The following publication X. Zhang, K. W. Chan, X. Yang, Y. Zhou, K. Ye and G. Wang, "A comparison study on electric vehicle growth forecasting based on grey system theory and NAR neural network," 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm), 2016, pp. 711-715 is available at https://doi.org/10.1109/SmartGridComm.2016.7778845.

A Comparison Study on Electric Vehicle Growth Forecasting based on Grey System Theory and NAR Neural Network

Xian Zhang, Ka Wing Chan
Department of Electrical Engineering
Hong Kong Polytechnic University
Hong Kong
eexianzhang@outlook.com

Xuesen Yang, Yangyang Zhou, Kexin Ye, Guibin Wang
Department of Mechatronics and Control Engineering
Shenzhen University
guibinwang@outlook.com

***Abstract*— Grey system forecasting theory model and nonlinear autoregressive (NAR) neural network model for forecasting the number of electric vehicles (EVs) in the city of Shenzhen are established in this paper separately. The number of EVs from 2006 to 2015 was used as the raw data in two models. The effectiveness of the two models are evaluated by various criteria. Afterward, the rationality, precision and the adaptability of the two models are compared. At last, the better model was used to forecasting the number of EVs in Shenzhen from 2016 to 2020.**

***Index Terms*—grey system-forecasting theory; NAR neural network; EV charging demand forecasting**

I. INTRODUCTION,

THE charging demand prediction of EVs explores the EV demanding characters and the short-term forecasting methods. Firstly, accurate EV demand forecasting is important for the unit commitment (UC), economic dispatch, optimal power flow(OPF) and electric power market transaction in the power system. On the other side, EV demand forecasting is the precondition for the planning, energy management and economy operation of charging stations. Thus it is quite necessary to explore the efficient method for the EV growth forecasting in the future by establishing reasonable models considering EV charging behaviors as well as charging methods basing on actual data.

Parametric technologies are widely used in the forecast of the number of EVs, such as autoregressive integrated moving average (ARIMA). Nonparametric methods are also broadly used such as artificial neural networks (ANN) and support vector regression (SVR). Recently, grey system forecasting theory was applied to the population forecasting [1], and power system load forecasting [2]. NAR neutral network is used in the establishment of aquaculture water nitrite prediction model, or the classification and prediction of audience rating [3-5].

In this paper, the two theories, grey system forecasting and NAR neutral network, are used separately to establish the model for forecasting the EV number growth in the future basing on the same history data. After the comparison of the accuracy, the rationality, and the adaptability of two models, the better one are chosen to forecast EVs in Shenzhen from 2016-2020.

This paper is organized as follows: The theory of the proposed two methods is stated in section II. A case study utilizing the real data in the city of Shenzhen is carried out in section III. Finally the conclusion is given in section IV.

II. ASSUMPTIONS AND MODELING

A. Model establishing basing on grey system forecasting theory

EV number growth data are complicated, orderly and have overall functionality. Grey system forecasting theory tries to find inherent laws in the seemingly disorganized data. Grey forecasting theory firstly distinguishes differences in trends of factors, and then finds hidden laws in history data in certain time horizon after processing the data. After that, grey system theory generates data sequence with regularity and forecasts the trend of data by setting up several differential equations.

GM (1, 1) is the most widely used grey model, which is a first-order differential model to forecast one variable. The mathematical model of GM (1, 1) is as follows [6]:

1) The general form of GM (1, 1)

We need to use a time series reflecting the character of the prediction object to structure the GM (1, 1) model, which forecast the features at a specific time in the future or the time when a feature increases to a certain value. In general, the raw time series $X_i^{(0)}$ can be written as follows:

$$X_i^{(0)}=\{x_1^{(0)}\ \ x_2^{(0)}\ \ \dots\ \ x_n^{(0)}\}\quad i\text{=1,2,3, …,n} \qquad (1)$$

At first, $X_i^{(1)}$ is generated by first-order accumulation and eliminating the randomness and volatility of the data:

$$X_i^{(1)}=\{x_1^{(1)}\ \ x_2^{(1)}\ \ \dots\ \ x_n^{(1)}\}\qquad i\text{=1,2,3, …,n} \qquad (2)$$

Among them, $x_s^{(1)} = \sum_{s=1}^{i} x_s^{(0)}$ (3)

The first-order differential is

$$\frac{dX^{(1)}}{dt} + \alpha X^{(1)} = \mu \quad (4)$$

2) Identification algorithm

Let grey number parameter series be $\hat{a}, \hat{a} = [a, \mu]^T$, $\hat{a}$ can be solved by the least square method:

$$\hat{a} = (\mathrm{B}^T\mathrm{B})^{-1}\mathrm{B}^T\mathrm{Y}_n \quad (5)$$

In the equation, B is the matrix after processing; Y_n is a data column.

$$\mathrm{B}=\begin{vmatrix} -\frac{1}{2} & (x_1^{(1)} + x_2^{(1)}) & 1 \\ & \cdots & \cdots \\ -\frac{1}{2} & (x_{n-1}^{(1)} + x_n^{(1)}) & 1 \end{vmatrix} \quad (6)$$

$$\mathrm{Y}_n = (x_2^{(0)} \ x_3^{(0)} \ \ldots \ x_n^{(0)})^T \quad (7)$$

The GM (1,1) model are as above.

The results of the GM (1, 1) model, the predicted value are:

$$\hat{X}_{\mathrm{k+1}}^{(1)} = \left(X_1^{(0)} - \frac{u}{a}\right) e^{-ak} + \frac{u}{a} \quad (8)$$

or

$$\hat{X}_{\mathrm{k}}^{(1)} = \left(X_1^{(0)} - \frac{u}{a}\right) e^{-a\,(\mathrm{k}-1)} + \frac{u}{a} \quad (9)$$

3) The transformation of the predicted value

The result of The GM (1, 1) model is the first-order cumulative result. To getting the result on the time $k \in \{n+1, n+2, \ldots\}$, $\hat{X}_{\mathrm{k+1}}^{(1)}$ getting from the GM (1,1) model should be transformed to $\hat{X}_{\mathrm{k+1}}^{(0)}$ as follows:

$$\hat{X}_{\mathrm{k+1}}^{(0)} = \left(X_{k+1}^{(1)} - X_k^{(1)}\right) \quad (10)$$

B. Model establishing basing on Nonlinear Autoregressive Neural Network

NAR neural network theory model could be defined as:

$$y_n = \mathrm{f}\left(y_{n-1}, \ \ldots, \ y_{n-k}, \ x_n, \ \ldots, \ x_{n-1}\right) \quad (11)$$

In the formula, x is the input data; y is the output data; n is a time series; f is a nonlinear function.

In the model, although these data are the factors influencing the predictable number of electric vehicles, there are default input data, such as socio-economic level, charging infrastructure, policies and regulations, which cannot be quantified. So we take the history output value as the input data. The model is constructed as follows:

$$y_n = f\left(y_{n-1}, \ \ldots, \ y_{n-k}\right) + k\varepsilon_n \quad (12)$$

In this formula, k is a constant; ε_n is a random variable obeying the Gauss distribution.

The output of each y in the NAR model will be the input data in the next calculation as the adjustment parameters for the next output, completing the adjustment of the neural network [3].

In the neural network training for the model, a neural network is created firstly, then the autoregressive order k should be set. Input sequence is $\{\mathrm{y_i}, \mathrm{y_{i+1}}, \ \ldots, \mathrm{y_{i+k-1}}\}$ i=1,2…n. And the target output is $\{\mathrm{y_{i+k}}\}$ i=1,2…n. For each input sample, the network output and the target output comparison algorithm will automatically adjust the network parameters to minimize mean square error. Actually in the MATLAB environment, fitness function is used to construct the network.

Next, network parameters of NAR neural network should be determined. NAR neural network is mainly composed of input layer, output layer and hidden layer. Because the model sample of our prediction model only has only one output variable, the number of neurons in the output layer is set to be 1. The order of auto-regression and the number of the hidden layer neurons are determined by a variety of factors. At present, there is no mature theoretical basis. The most reasonable approach is selecting a reasonable set of parameters to test in a number of comparative ways, and deciding the final parameters according to the test results.

C. Performance Criterion

The performance of the proposed approach could be comprehensively assessed in this paper via different indices.

1) Residual error: This index is the difference value between the actual value and the forecasted value

$$e_k = X_k - \hat{X}_{\mathrm{k}} \quad (13)$$

where the e_k is the residual error, while X_k and $\hat{X}_{\mathrm{k}}$ is the actual value and the forecasted value, respectively.

2) Relative error: The relative error could be stated as

$$e_k = (X_k - \hat{X}_{\mathrm{k}})/X_k * 100\% \quad (14)$$

3) The mean value of the residual error:

$$\bar{e} = \frac{1}{n}\sum_{i=1}^{n} e(i) \quad (15)$$

4) The mean value of the original data:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i^{(0)} \quad (16)$$

5) The standard deviation of the original data:

$$S_1 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i^{(0)} - \bar{x}\right)^2} \quad (17)$$

6) The standard deviation of the residual error

$$S_2 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e(i) - \bar{e})^2} \quad (18)$$

7) Rariance ratio

$$\mathrm{C} = \frac{S_2}{S_1} \quad (19)$$

III. CASE STUDIES

A. Test Case Settings and Methodologies

The total EV number in Shenzhen, Guangdong province, China from 2006 to 2015 is utilized as the test database in our study, which is demonstrated in Table I. The EVs are classified as electric buses (EBs) and non-EBs (including private electric cars, electric taxing, et al.) because they have obvious different application purposes and developing characteristics. From the table I we can also notice the fast growth rate of EVs in recent years.

In our forecasting model, every three years data of EV number is used to forecast the future EV number in the next year, and the forecasted number is compared by the actual EV number to testify the effectiveness of our proposed models.

Tab. I The EB and Non-EB number in Shenzhen from 2006-2015

Year	EB number	Non-EB number
2006	500	1000
2007	900	1250
2008	1200	3500
2009	1700	6200
2010	2100	7320
2011	2500	9300
2012	3100	10800
2013	4200	13500
2014	5400	19000
2015	7000	27500

B. The Forecasting Results of Grey System Theory

Firstly, the grey system theory is utilized to forecast the EV numbers. The forecasted results of EBs and non-EBs are demonstrated in Fig. 1 and Fig. 2, respectively.

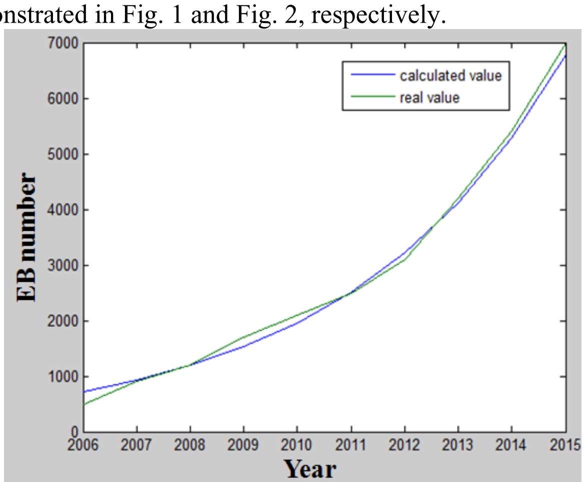

Fig 1. The forecasted results of EBs by the grey system theory.

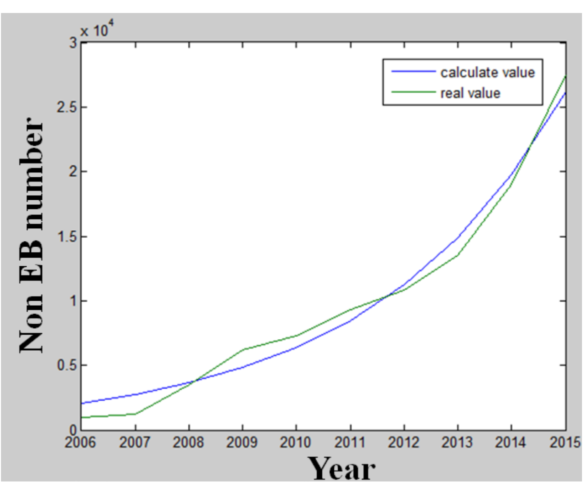

Fig 2. The forecasted results of non EBs by the grey system theory.

It could be roughly observed in Fig. 1 and Fig. 2 that the forecasting result is quite acceptable. To ensure that the established model has enough prediction accuracy in the practical application, three index, residual error, degree of association and variance ratio is calculated to check out the effectiveness of the proposed method. The residual error of EBs and non EBs is shown in Table II and Table III, respectively. And the association degree and variance ratio of EB and non EB forecasting is demonstrated in Table IV. According to the accuracy inspection level, which is shown in Table V, the forecasting relative error belongs to level 4, and the association rate belongs to level 1, while the variance ratio belongs to level 1, which indicates that the grey system method is suitable to forecast the EV numbers in Shenzhen, but the accuracy is not good enough for some situations.

Tab. II The EB forecasting result residual error

Year	Actual EB number	Forecasted EB number	residual error e	relative error %
2006	500	722	-222	44.4
2007	900	926	-26	2.9
2008	1200	1187	13	1.1
2009	1700	1523	177	7
2010	2100	1953	147	7
2011	2500	2505	-5	0.2
2012	3100	3202	-112	3.6
2013	4200	4119	81	1.9
2014	5400	5283	117	2.2
2015	7000	6775	225	3.2

Tab. III The non EB forecasting result residual error

Year	Actual EB number	Forecasted EB number	residual error e	relative error %
2006	1000	2068	-1068	106.8
2007	1250	2742	-1492	119.4
2008	3500	3635	-135	3.9
2009	6200	4820	1380	22.3
2010	7320	6390	930	12.7
2011	9300	8472	828	8.9
2012	10800	11232	-432	4.0
2013	13500	14891	-1391	10.3
2014	19000	19741	-741	3.9
2015	27500	26172	1328	4.8

Tab. IV The association degree and variance ratio of grey system method

	EB forecasting	Non EB forecasting
association degree	0.599	0.556
variance ratio C	0.066	0.136

Tab. V Accuracy level reference table

Accuracy level	relative error %	association degree	variance ratio C
Level 1	0.01	0.90	0.35
Level 2	0.05	0.80	0.50
Level 3	0.10	0.70	0.65
Level 4	0.20	0.60	0.80

C. The Forecasting Results of Nonlinear Autoregressive Neural Networks

Then the nonlinear autoregressive neural network is utilized in the EV number growth forecasting. The autoregressive process order is set to 3 and the hidden neuron number is set to 10 in this case. The forecasting setting interface of EBs and non-EBs are demonstrated in Fig. 3. The forecasted results of EBs and non-EBs are demonstrated in Fig. 4 and Fig. 5, respectively. The residual error of EBs and non EBs is shown in Table VI and Table VII.

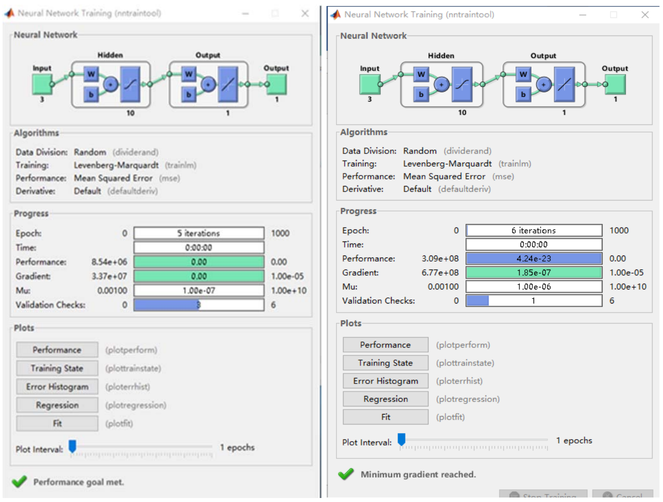

Fig 3. The forecasting setting interface of EBs and non-EBs.

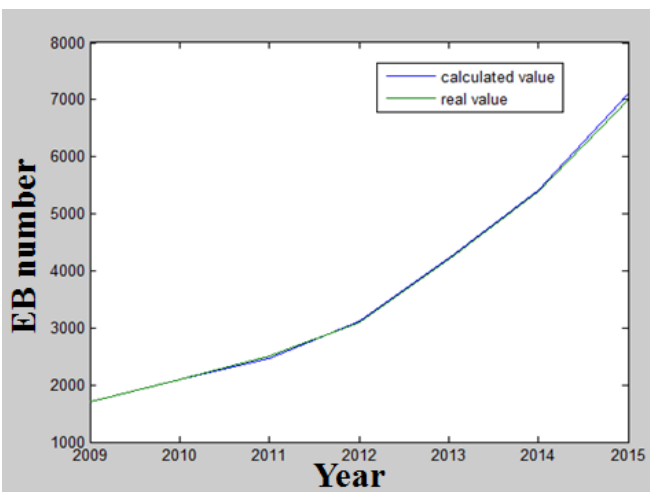

Fig 4. The forecasted results of EBs by NAR

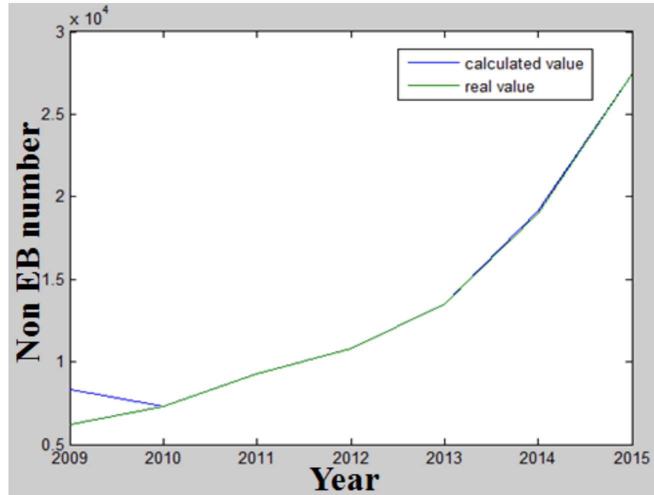

Fig 5. The forecasted results of non EBs by NAR

Tab. VI The EB forecasting result residual error of NAR

Year	Actual EB number	Forecasted EB number	residual error *e*	relative error %
2009	1700	1697	3	0.2
2010	2100	2098	2	0.1
2011	2500	2470	30	1.2
2012	3100	3108	-8	0.3
2013	4200	4221	-21	0.5
2014	5400	5407	-7	0.1
2015	7000	7111	-111	1.6

Tab. VII The non EB forecasting result residual error of NAR

Year	Actual EB number	Forecasted EB number	residual error *e*	relative error %
2009	6200	8323	-2123	34
2010	7320	7320	0	0
2011	9300	9300	0	0
2012	10800	10800	0	0
2013	13500	13500	0	0
2014	19000	19118	-118	1
2015	27500	27500	0	0

D. The Forecasting Results of Future 5 years by NAR

It could be concluded that the NAR method has a better performance in the forecasting of EV numbers. In this section the EV number in Shenzhen from 2016-2020 is forecasted according to historic data by the proposed NAR method. The forecasted results of EBs and non-EBs are demonstrated in Fig. 6, Fig. 7 and Table VIII.

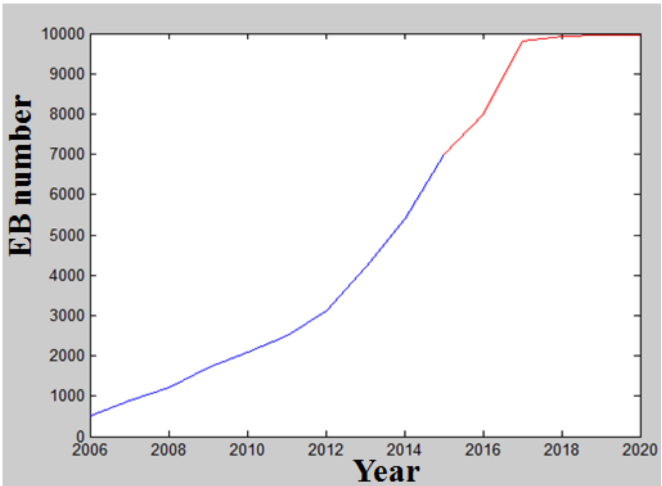

Fig 6. The forecasted results of EBs by NAR from 2016-2020

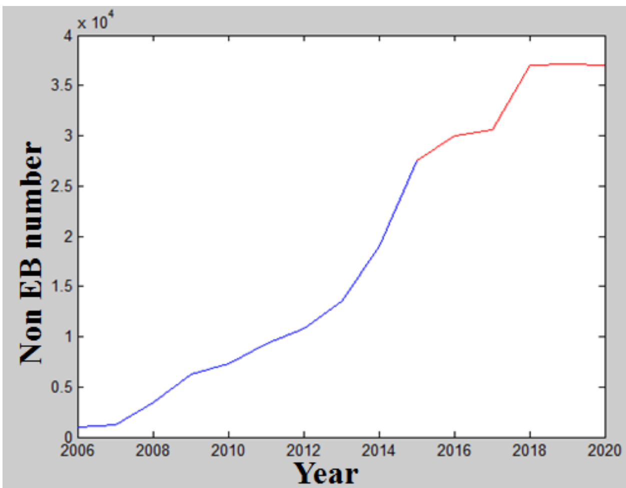

Fig 7. The forecasted results of non EBs by NAR from 2016-2020

Tab. VIII The EB and non EB forecasting results from 2016-2020

Year	EB forecasting	Non EB forecasting
2016	7911	30038
2017	9804	30576
2018	9922	37055
2019	9956	37091
2020	9996	37034

IV. CONCLUSION

The rapid social, environmental, economic and transportation development promotes the wide use of EVs in Shenzhen. Our study shows that the prediction accuracy of the grey system forecasting model is high only when the original EV demand data increases exponentially. Otherwise, the prediction appears a larger deviation. In comparison, it shown in this work that the NAR neural network model for EV charging demand forecasting has a good performance in future practical application. It should also be noticed that the disadvantage of the NAR neural network could be affected by the social uncontrollable factors, and the forecasting accuracy is not high when the original data fluctuation is obvious. How to overcome the mentioned disadvantages is our future work.

V. ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the Hong Kong Polytechnic University and Ms. Xian Zhang would also like to acknowledge the Hong Kong Polytechnic University Research Studentship for her PhD study. The authors gratefully acknowledge the support in part by the Shenzhen University Research and Development Startup Fund (No. 201530), National Basic Research Program (973 Program) (No. 2013CB228202), National Natural Science Foundation of China (No. 51507103, No. 51107114, No. 51177145), and Natural Science Foundation of Guangdong Province (No. 2015A030310316).

REFERENCES

[1] LAI Hongsong, ZHU Guorui, et al, “Population forecast based on combination of grey forecast and artificial neural networks,” *Economic Geography*, vol.24, no.2, pp. 197-201, Mar. 2004. (in Chinese).

[2] SHI Deming, LI Linchuan, et al, “Power system load forecasting based upon combination of grey forecast and artificial neural network,” *Power System Technology*, vol.25, no.12, pp. 14-17, Dec. 2001. (in Chinese).

[3] WANG Xiang, HE Jixiang, et al, “Establishment of aquaculture water nitrite prediction model based on NAR neural network,” *Fishery Modernization,* vol.42, no.4, pp. 30-34, Aug. 2015. (in Chinese).

[4] WEI Duan，LU Xiangyang, “An improved BP neural networks applied to classification,” *Energy Procedia*, vol.13, no.4, pp. 7065-7069, Sep.2012.

[5] Xiao Qiwei,Yang Xiuzhi, “Prediction of Audience Rating Based on Nolinear Auto-regressive(NAR) Model,” *Video Engineering*, vol.39, no.4, pp. 79-81, Sep.2015.

[6] ZHU Dengyuan, CHANG Xiaofeng, “Matlab realization of grey prediction GM (1,1) model,” *Journal of Henan University of Urban Construction*, vol.22, no.3, pp. 40-46, May. 2013. (in Chinese).

[7] Wei Yongqi, “The method of inspecting the forecasting precision of grey system with the correlation degree,” *Jiangxi Hydraulic Science & Technology*, vol.18, no.4, pp. 322-328, Dec. 1992. (in Chinese).