Analysis of Product sales Forecast based on time Series Model

Xiaoduo Zhang, Jinxuan Yang, Rui Zhai, Wanmei Cui

Yunnan University of Finance and Economics, Kunming, Yunnan 650221

Keywords: forecast demand; correlation analysis; moving average forecast; grey prediction

Abstract: This paper uses the neural network multi-layer perceptron, grey prediction method, moving average method, and other methods to construct grey prediction and time series models and comprehensively uses MATLAB, C language, to aim at forecasting demand SPSS, and Excel software to program and solve. First, predict the monthly sales of a target subclass in a specified month and calculate the MSRE. The data are preprocessed with Excel and C language, then the time series model is used, the moving average method is used to predict and analyze the data, and MATLAB is used to solve the problem and get the prediction results. Then forecast the weekly sales of the specified week in the target subcategory and calculate all MSRE. First of all, the data is preprocessed with Excel and C language. Then the grey prediction model is established and solved by MATLAB, and the prediction results are obtained.

1. Introduction.

With China's consumer market development, the consumption pattern in the market has gradually changed to "customer-oriented." With the continuous improvement of social productivity and customer income, their demand level has been further sublimated. To meet the needs of customers, the products produced by the enterprise are gradually moving forward to multi-variety and small batch. At the same time, it is more difficult for enterprises to manage inventory.

Therefore, at present, most retail enterprises need to focus on how to give accurate customer demand forecast at the regional level, small category level, and even store skc level according to complex and various sales data. This paper will solve the problem of "accurate demand forecasting" of new retail enterprises from three directions. First of all, considering the product sales characteristics, inventory information, holiday discounts, and other factors, analyze the impact of various factors during the 2018 National Day, double Day, double Twelve and New Year's Day period on the sales volume of the top 50 skc with cumulative sales from July 1, 2018, to October 1, 2018. Then, combined with the above analysis results, we forecast the monthly sales of the target subcategory in a given region in the three months after October 1, 2019, and give the monthly forecast MAPE. The target subcategory is the top 10 categories of cumulative sales from June 1, 2019, to October 1, 2019.

2. The establishment of target subclass prediction model

In this paper, the moving average method is selected to predict the monthly sales of the target subcategory in a given region in each of the three months after October 1, 2019.

Set up the view sequencing column and take the number of items on the moving average. The formula for calculating the one-time moving average is as follows:

$$\begin{aligned} M_t^{(1)} &= \frac{1}{N}(y_t + y_{t-1} + \ldots + y_{t-N+1}) \\ &= \frac{1}{N}(y_{t-1} + \ldots + y_{t-N}) + \frac{1}{N}(y_t - y_{t-N}) \\ &= M_{t-1}^{(1)} + \frac{1}{N}(y_t - y_{t-N}) \end{aligned} \tag{1}$$

When the primary trend of the predicted target is to fluctuate up and down at a certain level, the prediction model can be established by using the one-time moving average method:

$$y_{t_0+1} = M_t^{(1)} = \frac{1}{N}(y_{t_0} + ... + y_{t_0-N+1}), \quad t = N, \quad N+1,... \tag{2}$$

The standard deviation of prediction is:

$$S = \sqrt{\frac{\sum_{t=N+1}(y_{t_0} - y_t)^2}{T-N}} \tag{3}$$

The average value of the most recent series is used as the forecast result for future periods.

When the historical series's basic trend does not change much and more random components in the series, the value should be more considerable. Otherwise, the value should be smaller. In the data with a definite period change, the number of items on the moving average should take the length of the period. An effective way to select the best value is to compare the prediction errors of several models, and the one with the smallest standard error is better.

3. Application of target subclass prediction model

3.1 Forecast of the top three categories of sales

Combined with the results of the analysis of the impact of various related factors on the sales of the target skc during the four holidays of 2018 National Day, Singles Day, double Twelve, and New Year's Day, this paper forecasts the sales volume of the target subcategory in each of the three months after October 1, 2019.

(1) Historical sales analysis and future sales of the target subcategory (No. 27050401) rank first in sales.

I Analysis of historical sales volume

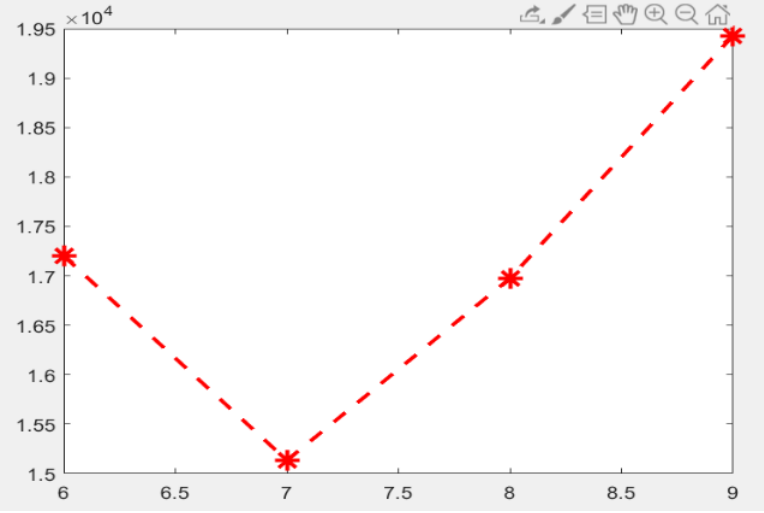

Fig 1. Historical sales chart

Table 1. Historical sales scale

Series	Month	Sales volume
1	6	17201
2	7	15136
3	8	16976
4	9	19428

II Forecast of future sales.

A sales forecast for October 2019
Respectively, the prediction formula:

$$y_{t_0+1}^{(1)} = \frac{y_t + y_{t-1}}{2}, \quad t = 2,3,4$$
$$y_{t_0+1}^{(2)} = \frac{y_t + y_{t-1} + y_{t-2}}{3}, \quad t = 3,4 \tag{4}$$

At $N = 2$ the time, the predicted value $y_{t_0=5}^{(1)} = 18664$.

At $N = 3$ the time, the predicted value $y_{t_0=5}^{(1)} = 18101$.

B sales forecast for November 2019
Respectively, the prediction formula $N = 2$ $N = 3$, :

$$y_{t_0+1}^{(1)} = \frac{y_t + y_{t-1}}{2}, \quad t = 2,3,4,5$$
$$y_{t_0+1}^{(2)} = \frac{y_t + y_{t-1} + y_{t-2}}{3}, \quad t = 3,4,5 \tag{5}$$

At $N = 2$ the time, the predicted value $y_{t_0=6}^{(1)} = 18282$.

At $N = 3$ the time, the predicted value $y_{t_0=6}^{(1)} = 17846$.

C sales forecast for December 2019.
Respectively, the prediction formula $N = 2$ $N = 3$:

$$y_{t_0+1}^{(1)} = \frac{y_t + y_{t-1}}{2}, \quad t = 2,3,4,5,6$$
$$y_{t_0+1}^{(2)} = \frac{y_t + y_{t-1} + y_{t-2}}{3}, \quad t = 3,4,5,6 \tag{6}$$

Based on the above results, the future forecast sales of the target subcategory (serial number: 27050401) ranked first in sales can be obtained:

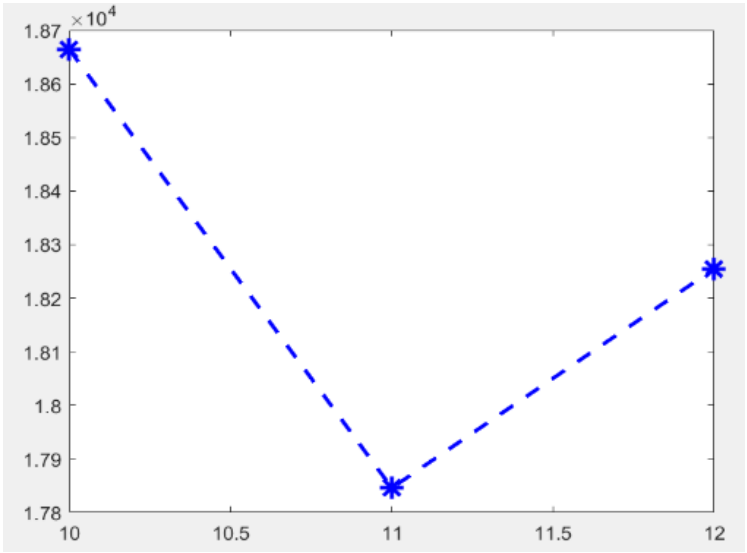

Fig. 2 Historical sales chart

Table 2 Historical sales chart

Month	Sales volume
10	18664
11	17846
12	18255

(2) The second target subcategory (serial number: 27060804) historical sales analysis and future sales.

I Analysis of historical sales volume

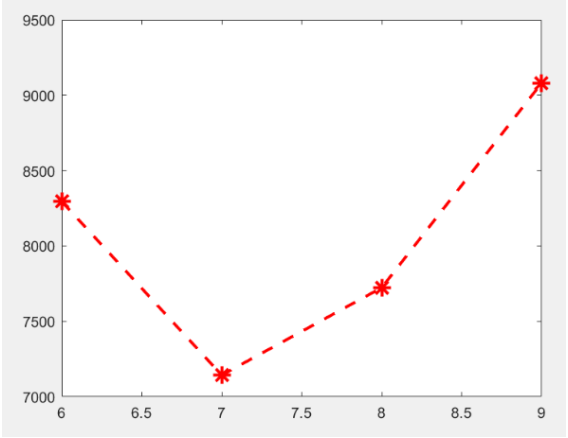

Fig. 3 Historical sales volume

Tab 3 Historical sales table

Series	Month	Sales Amount
1	6	8295
2	7	7143
3	8	7723
4	9	9081

II Forecast of future sales.

It can get the future forecast sales of the target subcategory (serial number: 27060804), which ranks second in terms of sales volume:

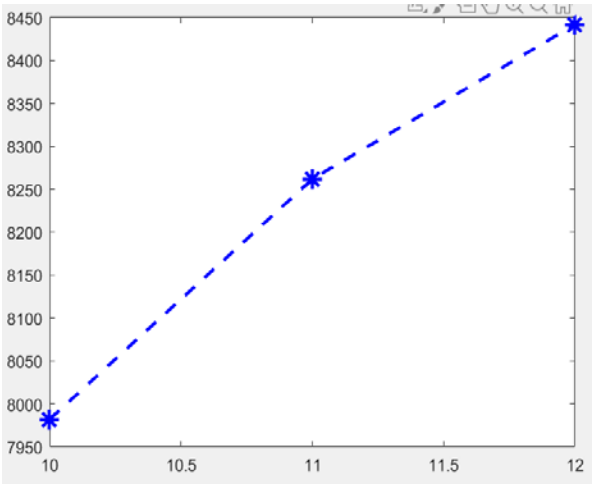

Fig. 4 Historical sales chart

Table 4 Historical sales chart

Month	Sales Amount
10	7982
11	8262
12	8442

3.2 Calculation of MAPE of the monthly forecast

According to the formula $MAPE = \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{n * y_i} = \sum_{i=1}^{n} \frac{1}{n} * \frac{|y_i - \hat{y}_i|}{y_i}$, the paper obtains the following data from excel:

Table 5 MAPE values

MAPE		
Oct	Nov	Dem
34.6%	78.5%	65.6%

4. Time series prediction model

4.1 Level ratio test

First of all, this paper establishes the weekly (1-18 weeks) time series of weekly sales of all skc in the target subcategory from June 1, 2019, to October 1, 2019:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), ..., x^{(0)}(18)) \tag{7}$$

According to the established skc sales time series, according to the grade ratio $\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}$,Get $\lambda = (\lambda(2), \lambda(3), \lambda(4), ..., \lambda(18))$, so $x(0)$ can be used for good modeling.

The original data is accumulated once, and the data matrix B and data vector Y is constructed.

$$B = \begin{bmatrix} \frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ \frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ \frac{1}{2}(x^{(1)}(17) + x^{(1)}(18)) & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(18) \end{bmatrix} \tag{8}$$

Calculate $\hat{u} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y$.

The value of a and b can be obtained. So we can get the model $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$: the solution is $x^{(1)}(k+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a}$:

4.2 Model solving

In this paper, the target subcategory is divided into 18 weeks from June 1 to October 1, 2019. Based on the historical data of these 18 weeks, the sales volume of the 12 weeks after October 1, 2019, is predicted.

Here is an example of historical sales analysis and the future sales of the target subcategory (No. 27050401), which ranks first in sales volume.

Table 6 comparison of historical sales and forecast

Historical sales		Forecast sales	
Series	Sales account	Series	Sales account
1	3681	19(The first week after October 1st)	5047
2	3175	20(The second week after October 1st)	5184
3	4596	21(The third week after October 1st)	5325
4	4185	22(The fourth week after October 1st)	5470
5	3534	23(The fifth week after October 1st)	5619
6	3685	24(The sixth week after October 1st)	5772
7	3344	25(The seventh week after October 1st)	5928
8	3375	26(The eighth week after October 1st)	6090
9	3561	27(The ninth week after October 1st)	6255
10	3948	28(The tenth week after October 1st)	6425
11	3353	29(The eleventh week after October 1st)	6600
12	4057	30(The twelfth week after October 1st)	6780
13	3754		
14	3014		
15	3908		
16	4114		
17	5446		
18	7111		

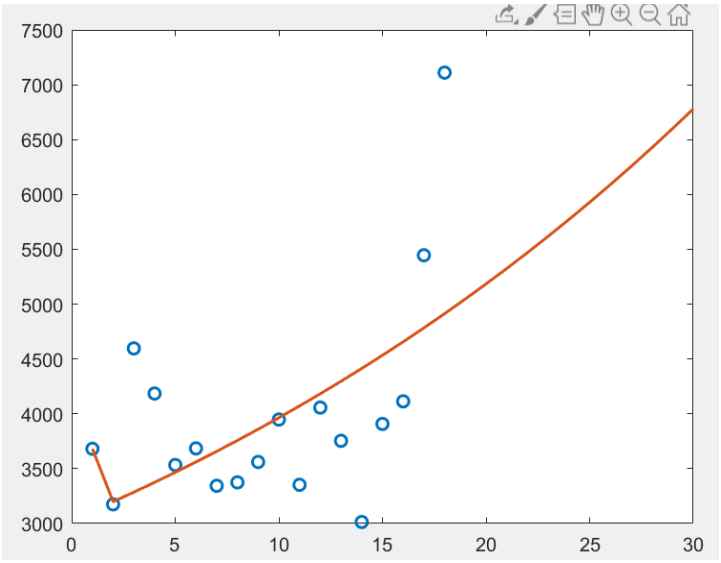

Fig. 5 comparison figure

5. Conclusion

In this paper, we choose the model according to the actual problem, forecast the sales volume in three months, the data requirement is less, the forecast period is short, we choose the time series theory and use the moving average to forecast. Then forecast the weekly sales volume of 12 weeks, the forecast period is extended, using grey system theory, using grey prediction. Use moving average prediction, grey prediction to predict the sales volume of a given region and get the predicted value of MAPE, combined with SPSS to get the correlation analysis data of the relevant factors affecting sales volume, making the results more accurate.

Because of the initial value of the model, the smoothness and grade deviation of the original data sequence affect the accuracy of the grey prediction model. Relative error tests, posterior error tests, and correlation tests can test the constructed model to achieve better results.

Reference

[1] Xiong Gang, Chen Zhangchao, defects and improvement of grey prediction model [J]. Systems Engineering, 1992 10 (06): 32-33.

[2] Tang Xiongying, talking about holiday consumption Tang Xiongying [J]. Mass Commerce, 2010 (109): 27-28.

[3] Zeng Chuwen, talking about the economic thinking of "Singles Day" Shopping Festival [J]. E-commerce, 2017 (28): 3-4.

[4] Zhang Jun, improvement and application of grey prediction model [D].

[5] Xu Zhiming, Wang Peng, Dong Yi. Evaluation of land ecological security in Hengyang City based on PSR model and grey prediction model [J]. Journal of Xinjiang normal University (Natural Science Edition), 2020, 39 (01): 49-56.

[6] Liu Dandan, Zhu Jiaming, Huang Tingting. Short-term power load forecasting based on time series and grey model [J]. Journal of Qiqihar University (Natural Science Edition), 2017, 33 (03): 7-12. 20.

[7] Zhang Haiyang, Gao Bai, Ma Wenjie, Li Yanmei, Lin Congye, Shen Wei, Zhu Liangzhu. Analysis of adsorption characteristics and influencing factors of U (VI) removal materials based on SPSS [J]. Non-ferrous Metals Engineering, 2020jue 10 (04): 109-116.

[8] Wang Ronghua. Analysis on the influencing factors of College students' participation in online Games based on SPSS [J]. China Journal of Multimedia and Network Teaching (first ten days), 2020 (05): 24-26.

[9] Chen Jiaqi. Based on SPSS software to realize the regression analysis of the factors affecting the employment of rural accompanying parents [J]. Journal of Beijing Institute of Printing, 2020, 28 (01): 58-60.