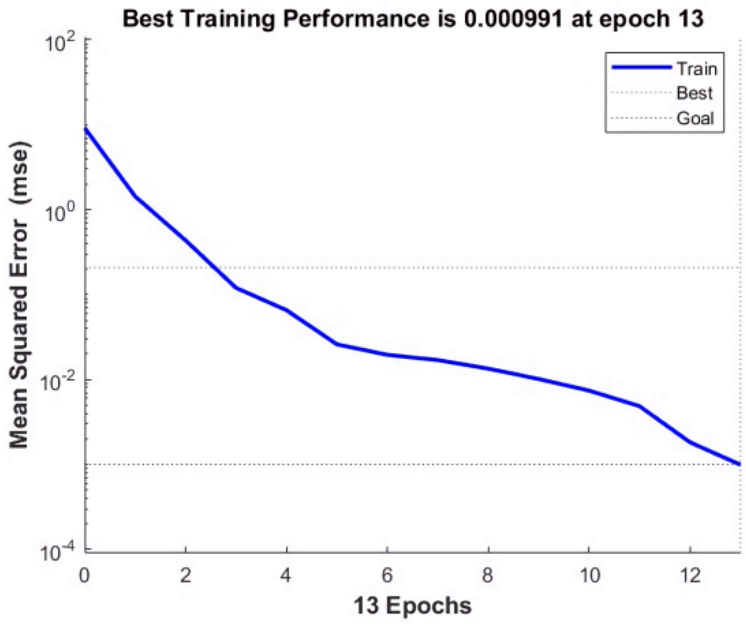

Figure 3. Combined model training error trend chart

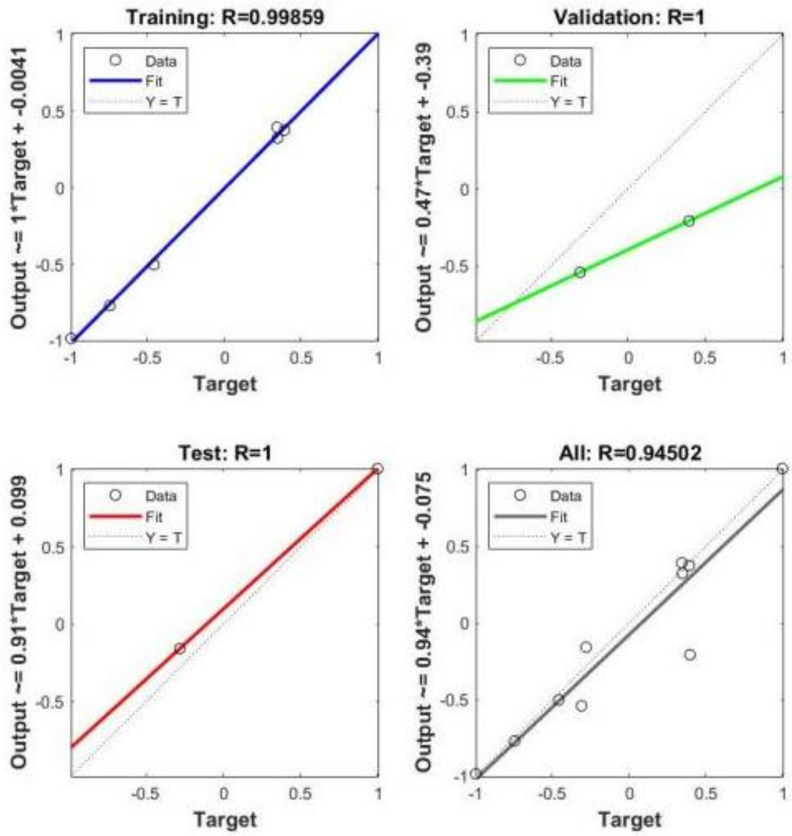

Figure 4. Combined model regression fit graph

4.2. Predictive analysis

The sample of medical alcohol demand test in universities from November 2022 to February 2023 was predicted and compared with the real value. The predicted fitting graph is shown in figure 5, and the simulation results are close to the actual data points. The simulation results are in good agreement with the actual data points.

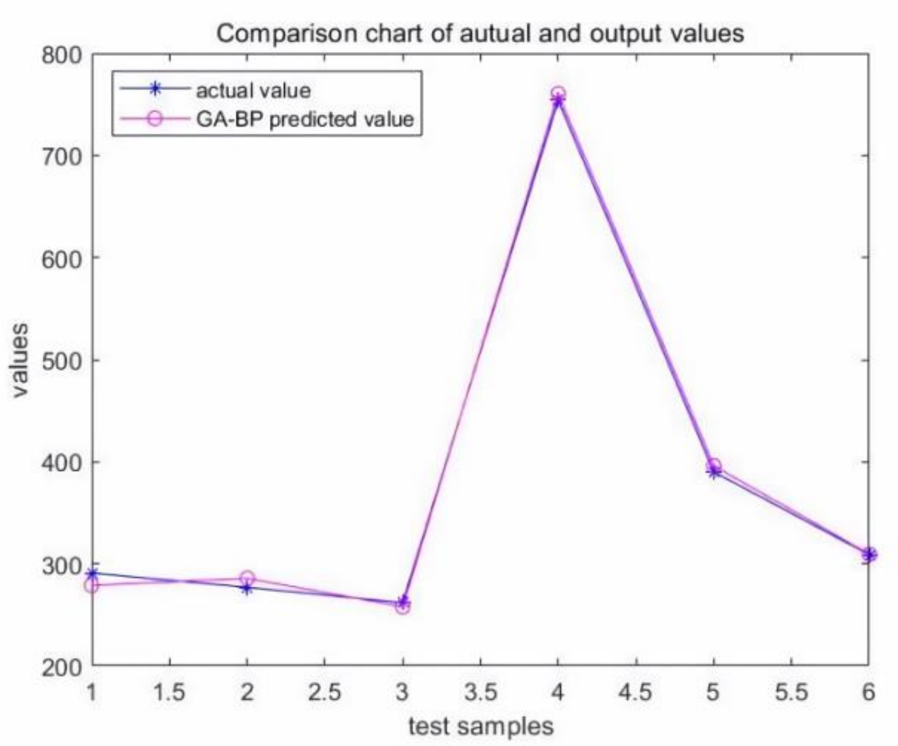

Figure 5. Combined model prediction fitting chart

The prediction residuals of GM-GA-BP model and unimproved BP model were compared intuitively. As shown in Table 2, it is found that the relative residuals of the combined model are within 4.1 %, which is generally consistent with the real demand trend of medical alcohol, and the prediction accuracy is higher.

Table 2. Comparison table of prediction results

Date	Actual value	BP Neural Network	Combined model	BP absolute percentage error	Combined model absolute percentage error
Early November 2022	291	295.897	278.979	1.68282%	4.13093%
Late November 2022	277	263.721	285.844	4.79386%	3.19278%
Early December 2022	262	230.791	258.054	11.11912%	1.50611%
Late December 2022	754	811.728	760.265	7.65623%	0.83090%
Early January 2023	390	393.908	396.006	0.23282%	1.54000%
Late February 2023	309	314.811	309.392	1.88058%	0.12686%

4.3. Experimental conclusion

Through many experiments, it is found that the GM-GA-BP model has smaller mean absolute percentage error and mean square error of training results under fewer iterations. The MAPE value is as low as 1.888 %. The MSE value exceeds the preset target by 0.000991. The correlation coefficient R is closer to 1, and the R of the training set is as high as 0.99859. In the prediction results of medical alcohol demand in universities from November 2022 to February 2023, the minimum absolute percentage error of the combined model is as low as 0.12686 %, and the maximum absolute percentage error is only 4.13093 %, which means that the prediction accuracy is higher and the comprehensive performance is better. The GM(1,1) with equal dimension and new information can take into account the replacement of old and new information under the condition of less and uncertain information, while the BP implicitly represents the nonlinear relationship, and can dynamically and adaptively approximate the optimal prediction effect after genetic algorithm optimization. Based on this, the combined model has certain generalizations in the prediction of medical alcohol sample data, which is suitable for the prediction of emergency materials in universities.

5. Conclusion

Based on the comprehensive analysis of the current situation of the demand for medical and health emergency materials in a university in a particular city, this paper has clarified the uncertainty of public health emergencies and the sudden increase in the demand for emergency materials, and established a safe and convenient model for predicting the demand for emergency materials in universities. In this paper, a GM-GA-BP neural network model was proposed.The BP prediction model based on conjugate gradient method was replaced with the updated data of GM(1,1) model with equal dimension and fresh information. The connection weight and threshold were universally optimized using the genetic algorithm. Combined with the data of medical alcohol demand in universities, the combined model was programmed to verify the objective and good prediction effect. Compared with the traditional BP neural network model, MAPE is reduced by about 3 %, and the absolute percentage error is as low as 0.12686 %. According to the above demand forecast results, the following suggestions are put forward for the improvement of emergency material management strategies in universities: After the release of COVID-19 epidemic control, the sharp rise in the consumption of sterilization materials represented by medical alcohol returned to stability after 2 months. Based on the background of influenza A and seasonal influenza, university administrators should refer to the model prediction to fortify the reserve of emergency materials in advance, formulate reasonable information monitoring strategies for procurement, storage and other processes, concentrate on in-depth excavation of the real-time data value fed back by logistics departments, and prepare ahead of time and adjust in real time to actively respond to campus public health emergencies.

References

[1] Xu Jie, Liu Yu. Analysis of prevention and control strategies for public health emergencies of infectious diseases in schools [J]. Rural Health in China, 2020, 12(14): 21.

[2] Yang Wenjun. Research on emergency management of public health emergencies in primary and secondary schools in K District of D City [D]. Jinan: Shandong university, 2022.

[3] Zou Lin. Construction of precise funding model based on BP neural network—Analysis of the impact on public health emergencies [J]. Journal of Southeast University(Philosophy and Social Science), 2021, 23(S2): 116-119+123.

[4] Xie Yiyun. Research on the optimization of public health emergency material donation process based on BP neural network [D]. Xian: Northwest University, 2022.

[5] Wang Jiaojiao, Chen Yuanyu, Zheng Ziwei, et al. Comparison of the efficacy of three risk prediction models in predicting carotid atherosclerosis in steel workers [J]. Chinese General Practice, 2022, 25(11): 1334-1339.

[6] Shi Jie, Liu Yuqing, Hui Jufen, et al. Application of GM (1,1) residual error correction model in prediction of gonorrhea incidence in China [J]. Medical Animal Control, 2023, 39(02): 103-105+110.

[7] Li Shouli. The application of time series model in the prediction of GDP in prefecture-level cities [D]. Zhengzhou: Zhengzhou university, 2013.

[8] Wu Lei, Xu Huaifu. The application of the new ARIMA-BP combination model in the sales management of pharmaceutical enterprises [J]. Shanghai Medicinal Monthly, 2016, 37(07): 68-72.

[9] Jiao Aonan, Shao Yiying, Mo Yingning, et al. Prediction of community health human resources in China [J]. Health Resources in China, 2022, 25(05): 644-649.

[10] Zheng Minggui, Yu Ming, Fan Qiurong, et al. Prediction of lithium carbonate demand in China from 2025 to 2035—A combined model based on grey correlation analysis and ARIMA-GM-BP neural network [J]. Advance in Earth Science, 2023, 38(04): 377-387.