Mathematics

AIMS Mathematics, 9(2): 3235–3252.
DOI: 10.3934/math.2024157
Received: 31 October 2023
Revised: 13 December 2023
Accepted: 18 December 2023
Published: 03 January 2024

http://www.aimspress.com/journal/Math

Research article

Huizhou resident population, Guangdong resident population and elderly population forecast based on the NAR neural network Markov model

Dewang Li[1], Meilan Qiu[1] and Zhongliang Luo[2,*]

[1] School of Mathematics and Statistics, Huizhou University, Huizhou 516007, China
[2] School of Electronic and Information Engineering, Huizhou University, Huizhou 516007, China

* **Correspondence:** Email: hk3333@163.com; Tel: +8613352648888.

Abstract: We propose a nonlinear auto regressive neural network Markov model (NARMKM) to predict the annual Huizhou resident population, Guangdong resident population and elderly population in China, and improve the accuracy of population forecasting. The new model is built upon the traditional neural network model and utilized matrix perturbation theory to study the natural and response characteristics of a system when the structural parameters change slightly. The delay order and hidden layer number of neurons has a greater effect the prediction result of NAR neural network model. Therefore, we make full use of prior information to constrain and test when making predictions. We choose reasonable parameter settings to obtain more reliable prediction results. Three experiments are conducted to validate the high prediction accuracy of the NARMKM model, with mean absolute percentage error (MAPE), root mean square error (RMSE), STD and R^2. These results demonstrate the superior fitting performance of the NARMKM model when compared to other six competitive models, including GM (1,1), ARIMA, Multiple regression, FGM (1,1), FANGBM and NAR. Our study provides a scientific basis and technical references for further research in the finance as well as population fields.

Keywords: NAR; neural network; fractional grey models; machine learning; parameter estimation and prediction; stochastic Markov process; time series
Mathematics Subject Classification: 62M05, 62M10, 62P05, 62P20

Nomenclature

Acronym	Full name
NAR	Nonlinear auto regressive
GM(1,1)	Traditional grey forecasting model
NARMKM	Nonlinear auto regressive neural network Markov model
MAPE	Mean absolute percentage error
RMSE	Root mean square error
STD	Standard deviation of absolute percentage error
R^2	Coefficient of determination
ARIMA	Auto regressive integrated moving average
FGM(1,1)	Fractional grey model
FANGBM	Fractional accumulation nonlinear grey Bernoulli model
ARMA	Auto regressive and moving average model
PSO	Particle swarm optimization
BP	Back propagation
PPAR	Projection pursuit auto regressive
CBD	Cairns-Blake-Dowd
SVM	Support vector machines
PID	Proportional integral derivative
GRNNMKM	General regression neural network Markov model

1. Introduction

Population is an important indicator to show the basic conditions of national conditions and national strength and it is the primary consideration in urban development analysis and planning. The number of population directly affects the economic and social development of a country as well as the allocation and utilization of resources. Therefore, the study of population is of great significance.

Population prediction takes the law of population development as the main body to determine the parameters, and the acquisition of these data and the choice of prediction algorithm greatly affect the accuracy of prediction results. Different models produce different results for population prediction. Jia et al. [1] used a BP neural network to establish a three-layer neural network structure model to predict the total population of the country. They established the method of combining the adaptive learning rate and the additional momentum method to learn and train the network, and the prediction results showed that the method was feasible. For the inner self-adaption， the solution can reach to true value in the way of any precision，and the BP method can reflect the essence of social or natural phenomenon. Li [2] provided the BP theory to forecast the U.S. population, which expressed the superiority of estimated population. Jiang et al. [3] constructed a multi-index grey PSO-BP neural network population prediction model when forecasting and analyzing China's total population, aiming at the problem that the prediction accuracy of a single index was not high, and experiments showed that the model had high prediction and accuracy. Ren et al. [4] proposed the ARMA model of time series to estimate the future fertility rate and mortality rate, and constructed the Leslie matrix accordingly. According to the truncation or trailing property of autocorrelation function and partial autocorrelation function, the ARMA model was recognized, and the random population prediction method combined with the two methods was proven to be robust. Yu et al. [5] applied the autoregressive analysis method to determine three predictors and establishes the projection pursuit autoregressive (PPAR) model, respectively. The data fitting accuracy of the modeling sample and the prediction accuracy of the verification sample are significantly higher than that of the GM and Logistic models, which can fully meet the needs of population prediction. Wang et al. [6] combined the population actuarial recurrence and the grey dynamic GM(1,1) model to forecast the population distribution of urban pension insurance in Shanghai. The conclusion of their study has important theoretical significance and application value for the estimation of pension fund income and expenditure gap and the formulation of new population policy. Li [7] analyzed the GM(1,1) modeling idea and combined the advantages of the equal dimension grey number replacement method, and applied the improved dynamic GM(1,1) model to forecast the total population of China from 2013 to 2022. Men et al. [8] explored a grey dynamic prediction model to forecast the population of China in the next 50 years. Wang et al. [9] established the Logistic two-population mortality model to predict the aging population in China, and the results manifested that the proposed model was superior to the single-population CBD model, and provided relevant suggestions for the old-age service. Aiming at the defect of low prediction accuracy of the traditional grey model, Guo et al. [10] proposed the adaptive filtering method to correct its residual error and forecast the floating population in Nanjing. The results indicated that the improved model had high prediction accuracy. Basing on chaotic operators, Zou and Xiu [11] suggested a prediction model which is composed of weighted summation of multiple chaotic operators to realize the effective prediction of population data in China. To optimize some critical parameters of the system dynamics model to reduce the subjectivity of the model construction, Li [12] offered an urban population prediction model based on a multi-objective lioness optimization algorithm and system dynamics. Moreover, population size of Xi’an from 2019 to 2050 was predicted by the model. Hou et al. [13] built a combined prediction model with GM(1,N) and SVM. The results demonstated that the combined model was more accurate and more stable than the single model. Hao and Wang [14] denoted the grey dynamic model and its application in population prediction. Wang et al. [15] explored the resident population of Xi'an by using the grey-weighted Markov prediction model. They found that the grey model had a good effect in predicting the change trend of the resident population, but due to the random volatility of the data, the prediction ability of the grey model needed to be improved. On the basis of the above research results, many scholars have carried out a large number of research models for population prediction; in fact, these models are easy to produce random errors. A NAR neural network Markov model (NARMKM) was proposed in order to capture the nonlinear trend analysis of the annual resident population data of Huizhou and obtain higher prediction accuracy.

Although NAR neural network has strong advantages in processing short-term data in time series, it has certain limitations that make it difficult to accurately make long-term predictions and mitigate the effects of random errors, especially in capturing relationships in large data sets when nonlinear. To overcome these limitations and improve prediction accuracy, we propose the NARMKM model to capture nonlinear trends. Thus, we improve the accuracy to forecast the annual resident population data of Huizhou and various population data of Guangdong province. The major contributions of this study can be summarized as follows:

1) NARMKM model was established based on NAR neural network. The optimal delay order of NARMKM model is determined by minimizing the average relative error and reasonably determining the number of hidden layer neurons.

2) The correlation coefficient R between the output data of the training set and the original data and the Ljung-Box Q test of the training error are used to determine whether the trained neural network is reliable.

3) The specific expressions of the estimated and predicted values of the NARMKM model are

constructed based on the Markov transition probability matrix and state division.

4) The validity of the model is verified by numerical examples, and the model is applied to predict the annual resident population of Huizhou. The proposed model and five comparison models are analyzed. The validity and robustness of each model are expressed by the statistics R^2 and STD.

5) In order to verify the practicability and robustness of the model, the model was applied to the permanent resident population and elderly population in Guangdong Province, and the application range of the model was extended.

This paper is organized as follows: Section 2 introduces NARMKM model and NAR neural network algorithm. It also illustrates the Markov model using modified predictions. Section 3 shows the application of the model and gives the results of using NAR neural network model and Markov model to estimate and predict the energy of the resident population. In section 4, the NARMKM model is applied to the estimation and prediction of permanent resident population and elderly population of Guangdong Province. Finally, the fifth section summarizes the full text and discusses the prospect of this paper.

2. The NAR neural network Markov model

2.1. The NAR neural network

According to the existence of feedback and memory, artificial neural networks can be divided into static neural networks and dynamic neural networks. Dynamic neural network has feedback and memory. Its output at every moment is based on the development and change law of system variables before the current moment, and it is a continuous deduction of the development of system variables. Dynamic neural networks can be divided into two categories according to the method of dynamic realization: One is the regression neural network composed of static neurons and the output feedback of the network, such as NAR dynamic regression neural network; the other is the neural network composed by neuronal feedback, such as the PID neural network, which is more commonly used. The variable relationship between the input and output of the model is not only a static mapping, but the output of each moment is synthesized based on the dynamic results of the system before the current moment, that is, it has the function of feedback and memory. The neural network is characterized by both dynamic and complete system information. An artificially generated neural network, which is established by simulating the human nervous system, can achieve specific functions. The network is constructed based on the structure connecting various neurons in the brain [16]. The network's input and output layers are determined by the background of the problem under study. The adjustment of the hidden layer and the number of neurons in the layer depends on the prediction results' accuracy [17]. During data learning and training, the NAR neural network discovers potential patterns and minimizes the difference between the final predicted outputs and actual records through a continuous and iterative fitting process [18]. Therefore, NAR neural network not only inherits the advantages of traditional time series model, but also has better adaptability and prediction effect for nonlinear data [19–24]. The model of NAR neural network can be defined as [25]:

$$y(t) = g\big(y(t-1), y(t-2), \cdots, y(t-d)\big) + \varepsilon_{t1}. \tag{1}$$

Where, $y(t)$ is the variable value of the current moment, $y(t-1), y(t-2), \cdots, y(t-d)$ are the variable values of the historical moment, d is the delay order, and ε_{t1} is the random error term.

A complete NAR neural network is generally composed of input layer, hidden layer and output layer. Data $y(t)$ is input from the input layer, enters the hidden layer, reaches the output layer after training, transmission and learning, and then obtains the expected result, as shown in Figure 1 below:

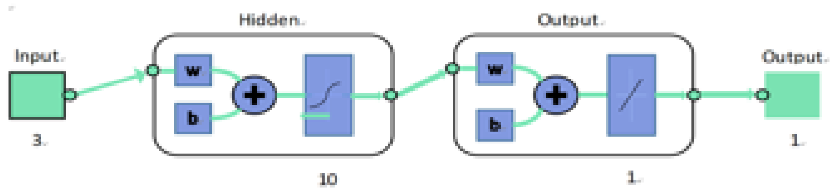

Figure 1. NAR neural network.

The output of each neuron can be expressed as:

$$H_i = g\left(\sum_{i=1}^{n} \omega_{ij}\, y_i + b_i\right) + \varepsilon_{t2}. \tag{2}$$

In Eq 2, g is the activation function, ω_{ij} is the connection weight between the ith output delay signal and the jth neuron in the hidden layer, and b_i is the threshold of the ith output delay signal, and ε_{t2} is the random error term. Our research ideas are shown in the following flow chart (Figure 2):

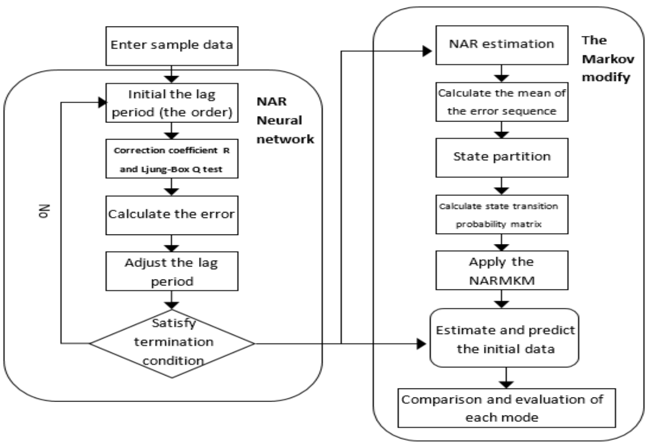

Figure 2. Model algorithm flow chart.

2.2. The Markov model

Markov process is a kind of stochastic process. Its original model, the Markov chain, was proposed by the Russian mathematician A.A. Markov in 1907. The Markov process is an important random process, which combines the initial probability of different states and the transition probability between states to determine the change of its future state, which is suitable for processing data with volatility and no after-effect. In the dynamic process with large random fluctuations, the Markov modified model can adjust the change degree of the predicted value according to the transition probability between states. A large amount of literatures study Markov model and their applications in energy, fiscal expenditure and economy [26–31]. The model is suitable to be applied to analyze the fluctuation data. It has been widely applied in military, biology, meteorology and so on [32–34]. The steps of the Markov model are described below:

Let P be a one-step state transition probability matrix, and the state probability vectors at time t and $t+1$ are Q_t and Q_{t+1}, respectively. The Markov model can be expressed as:

$$Q_{t+1} = PQ_t. \tag{3}$$

The prediction relative error of NAR model can be divided into several states. The number of states is determined according to the actual situation of the data. The state interval is divided as follows:

$$G_i = E_i(G_{i1}, G_{i2}], i = 1,2,\cdots,s. \tag{4}$$

Where, G_{i1} and G_{i2} are the lower and upper limits of the relative errors of the state intervals, and s is the number of intervals.

P_{ij} is the one-step transition probability of the state G_i moving to G_j over a period of time. The one-step transition probabilities of all states form the matrix P, thus:

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1r} \\ \vdots & \ddots & \vdots \\ P_{r1} & \cdots & P_{rr} \end{bmatrix}. \tag{5}$$

Where, $k = 1,2,\cdots,s$, the k-step transition probability matrix is:

$$P^{(k)} = P^{(k-1)}P = P^k. \tag{6}$$

Select the s groups of data closest to the predicted data, and determine the corresponding step q in order from near to far. The row vector of the q-step transition state matrix corresponding to each data is selected to form a new matrix, and the most likely state of the predicted data is determined by the maximum sum of each column vector of the new matrix, so as to determine the corresponding state interval, the center point of which is the Markov correction value, and the revised predicted value is

$$y_{NARMKM}(t) = \frac{\hat{y}_{NAR}(t)}{\frac{G_{i1}+G_{i2}}{2}+1}. \tag{7}$$

2.3. The test of the model error

In this section, the mean absolute percentage error ($MAPE$) and the root mean square error ($RMSE$) are used to evaluate model errors, respectively. By reference [35], we calculate the statistics STD and R^2, and obtain them as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|\hat{y}(t)-y(t)|}{y(t)}, \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{y}(t)-y(t))^2}, \tag{9}$$

$$STD = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\frac{|\hat{y}(t)-y(t)|}{y(t)} - MAPE)^2}, \tag{10}$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(\hat{y}(t)-y(t))^2}{\sum_{t=1}^{n}(\hat{y}(t)-\bar{y}(t))^2}. \tag{11}$$

Where, $\bar{y}$ is the mean value of the training data, and $\bar{y}(t) = \sum_{t=1}^{n} y(t)$.

3. The NAR neural network Markov model

3.1. Estimation and prediction of the NAR model

The NAR neural network was established using the data of 32 Huizhou resident population (unit: 10000) from 1990 to 2021. First, determine the order, that is, the lag period. If the order is too small, the application of historical data information is insufficient, and too many orders are susceptible to distortions in historical data. When the order is three, the first three values of the data in the time series are used to predict the value of the data and the prediction is the best. Second, the weights and thresholds are generated randomly. The network is trained with 70% sample data, cross-checked with 15% data, and tested with 15% data for network performance. The number of hidden layer neurons is 10. The maximum number of iterations is 1000. In the training stage of NAR neural network, the correlation coefficient R between the output data of training set and the original data and the Ljung-Box Q test according to the training error are used to judge whether the trained neural network is reliable. The training error of NAR neural network model is shown in Figure 3.

As can be seen from Figure 4 below, after 14 steps of network automatic training learning, the fitting variance of step 8 is the smallest.

It can be seen from Figure 5 that training data, cross-check data and test data all have a high degree of fit with the corresponding original data, and the correlation coefficient R between them and the original data is greater than 0.99. The result of the Ljung-Box Q test is 0, which indicates that the NAR neural network is reliable and can predict the population data.

The forecast curve is shown in Figure 6. The forecast zigzag shows a gentle upward trend as a whole, and is relatively smooth, which is basically consistent with the trend characteristics of the actual value, and reflects that the prediction results of the model have high reliability. The estimation results of NAR neural network on Huizhou resident population data are shown in Table 1.

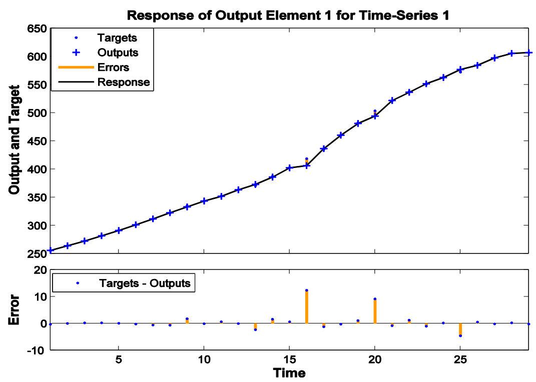

Figure 3. Training error of NAR neural network model.

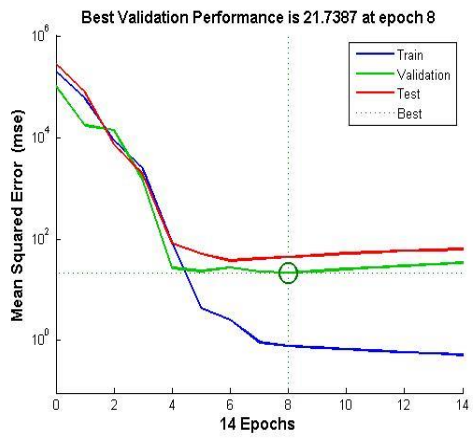

Figure 4. Network training and test results.

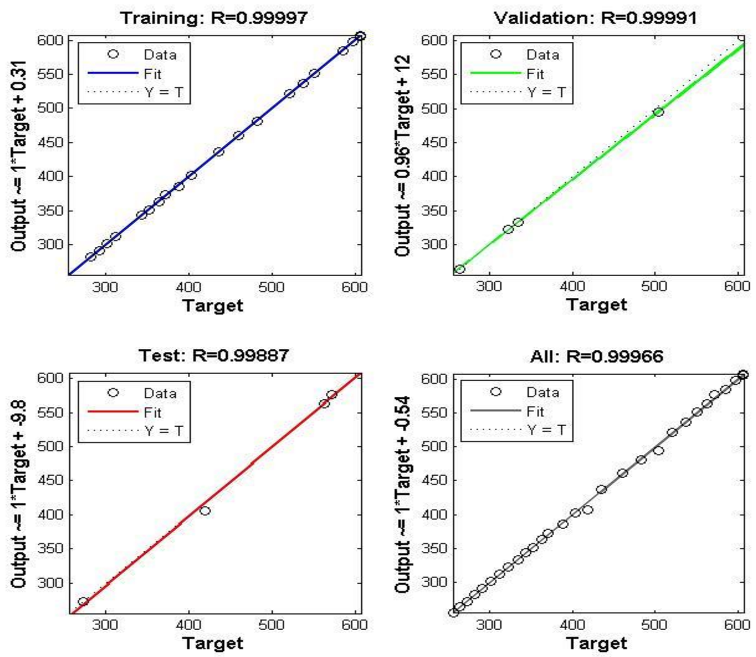

Figure 5. Neural network regression results.

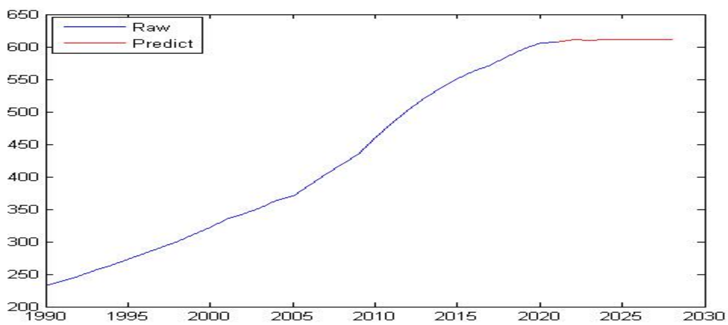

Figure 6. Prediction curve of the neural network.

Table 1. The comparison results between GM(1,1), FGM(1,1), ARIMA and NAR.

Year	Raw	NAR		FGM(1,1)		ARIMA		GM (1,1)	
		Predict value	Relative error	Predict value	Relative error	Predict value	Relative error	Predict value	Relative error
1990	231.25	231.25	0.00%	231.25	0.00%	231.20	-0.02%	231.25	-0.02%
1991	239.01	239.01	0.00%	237.72	-0.54%	239.21	0.08%	233.23	0.08%
1992	247.02	247.02	0.00%	245.55	-0.60%	246.71	-0.13%	241.67	-0.13%
1993	255.31	256.05	0.29%	254.02	-0.51%	255.29	-0.01%	250.42	-0.01%
1994	263.87	264.40	0.20%	262.95	-0.35%	263.87	0.00%	259.49	0.00%
1995	272.72	273.05	0.12%	272.28	-0.16%	272.70	-0.01%	268.89	-0.01%
1996	281.87	282.15	0.10%	282.00	0.05%	281.84	-0.01%	278.63	-0.01%
1997	291.32	291.84	0.18%	292.08	0.26%	291.30	-0.01%	288.72	-0.01%
1998	301.09	302.12	0.34%	302.53	0.48%	301.05	-0.01%	299.17	-0.01%
1999	311.19	312.77	0.51%	313.35	0.69%	311.15	-0.01%	310.00	-0.01%
2000	321.80	323.46	0.52%	324.54	0.85%	321.59	-0.07%	321.23	-0.07%
2001	334.77	334.04	-0.22%	336.12	0.40%	332.83	-0.58%	332.86	-0.58%
2002	343.41	345.73	0.68%	348.10	1.37%	348.04	1.35%	344.92	1.35%
2003	352.27	354.73	0.70%	360.47	2.33%	352.14	-0.04%	357.41	-0.04%
2004	363.18	362.48	-0.19%	373.27	2.78%	361.30	-0.52%	370.35	-0.52%
2005	370.69	372.66	0.53%	386.48	4.26%	374.20	0.95%	383.76	0.95%
2006	387.51	381.44	-1.57%	400.14	3.26%	378.50	-2.33%	397.66	-2.33%
2007	402.86	398.45	-1.09%	414.26	2.83%	404.87	0.50%	412.06	0.50%
2008	418.65	418.85	0.05%	428.83	2.43%	418.68	0.01%	426.98	0.01%
2009	435.08	438.81	0.86%	443.89	2.02%	434.92	-0.04%	442.44	-0.04%
2010	460.11	458.47	-0.36%	459.45	-0.14%	452.21	-1.72%	458.46	-1.72%
2011	482.21	481.08	-0.23%	475.52	-1.39%	485.95	0.78%	475.06	0.78%
2012	503.44	503.13	-0.06%	492.11	-2.25%	504.99	0.31%	492.27	0.31%
2013	521.06	520.51	-0.11%	509.25	-2.27%	525.21	0.80%	510.09	0.80%
2014	537.66	537.10	-0.10%	526.95	-1.99%	539.10	0.27%	528.57	0.27%
2015	550.41	550.81	0.07%	545.23	-0.94%	554.56	0.75%	547.71	0.75%
2016	562.73	564.40	0.30%	564.12	0.25%	563.35	0.11%	567.54	0.11%
		RMSE	1.9418	RMSE	7.1771	RMSE	2.9951	RMSE	6.4415
		MAPE	0.35%	MAPE	1.31%	MAPE	0.42%	MAPE	1.43%
		STD	0.0036	STD	0.0114	STD	0.0058	STD	0.0086
		R^2	0.9996	R^2	0.9949	R^2	0.9991	R^2	0.9960
2017	572.22	573.99	0.31%	583.62	1.99%	575.23	0.53%	588.09	2.77%
2018	584.72	583.62	-0.19%	603.75	3.25%	587.92	0.55%	609.39	4.22%
2019	597.23	590.48	-1.13%	624.55	4.57%	600.78	0.59%	631.46	5.73%
2020	605.72	600.45	-0.87%	646.03	6.65%	613.83	1.34%	654.32	8.02%
2021	606.60	608.39	0.30%	668.20	10.15%	627.06	3.37%	678.02	11.77%
		RMSE	4.0220	RMSE	36.4909	RMSE	10.1616	RMSE	43.5774
		MAPE	0.56%	MAPE	5.33%	MAPE	1.28%	MAPE	6.50%
		STD	0.0037	STD	0.0587	STD	0.0109	STD	0.0316
		R^2	0.8926	R^2	0.3043	R^2	0.7383	R^2	0.2491

In order to better reflect the superiority of the NAR model, we choose the grey model (GM(1,1)), fractional grey model (FGM(1,1)), Multiple regression and ARIMA model as comparison models. Each model estimates the resident population data of Huizhou from 1990 to 2016 and forecasts the original data from 2017 to 2021. It can be seen from Table 1 that the values of MAPE, RMSE and STD of the NAR model are smaller than those of the other three models. The R^2 value of the NAR model is larger than that of the other three models. This shows that the NAR model is more efficient.

3.2. Markov model modified Huizhou resident population data

The state interval is divided according to the relative error of NAR model. As can be seen from Table 1, the minimum and maximum relative error of the first 27 fitting data of this model are -1.57% and 0.87%. Therefore, according to the rule of equal spacing, four state intervals are divided into:

$$E_1(-1.57\%, -0.96\%],\ E_2(-0.96\%, -0.0.35\%],\ E_3(-0.35\%, 0.26\%],\ E_4(0.26\%, 0.87\%].$$

From the above state intervals and the probability of moving from the current state to the next state, the following 1, 2, 3, 4 steps state transition probability matrixes are obtained:

$$P^{(1)} = \begin{bmatrix} 0.5000 & 0.0000 & 0.5000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.6000 & 0.4000 \\ 0.1250 & 0.1250 & 0.3750 & 0.3750 \end{bmatrix}, P^{(2)} = \begin{bmatrix} 0.2500 & 0.0000 & 0.5500 & 0.2000 \\ 0.0000 & 0.0000 & 0.6000 & 0.4000 \\ 0.0500 & 0.0500 & 0.5100 & 0.3900 \\ 0.1094 & 0.0469 & 0.5531 & 0.2906 \end{bmatrix},$$

$$P^{(3)} = \begin{bmatrix} 0.1500 & 0.0250 & 0.5300 & 0.2950 \\ 0.0500 & 0.0500 & 0.5100 & 0.3900 \\ 0.0738 & 0.0488 & 0.5273 & 0.3503 \\ 0.0910 & 0.0363 & 0.5424 & 0.3302 \end{bmatrix}, P^{(4)} = \begin{bmatrix} 0.1119 & 0.0369 & 0.5286 & 0.3226 \\ 0.0737 & 0.0488 & 0.5273 & 0.3503 \\ 0.0807 & 0.0438 & 0.5333 & 0.3422 \\ 0.0868 & 0.0413 & 0.5531 & 0.3408 \end{bmatrix}.$$

Constructed a new state transition matrix using the most recent sets of data, we get the state of 2017, which is listed in Table 2.

Table 2. The prediction status of 2017.

Year	Initial status	Transferring steps	P_{ij}	E_1	E_2	E_3	E_4
2016	4	1	P_{14}	0.1250	0.1250	0.3750	0.3750
2015	3	2	P_{23}	0.0500	0.0500	0.5100	0.3900
2014	3	3	P_{33}	0.0738	0.0488	0.5273	0.3503
2013	3	4	P_{43}	0.0807	0.0438	0.5333	0.3422
Total				0.3295	0.2676	1.9456	1.4575

From Table 2, it can be concluded that the most likely state of Huizhou resident population in 2017 is E_3, because E_3 has the largest value in the total. The predicted value of NAR model in 2017 is 573.99, and the predicted value of Markov model is 574.24 according to formula (7). In terms of the same method, the predicted value of Markov model for 2018–2021 can be obtained, and the specific results are shown in Table 3.

Table 3. The comparison results between NAR, Multiple regression, FANGBM and NARMKM.

Year	Raw	NAR			FANGBM		Multiple regression		NARMKM	
		Predict value	Relative error	Stat-ion	Predict value	Relative error	Predict value	Relative error	Predict value	Relative error
1990	231.25	231.25	0.00%	3	231.25	0.00%	226.42	-2.09%	231.35	0.04%
1991	239.01	239.01	0.00%	3	230.71	-3.47%	234.80	-1.76%	239.11	0.04%
1992	247.02	247.02	0.00%	3	240.10	-2.80%	243.51	-1.42%	247.13	0.04%
1993	255.31	256.05	0.29%	4	249.43	-2.30%	252.53	-1.09%	254.61	-0.27%
1994	263.87	264.40	0.20%	3	258.91	-1.88%	261.89	-0.75%	264.51	0.24%
1995	272.72	273.05	0.12%	3	268.61	-1.51%	271.56	-0.43%	273.17	0.17%
1996	281.87	282.15	0.10%	3	278.58	-1.17%	281.57	-0.11%	282.27	0.14%
1997	291.32	291.84	0.18%	3	288.85	-0.85%	291.89	0.20%	291.97	0.22%
1998	301.09	302.12	0.34%	4	299.44	-0.55%	302.54	0.48%	300.42	-0.22%
1999	311.19	312.77	0.51%	4	310.39	-0.26%	313.51	0.75%	311.01	-0.06%
2000	321.80	323.46	0.52%	4	321.69	-0.03%	324.81	0.94%	321.64	-0.05%
2001	334.77	334.04	-0.22%	3	333.38	-0.42%	336.43	0.50%	334.19	-0.17%
2002	343.41	345.73	0.68%	4	345.47	0.60%	348.37	1.44%	343.78	0.11%
2003	352.27	354.73	0.70%	4	357.97	1.62%	360.64	2.38%	352.73	0.13%
2004	363.18	362.48	-0.19%	3	370.90	2.13%	373.23	2.77%	362.64	-0.15%
2005	370.69	372.66	0.53%	4	384.29	3.67%	386.15	4.17%	370.56	-0.04%
2006	387.51	381.44	-1.57%	1	398.14	2.74%	399.39	3.07%	386.32	-0.31%
2007	402.86	398.45	-1.09%	1	412.47	2.39%	412.95	2.50%	403.55	0.17%
2008	418.65	418.85	0.05%	3	427.31	2.07%	426.84	1.96%	419.03	0.09%
2009	435.08	438.81	0.86%	4	442.67	1.74%	441.05	1.37%	436.34	0.29%
2010	460.11	458.47	-0.36%	2	458.57	-0.33%	455.58	-0.98%	461.49	0.30%
2011	482.21	481.08	-0.23%	3	475.03	-1.49%	470.44	-2.44%	481.29	-0.19%
2012	503.44	503.13	-0.06%	3	492.07	-2.26%	485.62	-3.54%	503.35	-0.02%
2013	521.06	520.51	-0.11%	3	509.71	-2.18%	501.13	-3.82%	520.74	-0.06%
2014	537.66	537.10	-0.10%	3	527.97	-1.80%	516.96	-3.85%	537.34	-0.06%
2015	550.41	550.81	0.07%	3	546.88	-0.64%	533.12	-3.14%	551.05	0.12%
2016	562.73	564.40	0.30%	4	566.46	0.66%	549.59	-2.34%	561.22	-0.27%
		RMSE	1.9418		RMSE	6.8492	RMSE	9.8889	RMSE	0.6803
		MAPE	0.35%		MAPE	1.54%	MAPE	1.86%	MAPE	0.15%
		STD	0.0036		STD	0.0101	STD	0.0119	STD	0.0009
		R^2	0.9996		R^2	0.9955	R^2	0.9897	R^2	0.9999
2017	572.22	573.99	0.31%	4	586.74	2.54%	566.40	-1.02%	574.24	0.35%
2018	584.72	583.62	-0.19%	3	607.73	3.94%	583.52	-0.21%	583.88	-0.14%
2019	597.23	590.48	-1.13%	1	629.46	5.40%	600.97	0.63%	590.74	-1.09%
2020	605.72	600.45	-0.87%	2	651.97	7.64%	618.74	2.15%	600.72	-0.83%
2021	606.60	608.39	0.30%	4	675.27	11.32%	636.84	4.99%	608.66	0.34%
		RMSE	4.0220		RMSE	41.5540	RMSE	15.0551	RMSE	3.9026
		MAPE	0.56%		MAPE	6.17%	MAPE	1.80%	MAPE	0.55%
		STD	0.0037		STD	0.0308	STD	0.0172	STD	0.0035
		R^2	0.8926		R^2	0.2634	R^2	0.6687	R^2	0.8984

It can be seen from the results in Table 3 that the statistical values of MAPE, RMSE and STD predicted by NARMKM model are smaller than those of other models. The R^2 value of NARMKM model is larger than that of other models. This shows that NARMKM model is more effective.

The predicted resident population of Huizhou in 2022–2028 is 610.59, 610.00, 611.56, 610.72, 611.29, 610.68 and 610.95 (unit: 10000).

4. Research on the prediction of permanent resident population and elderly population in Guangdong

In the same way as Part 3, NAR model and NARMKM model are established to predict the permanent resident population and the elderly population in Guangdong. The results are shown in Tables 4 and 5.

Table 4. The comparison results between NAR, GM(1,1), FANGBM and NARMKM of resident population in Guangdong.

Year	Raw	NAR			FANGBM		GM (1,1)		NARMKM	
		Predict value	Relative error	Stat-ion	Predict value	Relative error	Predict value	Relative error	Predict value	Predict value
1990	6347.19	6347.19	0.00%	5	6347.19	0.00%	6347.19	0.00%	6341.98	-0.08%
1991	6527.01	6527.01	0.00%	5	6478.65	-0.74%	6755.51	3.50%	6521.66	-0.08%
1992	6706.45	6647.81	-0.87%	4	6747.47	0.61%	6912.28	3.07%	6684.84	-0.32%
1993	6936.69	6796.66	-2.02%	2	6976.82	0.58%	7072.68	1.96%	6923.07	-0.20%
1994	7209.58	7039.41	-2.36%	1	7188.77	-0.29%	7236.80	0.38%	7217.09	0.10%
1995	7387.49	7341.64	-0.62%	4	7391.78	0.06%	7404.74	0.23%	7382.53	-0.07%
1996	7569.78	7518.71	-0.67%	4	7590.04	0.27%	7576.57	0.09%	7560.59	-0.12%
1997	7779.69	7729.80	-0.64%	4	7785.94	0.08%	7752.38	-0.35%	7772.86	-0.09%
1998	7990.03	7969.81	-0.25%	4	7981.02	-0.11%	7932.28	-0.72%	8014.20	0.30%
1999	8217.91	8188.58	-0.36%	4	8176.29	-0.51%	8116.35	-1.24%	8234.19	0.20%
2000	8650.03	8409.52	-2.78%	1	8372.50	-3.21%	8304.69	-3.99%	8621.78	-0.33%
2001	8733.18	8730.20	-0.03%	5	8570.21	-1.87%	8497.41	-2.70%	8723.04	-0.12%
2002	8842.08	8861.49	0.22%	5	8769.84	-0.82%	8694.59	-1.67%	8854.22	0.14%
2003	8962.69	8967.74	0.06%	5	8971.76	0.10%	8896.36	-0.74%	8960.39	-0.03%
2004	9110.66	9085.50	-0.28%	4	9176.24	0.72%	9102.80	-0.09%	9136.11	0.28%
2005	9194.00	9230.93	0.40%	5	9383.55	2.06%	9314.03	1.31%	9223.36	0.32%
2006	9442.07	9336.09	-1.12%	3	9593.90	1.61%	9530.17	0.93%	9448.52	0.07%
2007	9659.52	9538.45	-1.25%	3	9807.48	1.53%	9751.32	0.95%	9653.32	-0.06%
2008	9893.48	9813.36	-0.81%	4	10024.47	1.32%	9977.60	0.85%	9868.02	-0.26%
2009	10130.19	10086.05	-0.44%	4	10245.03	1.13%	10209.14	0.78%	10142.23	0.12%
2010	10440.94	10369.31	-0.69%	4	10469.31	0.27%	10446.04	0.05%	10427.07	-0.13%
2011	10756.00	10618.58	-1.28%	3	10697.46	-0.54%	10688.45	-0.63%	10746.46	-0.09%
2012	11041.00	10905.43	-1.23%	3	10929.62	-1.01%	10936.48	-0.95%	11108.26	0.61%
2013	11270.00	11200.51	-0.62%	4	11165.91	-0.92%	11190.26	-0.71%	11262.90	-0.06%

Continued on next page

Year	Raw	NAR			FANGBM		GM (1,1)		NARMKM	
		Predict value	Relative error	Stat-ion	Predict value	Relative error	Predict value	Relative error	Predict value	Predict value
2014	11489.00	11467.39	-0.19%	5	11406.47	-0.72%	11449.94	-0.34%	11457.99	-0.27%
2015	11678.00	11678.51	0.00%	5	11651.42	-0.23%	11715.64	0.32%	11668.54	-0.08%
2016	11908.00	11894.54	-0.11%	5	11900.89	-0.06%	11987.50	0.67%	11884.79	-0.19%
2017	12141.00	12087.46	-0.44%	4	12154.99	0.12%	12265.68	1.03%	12145.74	0.04%
		RMSE	86.715		RMSE	97.832	RMSE	122.503	RMSE	20.420
		MAPE	0.71%		MAPE	0.77%	MAPE	1.08%	MAPE	0.17%
		STD	0.007		STD	0.0074	STD	0.0104	STD	0.0013
		R^2	0.9975		R^2	0.9967	R^2	0.949	R^2	0.9999
2018	12348.00	12307.41	-0.33%	4	12413.85	0.53%	12550.31	-1.02%	12375.97	0.23%
2019	12489.00	12472.61	-0.13%	5	12677.59	1.51%	12841.54	-0.21%	12542.09	0.43%
2020	12624.00	12573.43	-0.40%	4	12946.34	2.55%	13139.53	0.63%	12643.47	0.15%
2021	12684.00	12548.15	-1.07%	3	13220.21	4.23%	13444.44	2.15%	12618.05	-0.52%
2022	12656.80	12577.78	-0.62%	4	13499.32	6.66%	13756.42	4.99%	12647.84	-0.07%
		RMSE	76.384		RMSE	477.737	RMSE	666.095	RMSE	41.013
		MAPE	0.51%		MAPE	3.10%	MAPE	4.56%	MAPE	0.28%
		STD	0.0032		STD	0.0216	STD	0.0248	STD	0.0017
		R^2	0.5965		R^2	0.2398	R^2	0.1555	R^2	0.8390

Table 5. The comparison results between NAR, GM(1,1), FANGBM and NARMKM of elderly population in Guangdong.

Year	Raw	NAR			FANGBM		GM (1,1)		NARMKM	
		Predict value	Relative error	Stati -on	Predict value	Relative error	Predict value	Relative error	Predict value	Relative error
2002	603.33	603.33	0.00%	2	603.33	0.00%	603.33	0.00%	606.18	0.47%
2003	610.15	610.15	0.00%	2	609.56	-0.10%	613.54	0.56%	613.03	0.47%
2004	656.82	548.37	-16.51%	1	635.13	-3.30%	636.84	-3.04%	617.28	-6.02%
2005	681.28	656.90	-3.58%	2	660.59	-3.04%	661.03	-2.97%	660.01	-3.12%
2006	698.71	715.56	2.41%	2	686.49	-1.75%	686.14	-1.80%	718.94	2.90%
2007	734.12	737.61	0.48%	2	713.04	-2.87%	712.20	-2.99%	741.1	0.95%
2008	781.58	793.34	1.50%	2	740.38	-5.27%	739.25	-5.42%	797.09	1.98%
2009	810.42	869.10	7.24%	3	768.57	-5.16%	767.33	-5.32%	783.96	-3.26%
2010	708.62	894.67	26.26%	4	797.70	12.57%	796.47	12.40%	739.75	4.39%
2011	731.41	746.60	2.08%	2	827.81	13.18%	826.72	13.03%	750.13	2.56%
2012	772.87	787.50	1.89%	2	858.97	11.14%	858.12	11.03%	791.2	2.37%
2013	920.76	857.00	-6.92%	1	891.22	-3.21%	890.72	-3.26%	964.69	4.77%
2014	950.14	938.89	-1.18%	2	924.61	-2.69%	924.55	-2.69%	943.33	-0.72%
2015	990.29	946.89	-4.38%	2	959.19	-3.14%	959.66	-3.09%	951.37	-3.93%

Continued on next page

Year	Raw	NAR		Stati-on	FANGBM		GM (1,1)		NARMKM	
		Predict value	Relative error		Predict value	Relative error	Predict value	Relative error	Predict value	Relative error
2016	1018.13	970.36	-4.69%	2	995.01	-2.27%	996.11	-2.16%	974.95	-4.24%
2017	1046.55	1003.7	-4.09%	2	1032.11	-1.38%	1033.95	-1.20%	1008.45	-3.64%
		RMSE	61.991		RMSE	45.289	RMSE	44.941	RMSE	27.264
		MAPE	5.20%		MAPE	4.44%	MAPE	4.44%	MAPE	2.86%
		STD	0.0670		STD	0.0404	STD	0.0397	STD	0.0159
		R^2	0.7915		R^2	0.8874	R^2	0.8889	R^2	0.9584
2018	1064.40	1041.59	-2.14%	2	1070.56	0.58%	1073.22	0.83%	1046.52	-1.68%
2019	1081.30	1088.85	0.70%	2	1110.40	2.69%	1113.98	3.02%	1094.01	1.18%
2020	1124.01	1124.69	0.06%	2	1151.69	2.46%	1156.29	2.87%	1130.02	0.53%
2021	1157.00	1146.26	-0.93%	2	1194.48	3.24%	1200.21	3.73%	1151.69	-0.46%
2022	1214.40	1234.79	1.68%	2	1238.83	2.01%	1245.80	2.59%	1240.64	2.16%
		RMSE	14.892		RMSE	27.027	RMSE	31.751	RMSE	15.710
		MAPE	1.10%		MAPE	2.20%	MAPE	2.61%	MAPE	1.20%
		STD	0.0073		STD	0.0090	STD	0.0097	STD	0.0066
		R^2	0.9406		R^2	0.8245	R^2	0.7817	R^2	0.9413

As can be seen from the prediction results in Table 4, the NARMKM model's statistical values of MAPE, RMSE and STD are smaller than those of NAR, FANGBM and GM(1,1) models. The R^2 value of NARMKM model is larger than the corresponding values of these three models. This shows that NARMKM model is more efficient. The results of Figure 7 show that the curves of the NARMKM model are closer to the true values than those of the NAR model. The forecast permanent population of Guangdong from 2023 to 2026 is 12742.92, 12828.69, 12806.52 and 12800.32 (unit: 10000).

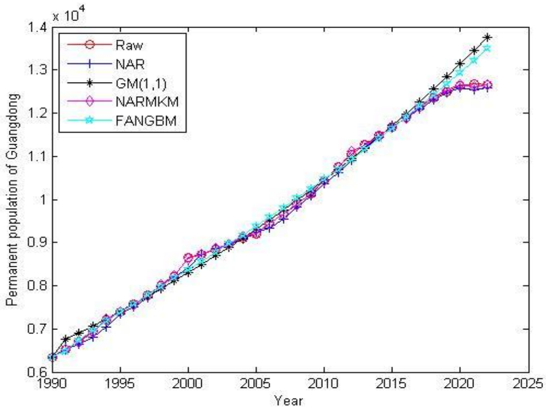

Figure 7. The forecast results of different models for permanent population of Guangdong.

Based on the elderly population data of Guangdong Province, it can be seen from the prediction results in Table 5 that the statistical values of MAPE, RMSE and STD of NARMKM model are smaller than those of GM(1,1) model, but basically the same as those of NAR model. The R^2 value of

NARMKM model is larger than the corresponding values of these three models. From the estimation results, NARMKM model is better than NAR, FANGBM and GM(1,1) model. This indicates that NARMKM model is more effective. The results of Figure 8 testify that the curve of the NARMKM model is closer to the true values than those of the NAR model.

The projected values of the elderly population in Guangdong from 2023 to 2026 is: 1269.55, 1266.91, 1259.83 and 1259.16 (unit: 10000).

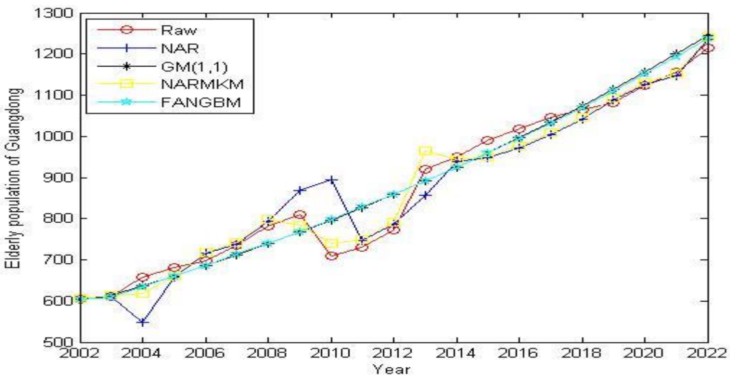

Figure 8. The forecast results of different models for the elderly population of Guangdong.

5. Conclusions

We construct a NAR neural network Markov model (NARMKM). The NAR neural network model is provided to estimate and forecast the resident population of Huizhou, the resident population and the elderly population of Guangdong. It does not need too much preprocessing operation on the original time series, and the convergence speed is fast and easy to use. The prediction results of NAR neural network model are greatly affected by the delay order and the number of hidden layer neurons. Therefore, we should make full use of prior information to constrain and test when making prediction. We choose reasonable parameter settings to obtain more reliable prediction results. NARMKM model is compared with NAR model. The results show that the NARMKM model has better fitting effect and estimation accuracy than the other six competing models, including FGM model, GM, multiple linear model, NAR, FANGBM and ARIMA model. Finally, the NARMKM model is used to predict some kinds of population, and the results show that the new model has more accurate and effective prediction and evaluation effect than the NAR model. According to the comparison between the predicted result and the actual population, the predicted result has a high accuracy. From the forecast results after 2023, the growth trend of these populations is not large.

In order to improve the prediction accuracy, we plan to further study the BP neural network Markov model and its application. We will use the general regression neural network Markov model (GRNNMKM) to deal with the unstable data. It is hoped that further progress can be made.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by the NSF of Huizhou University (Grant No. hzu201806), the Project of Guangdong Provincial Department of Education (Grant No. 2021ZDJS080). This work was also supported by the Ministry of Education's "Blue Fire Program" (Huizhou) Industry-University-Research Joint Innovation Fund 2018 Annual Project: High-power Intelligent Dimming Control System (CXZJHZ201812), the Huizhou University School Level Undergraduate Teaching Quality Engineering Project (X-JYJG2021045), the Higher Education Teaching Reform Project of Huizhou University ([2022]163), the Huizhou Philosophy and Social Sciences Discipline Co-Construction Project (2022ZX046, 2023ZX056), as well as the Guangdong Philosophy and Social Sciences Discipline Co-Construction Project "Research on the evolution characteristics and development trends of the permanent resident population, floating population and population aging in Guangdong Province" (GD23XSH27).

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

1. N. Jia, H. P. Hu, Y. P. Bai, Population prediction based on BP neural network, *J. Shandong Uni. Techn. (Natural Science Edition)*, **25** (2011), 22–24. https://doi.org/10.3969/j.issn.1672-6197.2011.03.006
2. Y. Li, U. S. Population forecast based on BP algorithm neural network, in Chinese, *Sci. Technl. Eng.*, **34** (2011), 8665–8667.
3. R. F. Jiang, Y. M. Jiang. F. Y. Li, Study and application of population forecast model based on grey system and PSO-BP neural network, in Chinese, *Northwest Popul.*, **32** (2011), 23–26. https://doi.org/10.15884/j.cnki.issn.1007-0672.2011.03.018
4. Q. Ren, D. D. Hou, Stochastic model for population forecast: based on leslie matrix and arma model, *Popul. Res.*, **35** (2011), 28–41.
5. X. H. Yu, H. Y. Xu, W. G. Lou, Construction and empirical study of projection pursuit autoregressive model of population prediction, in Chinese, *Stat. Decis.*, **23** (2022), 38–42. https://doi.org/10.13546/j.cnki.tjyjc.2022.23.007
6. Y. X. Wang, H. Wang, J. Xiao, Forecast on population distribution of Shanghai pension system based on the gray GM(1,1) model, *Syste. Engin. Theo. Pract.*, **30** (2010), 2244–2253. https://doi.org/10.12011/1000-6788(2010)12-2244
7. F. R. Li, Application of improved dynamic GM(1,1) model to population forecasting, in Chinese, *Stat. Decis.*, **19** (2013), 72–74. https://doi.org/10.13546/j.cki.tjyjc.2013.19.015
8. K. P. Men, W. Zen, Research on population development forecast of China in the next 50 years, *J. Quant. Techn. Econo.*, **3** (2004), 12–17. https://doi.org/10.13653/j.cnki.jqte.2004.03.003
9. X. J. Wang, H. M. Chen, X. Y. Zhao, Joint modeling and prediction of the mortality of male and female aged population in China, in Chinese, *Stat. Res.*, **38** (2021), 151–160. https://doi.org/10.19343/j.cnki.11－1302/c.2021.10.012

10. X. F. Guo, J. Y. Huang, H. Wang, Application of improved grey model in floating population prediction, in Chinese, *Stat. Decis.*, **34** (2018), 76–79. https://doi.org/10.13546/j.cnki.tjyjc.2018.08.018
11. X. M. Zou, C. B. Xiu, Application of chaotic operator model in population prediction, in Chinese, *Stat. Decis.*, **15** (2011), 169–171. https://doi.org/10.13546/j.cnki.tjyjc.2011.15.025
12. D. Li, Y. Y. Yu, B. Wang, Urban population prediction based on multi-objective lioness optimization algorithm and system dynamics model, *Sci. Rep.*, **13** (2023), 11836–11861. https://doi.org/10.1038/S41598-023-39053-1
13. R. H. Hou, X. Y. Xu, Population prediction based on improved multi-dimensional grey model and support vector machine, in Chinese, *Stat. Decis.*, **18** (2021), 41–51. https://doi.org/10.13546/j.cnki.tjyjc.2021.18.009
14. Y. H. Hao, X. M. Wang, The dynamic model of gray system and its application to population forcasting, *Math. Pract. Theory*, **32** (2002), 813–820. https://doi.org/10.3969/j.issn.1000-0984.2002.05.022
15. Z. Y. Wang, X. L. Xu, Z. Hu, F. Ye, Y. F. Wang, Prediction of permanent resident population in Xi'an based on grey-weighted markov prediction model, *Modern Inform. Technol.*, **6** (2022), 118–121. https://doi.org/10.19850/j.cnki.2096-4706.2022.18.029
16. Sunayana, S. Kuma, R. Kumar, Forecasting of municipal solid waste generation using non-linear autoregressive (NAR) neural models, *Waste Manag.,* **121** (2021), 206–214. https://doi.org/10.1016/j.wasman.2020.12.011
17. M. Adil, R. Ullah, S. Noor, N. Gohar, Effect of number of neurons and layers in an artificial neural network for generalized concrete mix design, *Neural Comput. Appl.,* **34** (2020), 8355–8363. https://doi.org/10.1007/s00521-020-05305-8
18. S. Wei, D. Zuo, J. Song, Improving prediction accuracy of river discharge time series using a Wavelet-NAR artificial neural network, *J. Hydro Inform.,* **14** (2012), 974–991. https://doi.org/10.2166/hydro.2012.143
19. Y. T. Bai, X. B. Jin, X. Y. Wang, J. Kong, Y. Lu, Compound autoregressive network for prediction of multivariate time series, *Complexity*, **2019** (2019), 9107167. https://doi.org/10.1155/2019/9107167
20. J. T. Song, Y. C. Chen, J. Yan, A novel outlier detection method of long-term dam monitoring data based on SSA-NAR, *Wirel. Commun. Mobile Comput.*, **2022** (2022), 6569367. https://doi.org/10.1155/2022/6569367
21. F. Ma, Y. Jin, C. Sun, Short-term prediction model of subway passenger flow based on EMD optimized NAR dynamic neural network, *J. Appl. Sci.,* **38** (2020), 936–943. https://doi.org/10.3969/j.issn.0255-8297.2020.06.010
22. M. Lydia, S. SureshKumar, A. I. Selvakumar, G. E. P. Kumar, Linear and non-linear autoregressive models for short-term wind speed forecasting, *Energy Conver. Manage.*, **112** (2016), 115–124. https://doi.org/10.1016/j.enconman.2016.01.007
23. M. S. Tanvir, I. M. Mujtaba, Neural network based correlations for estimating temperature elevation for seawater in MSF desalination process, *Desalination*, **195** (2006), 251–272. https://doi.org/10.1016/j.desal.2005.11.013
24. R. Sarkar, S. Julai, S. Hossain, W. T. Chong, M. Rahman, A comparative study of activation functions of NAR and NARX neural network for long-term wind speed forecasting in Malaysia, *Math. Probl. Engin.*, **2019** (2019), 6403081. https://doi.org/10.1155/2019/6403081

25. G. Benrhmach, K. Namir, A. Namir, J. Bouyaghroumni, Nonlinear autoregressive neural network and extended Kalman filters for prediction of financial time series, *J. Appl. Math.*, **2020** (2020), 5057801. https://doi.org/10.1155/2020/5057801
26. W. C. Fan, Y. Jiang, S. Y. Huang, W. G. Liu, Research and prediction of opioid crisis based on BP neural network and Markov chain, *AIMS Math.*, **4** (2019), 1357–1368. https://doi.org/10.3934/math.2019.5.1357
27. A. Souissi, E. G. Soueidy, M. Rhaima, Clustering property for quantum Markov chains on the comb graph, *AIMS Math.*, **8** (2023), 7865–7880. https://doi.org/10.3934/math.2023396
28. Y. H. Lin, H. Y. Liu, Inverse problems for fractional equations with a minimal number of measurements, *Commun. Anal. Comput.*, **1** (2023), 72–93. https://doi.org/10.3934/cac.2023005
29. D. W. Li, M. L. Qiu, S. P. Yang, C. Wang, Z. L. Luo, An optimal fractional-order accumulative Grey Markov model with variable parameters and its application in total energy consumption, *AIMS Math.*, **8** (2023), 26425–26443. https://doi.org/ 10.3934/math.20231349
30. D. W. Li, D. M. Xu, M. L. Qiu, S. P. Yang, Forecasting the public financial budget expenditure in Dongguan with an optimal weighted combination Markov model, *AIMS Math.*, **8** (2023), 15600–15617. https://doi.org/ 10.3934/math.2023796
31. M. L. Qiu, D. W. Li, Z. L. Luo, X. J. Yu, Huizhou GDP forecast based on fractional opposite-direction accumulating nonlinear grey bernoulli markov model, *ERA.*, **31** (2023), 947–960. https://doi.org/10.3934/era.2023047
32. M. C. Şahingil, R. Yurttaş, The determination of flare launching programs to use against pulse width modulating guided missile seekers via hidden Markov models, In: *2012 20th Signal Processing and Communications Applications Conference (SIU),* 2012. https://doi.org/10.1109/SIU.2012.6204715
33. A. Krogh, B. Larsson, G. H. Von, E. L. L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes, *J. Mol. Biol.*, **305** (2001), 567–580. https://doi.org/10.1006/jmbi.2000.4315
34. M. Thyer, G. Kuczera, Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model, *Water Resour. Res.*, **36** (2000), 3301–3310. https://doi.org/10.1029/2000WR900157
35. M. Gao, H. Yang, Q. Xiao, M. Goh, A novel method for carbon emission forecasting based on Gompertz's law and fractional grey model: evidence from American industrial sector, *Renew. Energy*, **181** (2022), 803–819. https://doi.org/10.1016/j.renene.2021.09.072

AIMS Press