A Forecasting Analysis of Health Technicians Demand in Hainan Province based on Several Combination Forecasting Models*

Shihong Wang, Lianhua Liu
International School of Public Health and One Health, Hainan Medical University, Haikou 571199, China.

Abstract: **Objective:**Several combination models were used to forecast the demand for health technicians in Hainan Province in order to find the best combination of forecasting models and thus provide the relevant departments with a more scientific basis for their planning. **Methods:**First using ARIMA method, GM(1,1) method and trend extrapolation method to establish single forecasting models, and then based on single models, adopt equal weight method , reciprocal errors method ,odds-matrix method and artificial neural networks method to establish four kinds of combination forecasting models, finally evaluate all the prediction models and select the best model to make a short -term forecasting. **Results:**The combined forecasting model generally has a lower forecasting error than the single forecasting model. The combined artificial neural network model has the lowest forecasting error and is the relatively optimal model for forecasting health technicians. It is predicted that the demand for health technicians in Hainan Province will increase steadily from 2021 to 2023. **Conclusion:**The demand for health technicians in Hainan Province is still high, and training efforts should be further strengthened to lay a solid foundation of health care for the construction of the Hainan Free Trade Port.

Keywords: HealthTechnician; Combination Forecasting Model; Hainan

1. Introduction

As the first resource in health resources, health human resources are the key to protecting the health of citizens and promoting the healthy development of health services. Health technicians are the most important representative of health human resources, and the number of health technicians is an important indicator to measure the level of public health and medical services in a region[1]. In recent years, with the establishment of the Bo'ao Le Cheng Medical Pilot Zone and the construction of the Free Trade Port in Hainan Province, the number of migrants is increasing year by year, and the demand for medical and health services is also increasing. Therefore, it is of great importance to establish a scientific and reasonable prediction model of the number of health technicians for the rational allocation of health human resources in Hainan Province and even to guarantee the construction of the free trade port.

The issue of forecasting for health technicians has been a hot topic of research in recent years. Many researchers have used single forecasting models such as gray neural network model, ARIMA model, and gray GM(1,1) model to forecast the demand of health technicians in various provinces and cities[2-5]. However, single prediction model contains only part of the information of the predicted object, and the accuracy and stability of the prediction results are poor, while combining each single prediction model by certain rules can contain more comprehensive information of the original data,thus further improving the accuracy and stability of the prediction [6]. Therefore,many researchers have started to apply Combination Models for prediction[7-9]. However, many of these researchers have only used a certain combination of models, and most of them are linear combinations. But there are many kinds of combination prediction methods, the common ones are linear

combination methods such as equal weight method, minimum variance method, error inverse method, dominance matrix method, etc. There are also non-linear combination methods such as artificial neural network, weighted harmonic average, etc. And the prediction accuracy of different combination methods will also be different.

Therefore, this study first used the data of health technicians in Hainan Province from 2005 to 2020 to establishing three single prediction models, and then used four combined methods to establish combination prediction models. Finally, the error evaluation of all the prediction models was conducted, and the best model was selected to make a short-term prediction of the needs of health technicians in Hainan Province, in order to provide a more accurate and reliable decision-making basis for the government and health administration departments to allocate health resources.

2. Materials and methods

The data of health technicians were derived from the Statistical Yearbook of Hainan Province from 2005 to 2020.Based on the data , three single prediction methods: ARIMA, GM (1,1), and trend extrapolation method were established[10-12]. and then four combined forecasting models were established based on the single model using equal weight method, error inverse method, dominance matrix method, and artificial neural network method,and finally three forecasting accuracy evaluation indexes, MSE, MAE, and MAPE were used to evaluate the accuracy of all the prediction models.

2.1 Combination models

Combined model with equal weighting method is the simplest combined prediction method, that is, to assign the same weight to each single prediction model for linear combination [6]. For example, if there are three single prediction models, each single model has a weight of 1/3 in the combined model, and although the combined model with equal weights is simple to operate, the prediction accuracy is often higher than that of many complex combined models,This phenomenon has been regarded as the "combined prediction puzzle"[13].

The error reciprocal method is to give different weights to each single prediction model according to the error size of each single prediction model. The smaller the single prediction model error is, the larger the weight in the combined model is. For example, there are three single models, and Q_i is the sum of error squares of the ith single prediction model, so the weight of the ith single model in the combined model is: $w_i = \frac{1}{Q_i} \Big/ \sum_{i=1}^{3} \frac{1}{Q_i} \quad (i = 1,2,3)$.

The dominance matrix method is to assign weights according to the prediction accuracy of each single prediction method in each sample[14]. For example, A, B and C represent three single prediction models, n_A represents the times that error of model A is smaller than model B and C in the whole sample set, and n_B and n_C follow similar definition.s. The allocation rules of the weigh of A, B and C in the combined model is as follows:

$$w_A = \frac{n_A}{n_A + n_B + n_C}, w_B = \frac{n_B}{n_A + n_B + n_C}, w_C = \frac{n_C}{n_A + n_B + n_C}$$

Artificial neural network is a mathematical model similar to the structure of neural synaptic connections in the brain. Artificial neural network combination is a nonlinear combination method that is capable of learning a highly nonlinear mapping relationship between input and output data.

With one hidden layer, the prediction results of three single prediction methods are used as the network input layer, the

actual value as the output layer. The weights are learned by forward and back propagation to determine the weights of the neurons in each layer. Finally, the trained model is used for prediction[15].

2.2 Evaluation index

Evaluating the accuracy and reliability of prediction results is an important part of prediction analysis. and several evaluation indicators are commonly used to evaluate the prediction results, while using only one error evaluation index to evaluate the model is not reliable enough. The commonly used evaluation indexes are calculated as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|$$

In the above equation, $\hat{y}_i$ is the predicted value, and y_i is the actual value.

3. Results

3.1 Current status of the number of health technicians

During 2005-2020, the number of health technicians in Hainan Province increased from 30,056 to 74,378 (Table 1), with an average annual growth rate of 6.20% and an average annual growth of 2,770 people.

Table 1. Raw data and predicted values of health technicians in Hainan Province from 2005 to 2020

Year	Health technicians	Three single prediction methods			Four combined prediction methods			
		ARIMA	GM(1,1)	trend extrapolation	equal weight	error reciprocal	dominance matrix	artificial neural network
2005	30056	30049	30056	29229	29778	29728	29898	29905
2006	30787	30813	31010	31078	30967	31003	30949	30814
2007	32545	32496	32972	33045	32838	32918	32807	32445
2008	33875	34566	35058	35138	34921	35004	34889	34588
2009	37857	36057	37277	37364	36899	37099	36836	37329
2010	39520	41781	39636	39732	40383	40047	40458	39684
2011	43295	43237	42144	42250	42544	42376	42574	42697
2012	44720	45284	44811	44926	45007	44939	45010	45373
2013	48192	47721	47647	47768	47712	47708	47697	47943
2014	50557	49843	50662	50785	50430	50568	50378	50752
2015	54677	54590	53868	53984	54147	54039	54160	54481
2016	57784	58034	57276	57375	57562	57447	57579	57786
2017	60579	62313	60901	60964	61393	61172	61442	60758
2018	63663	62445	64755	64761	63987	64355	63890	63441
2019	67695	66507	68853	68773	68044	68413	67958	67877
2020	74378	72095	73210	73008	72771	72936	72754	74306

3.2 Single prediction model

R4.1.0 was used for establishing the ARIMA model of health technicians. By ADF test, white noise test and model order, The optimal model is ARIMA (1,3,0). The model equation is $\nabla^3 x_t = -0.88\nabla^3 x_{t-1} + \varepsilon_t$.

Base on MATLAB7.0, GM (1,1) model of health technicians was fitted and the accuracy rating was evaluated(Table 2). With C value<0.35 and P value=1, indicate that the GM (1,1) model had a good fitting accuracy.

Table 2. The GM (1,1) model of health technicians

Index	GM(1,1)	a	μ	C	P	Evaluation results
health technicians	$x^{(1)}(k+1) = 500451.17e^{0.06k} - 470395.17$	-0.06	28223.7	0.03	1	Good

Trend extrapolation models such as linear, quadratic, cubic, logarithmic, exponential, and composite functions were fitted to the health technician data using SPSS 20.0. Among the models, the cubic function had the best effect (Table 3), and R^2 of the model was above 0.99, indicating that the trend extrapolation model is reliable and significant (P<0.05).

Table 3. Trend extrapolation model for health technicians

Index	F value	P value	R^2	Model
health technicians	1241.5	<0.001	0.997	$y = 27487.80+1688.96t+50.45t^2+1.36t^3$

3.3 Combined prediction models

based on three single prediction models, three weight assignment methods:equal weight method, error inverse method and dominance matrix method, were used to establish linear combination models (Table 4) , and also used artificial neural network method to establish nonlinear combination model. y_{1t}, y_{2t}, y_{3t} represent the predicted value of three single predicted models in t year respectively, and y_t represents the predicted value of the combined model.

Table 4. Three linear combined models

Combined models	weight of ARIMA	weight of GM（1,1）	weight of trend extrapolation model	Combined models
Equal weight method	1/3	1/3	1/3	$y_t = 1/3*y_{1t} + 1/3*y_{2t} + 1/3*y_3$
Error reciprocal method	0.174	0.431	0.395	$y_t = 0.174y_{1t} + 0.431y_{2t} + 0.395y$
Dominance matrix method	6/16	7/16	3/16	$y_t = 6/16*y_{1t} + 7/16*y_{2t} + 3/16*y$

3.4 Comparison of prediction results

The prediction values of each single prediction model and combined prediction model are shown in Table 1.The

prediction effects are shown in Table 5.The results show that among the three single prediction models, the GM (1,1) model has the smallest prediction error, followed by the trend extrapolation models. The error of the combined prediction models was generally smaller than the single prediction model, and the artificial neural network combined model has the smallest error among all the combined models. Many researchers have also proved that the combination prediction method of artificial neural network has a higher prediction accuracy[16], so we finally choose combination model of artificial neural network as the short-term prediction model for health technicians in Hainan Province.

Table 5 Comparison of the effects of the different prediction models

	MAE	MAPE	MSE
ARIMA	837.52	1.67	1316862
GM（1,1）	592.34	1.21	533562
trend extropolation	657.63	1.42	581721
Equal weight combination	569.29	1.23	477887
Error reciprocal combination	586.44	1.25	468288
dominance matrix combination	558.70	1.20	485287
Artificial neural network combination	**264.43**	**0.62**	**118133**

3.5 Prediction results of the best combined model

Combined prediction model of artificial neural network was used to make a short-term prediction of health technician needs in Hainan province from 2021 to 2023, the results are shown in Table 6.

Table 6. Short-term prediction value of health technicians in Hainan Province

Year	Predicted values by artificial neural networks model
2021	81967
2022	87193
2023	89707

4. Discussion

The combined forecasting model is, to some extent, an improvement over the single forecasting model.It improves the accuracy of a single prediction model and reduces the error of prediction.Using multiple combination methods can extract the advantages of a single model from different sides, so as to select the best combination and obtain more accurate prediction results, thus providing a more accurate and reliable decision-making basis for scientific allocation of health human resources.Later stage, the accuracy of the prediction model can be further improved from three aspects: The first is to optimize the selection strategy of single model, try to select some new single prediction models, such as SVM [17-18] and other machine learning models;the second is the selection of combination methods, whether the selection of combination methods is appropriate directly affects the combination prediction effect, so other combination methods can also be tried; the Third is other social factors should also be considered, such as the aging of the society's population, health policy, economic development and so on.

The forecast results show that from 2021 to 2023, the demand of health technicians in Hainan province will increase year by year. And with the establishment of Hainan Boao Le Cheng Medical Pilot Zone and the construction of Hainan Free Trade Port, the demand of health technicians in Hainan province will continue to increase. However, there is only one undergraduate medical college in Hainan province, and its economic development is relatively backward, so the attraction of medical talents from other provinces is weak. Therefore, it is recommended that the relevant departments use the forecast

results as a scientific basis for policy formulation and talent training, vigorously develop local medical education, and further enhance the coordination between health administration departments and education departments,adjust the scale and training direction of medical education according to the number of health technicians needed, and actively promote a balance between supply and demand of health human resources [19].

Author contributions

W-SH and L-LH contributed equally to this study and approved the final version.

References

[1] Chen Y, Xue Q. A Forecasting Analysis of Health Human resources in Inner Mongolia based on GM (1，1) Model. Medicine and Society, 2019,32 (09): 20-23.

[2] Zhang Y. A gray neural network model for predicting trends in the number of health technicians. Chinese Journal of Health Statistics, 2016,33 (02): 331-332 + 335.

[3] Sun J. Prediction Analysis on Demand of Health Workers in Guangxi Zhuang Autonomous Region Based on ARIMA Model. Modern Preventive medicine, 2017,44 (12): 2196-2201.

[4] Liu Q, Zou B, Zhu P, Zhang F, Chang R, Dong Y, Wan H. Forecast of Health Professionals Demand Based on ARIMA Model in Hubei Province. Medicine and Society, 2020,33 (02): 18-21.

[5] Li S, Liu X, Zhang X.Study on the prediction of health human resource demand in Hubei province based on grey system . Soft science of Health, 2020,34 (03): 49-53.

[6] Ling L, Zhang D. A Review of Combined Prediction Model Construction Methods and Their Applications. Statistics & Decision, 2019,35 (01): 18-23.

[7] Hou Y, Wang D, Chen Y, Zheng X, Gao Y, Zheng W. Construction and application of combination forecasting model of health human resources in Shandong Province. Soft science of Health, 2021,35 (04): 72-75 + 79.

[8] Bai Y, Wang D, Han D. Trend prediction of the number of urban and rural health technicians in China based on the combination model [J]. Chinese Journal of Health Statistics, 2019,36 (01): 111-114.

[9] Han C, Hu X, Zhao Y, Dong Z. Combination Model in Forecasting the Total Number of Health Technical Staffs of Our Country. Chinese Journal of Health Statistics, 2011,28 (04): 391-393.

[10] Zhang R, Zhao D, He S, Liu Z, Sun B, Xu X. Health human resource prediction based on Grey GM (1,1) model and ARIMA model,Sichuan Province.Modern Preventive Medicine, 2017,44 (07): 1242-1247.

[11] Yang Y, Mao S, Xue Y, Tian X, Shi X. Prediction on the Incidence rate of AIDS in China with Gray Model(1,1) and Trend Extrapolation Model. Chinese Journal of Health Statistics, 2014,31 (06): 952-954.

[12] Li W, Cui Q, Zhang L. Trend extrapolation and ARIMA forecast hospital diagnosis and hospitalization in China. Chinese Journal of Health Statistics, 2016,33 (03): 477-478 + 481.

[13] Gerder Claeskens, Jan R, Magnus, Andrey L.Vasnev,et al.The forecast combination Puzzle:A Simple Theoretical Explanation[J].International Journal of Forecasting,2016,32.

[14] Dou J, Ma H. Prediction of Building Energy Consumption Based on Ensemble Artificial Neural Networks. Computer Simulation, 2022,39 (05): 438-443.

[15] Zhang Q. Application Research on an optimal mix forecasting method on ANN. Systems Engineering-Theory & Practice, 2001 (09): 90-93.

[16] Li S, Liu L, Zhai M. Prediction for short-term traffic flow based on modified PSO optimized BP neural network. Systems Engineering-Theory & Practice, 2012,32 (09): 2045-2049.

[17] Bao X, Xiang G, Shi L, Wang D. Prediction of total health expenditure in Guangdong,based on GM(1,1)-SVM combination model. Modern Preventive medicine, 2022,49 (05): 856-859.

[18] Shen G, Wang X, Kong X. Short-term traffic volume intelligent hybrid forecasting model and its application. Systems Engineering-Theory & Practice, 2011,31 (03): 561-568.

[19] Liang S, Feng Q, Wang Y, Luo H, Qin Q, Zhang Y. Studying on the equity of health resources allocation in the minority areas sampled with Guangxi. Chinese Health Service Management, 2015,32 (09): 677-680.

Funding:This study was funded by the Natural Science Foundation of Hainan Province (No. 821QN0895).

Corresponding author:LIU Lian-hua,E-mail:liulianhua.009@163.com,Tel:86-137-07521716