Guizhou Cigarette Sales Prediction based on Seasonal Decomposition MLP

Jian Tan

School of Management Science and Project Management, Guizhou University of Finance and Economics, Guizhou 550025, China

tanjian123@126.com

Keywords: Cigarette Sales; Prediction; Multilayer Perceptron; Seasonal Decomposition.

Abstract. We combine with seasonal decomposition method, and employ multi-level sensor prediction (MLP, Multilayer Preceptor) method, take prediction of 2014 cigarette sales in Guizhou as example, monthly and quarterly cycle MLP prediction models were put forward, results of the prediction precision were analyzed, through the prediction accuracy of each model, the importance of the independent variables and the average relative error rate, can be drawn on the quarterly MLP model is highest accuracy, it better suited to predict cigarette sales in Guizhou.

1. Introduction

Demand forecasting is the overall analysis and judgment for the future market demand. Cigarette demand forecasting is an important part of the tobacco industry's marketing activities, and it's the cigarette market forecast commodity central goal. Based on the market, adhere to market-oriented, starting from the market reality, overcoming the non-market factors, scientific prediction is an effective means to accurately grasp the market demand, which also the basis for supply work orders.

In recent years, Cigarette market forecasts focused on traditional forecasting methods, for example Xi Hao (1996) using multiple regression model to predict the medium and long term national cigarette sales[1], Wang Weimin (2009) utilizing the gray Markov model to predict national cigarette demand[2], Zou Liang (2009) composing the trend push method, time series decomposition, multiple regression method to forecast Huaihua cigarette demand[3],Shi Yan feng (2010) composing the least square method, exponential smoothing and ARIMA to forecast annual and monthly cigarette demand of An Kang city [4], Zhong Dongting(2007) applying Levenberg Marquardt algorithm to improve BP neural network for prediction cigarette sales [5].Overall, these methods are regression model, time series model, seasonal decomposition model, neural network model, gray system model, Markov chain model, combined model and so on, as specified in the literature [6] shows.

In fact, most of the forecasting methods used in the literature based on the linear variation of cigarette sales, but most of the historical data above have shown that cigarette sales are provided with seasonal and cyclical dynamic changes, and cigarette sales with significant non-linear is affected by the regional economic level, population ,various macro and micro environmental. However, these traditional forecasting methods can not reflect the nonlinear characteristics of cigarette sales. Neural network algorithm is a nonlinear predictive ability of intelligent learning algorithm; it can accurately describe the variation of nonlinear systems. Although some scholars applying neural network algorithm to predict the cigarette market, but they still ignore the important characteristics of cigarette sales cyclical. In this paper, for the seasonal, cyclical and non-linear characteristics of cigarette sales, we combine with seasonal decomposition method, and employ multi-level sensor prediction method, to forecast cigarette sales of Guizhou Province.

2. Theory and Data

Seasonal Decomposition Model. Seasonal time series decomposition model will be divided into four categories of factors, such as long-term trend factor, seasonal variable, cycle variable and

irregular variable. Suppose T denotes the long-term trend, seasonal variation is S, C denotes cyclic variation; I is irregular variable, then optional forecasting models are plus model $Y = T + S + C + I$, multiplication model $Y = T \cdot S \cdot C \cdot I$, and hybrid model $Y = T \cdot S \cdot C + I$.

MLP Model. Multilayer Perceptron is a feed forward supervised structure, a plurality of input data sets can be mapped to a single output data set.

Data. According to Guizhou Province 2006.1-2013.12 cigarette sales data of each month, we get Figures 1 and 2. As can be seen from Figure 1, the trend of cigarette sales in Guizhou Province is overall upward, but sales in January were significantly higher than other sales month of the year. As can be seen from Figure 2, for every year, the first quarter sales is largest, and also the trend is upward. As can be seen from the figures, whether it is the month or quarter sales, they are showing the seasonal and cyclical law, which provide a theoretical basis for the following sales forecasts.

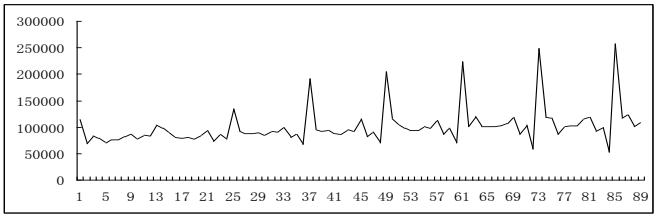

Fig. 2 2006 - 2013 monthly actual sales graph

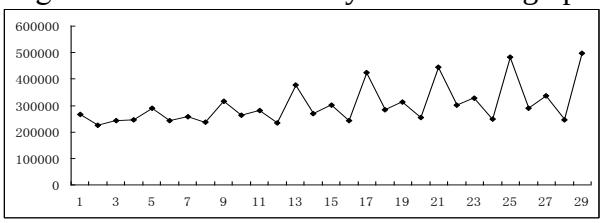

Fig. 3 2006 -2013 quartly actual sales graph

3. Model Predicting

Seasonal Decomposition Model Prediction .Applying seasonal factors decomposition, respectively, for the month and quarter cycle, choose a multiplicative model, factor out the variables, measure the change laws, then comprehensively reflect variation factors affecting changes in the time series.

4. Monthly Seasonal Decomposition

Long-term Trend Analysis. Long-term trend factor reflect a phenomenon development direction in a long period, it can exhibit an approximate continuous upward, downward trend or balance. Time t is the independent variable ($t = 1$ means January 2006), sales y is the dependent variable, we employ the linear regression model $\hat{y} = b_0 + b_1 t$ to fit the long-term trend. Using the least square method, the results obtained are shown in Table 1. According to Table 1, the non-standardized coefficients Sig t test was less than 0.05, the test results significantly, the long-term trend model regression equation is $\hat{y} = 79621.969 + 466.756t$.

Table 1 Long-term results of linear regression trend

Model		Non-standardized Coefficients		t	Sig.
		B	Standard Error		
1	(Constant)	79621.969	6963.108	11.435	.000
	Sales	466.756	134.379	3.473	.001

Seasonal Factors. Seasonal factors are phenomenon that the length and magnitude of changes is fixed cycle formed by seasonal fluctuations. Cigarette sales of the time series show the month or quarter cyclical changes. Based on historical data, the seasonal factors are shown in Table 2.

Cyclical Changes. Known as loop variables, due to some physical effects or economic reasons, it shows fixed cycle changes. According to the residual method, using the moving average method

$\hat{y}_n = \sum_{i=n-11}^{n} y_i / 12$ to exclude long-term trend and cycle change, we obtain the sequence $T \cdot C$, $T \cdot C$ divided by the long-term trend T, we get cyclical variable C. Taking the average value for each of the cyclical movements in the same month as the factor S, we get concrete results as shown in Table 2.

Table 2 Seasonal cycle in the month decomposition predictions(Unit: Box)

t	T	C	S（%）	Predictions	t	T	C	S（%）	Predictions
201401	124897.3	0.957	186.9	223449.5	201407	127697.8	0.972	93.3	115747.8
201402	125364.1	0.969	103.9	126277.8	201408	128164.6	0.971	95.6	118949.6
201403	125830.8	0.971	100.7	123048.5	201409	128631.3	0.971	104.8	130864.9
201404	126297.6	0.971	93	114071.7	201410	129098.1	0.971	81.4	102037.7
201405	126764.3	0.969	90.3	110957.7	201411	129564.9	0.969	91.6	114948
201406	127231.1	0.972	90.3	111711.9	201412	130031.6	0.966	68.2	85687.5
					Total sales				1477752

Because of the randomness cannot be directly predicted, the predicted relationship is $Y = T \cdot S \cdot C$, the results shown in Table 2.

Quarterly Seasonal Decomposition. The same as monthly seasonal decomposition prediction methods, by calculation, the quarterly seasonal decomposition prediction are shown in Table 3.

Table 3 Quarterly seasonal cycle decomposition predictions

Time	T	C	S (%)	Predictions
20141	383303.8	0.86939	129.3	430879.92
20142	387845.4	1.001815	91	353579.99
20143	392387.1	0.974912	99.5	380630.07
20144	396928.8	1.020805	80.2	324959.94
Total sales				1490050

5. MLP model predictions

Monthly Seasonal Decomposition MLP. Applying Multilayer Perceptron model, taking the total sales of each month as the dependent variable, year and month as factors, using batch conjugate gradient algorithm training and adjustment, and calculate output importance of each factor, the results are shown in Table 4. It shows that the relative error rate of total sales for the month is 32.5%, the relative error of the dependent variable in test set is 17.3%, indicating that prediction accuracy rate of monthly MLP model is only 67.5%. 2014 prediction results are shown in Table 5.

Table 4 Monthly Seasonal Decomposition MLP Summary

Training	Squared error	12.019	Test	Squared error	2.180
	relative error	.325		relative error	.173
	training time	00:00:00.016			

Table 5 Monthly decompose MLP seasonal predictions for2014

Month	forecasting	Month	forecasting	Month	forecasting	Month	forecasting
1	259608	4	130652	7	133347	10	141262
2	156411	5	137620	8	148828	11	135230
3	182510	6	154923	9	162391	12	84696
				Total sales		1827478	

Quarterly Seasonal Decomposition MLP .Using Multilayer Perceptron model, taking the total sales of each quarter as the dependent variable, year and quarter as a factor, using batch conjugate gradient algorithm to train and adjustment, and calculate output importance of each factor. The results are shown in Table 6.Table 6 shows that the relative error rate of total sales in the quarter is 0.7%, due to variable relative error in the test set is 14.5%, indicating that prediction accuracy of quarterly

seasonal decomposition MLP model is up to 97.3%.However, because of the test relative error is significantly higher than the training set, indicating the current sample for establishing such a complex model has slightly difficult, there may be a slight model overfitting problems. According to the season decomposition MLP model, we get the 2014 cigarette sales, as shown in Table 7.

Table 6 Quarterly Seasonal Decomposition MLP Summary

Training	Squared error	.074	Test	Squared error	.629
	relative error	.007		relative error	.145
	training time	00:00:01.120			

Table 7 Quarterly Decompose MLP Seasonal Predictions for 2014

Season	First Quarter	Second Quarter	Third Quarter	Fourth quarter	Total sales
Sale	497593.4	323707.3	373642.4	256153.1	1451096

6. Outcome Discussions

Due to the seasonal decomposition model used exponential smoothing, and therefore it reduced the predictive value of some initial data. In order to analyze the accuracy of each model, both predicted and actual values start from January 2007.

From the MLP importance of the independent variables point of view, as shown in Table 8, for the monthly MLP model, the role of the month is far greater important than the role of the quarter in the forecast; for the quarterly MLP model, the role of the year slightly more important than the role of the quarter. By Figures 2 and 3 can be seen, from the annual point of view, it shown an overall upward trend in cigarette sales, but in each year, there will appear regular cyclical fluctuations by month and quarter, therefore, the role of year should be greater important than the role of quarter and month . Therefore, from the perspective of the importance of the independent variables, quarterly MLP model is more realistic.

Furthermore, the predicted results by calculating the mean relative error rate (Table 8), the average relative error rate of the quarterly MLP model is the minimum, which is only 2.01%.

Through the prediction error figures, prediction accuracy of each model, the importance of the independent variables and the average relative error rate, can be drawn on the quarterly MLP model is highest accuracy, it better suited to predict cigarette sales in Guizhou.

Table 8 The importance of the independent variables of MLP models

	indicators	importance	standardized importance
Monthly MLP	YEAR, not periodic	0.2	25.00%
	MONTH, period 12	0.8	100.00%
Quarterly MLP	YEAR, not periodic	0.539	100.00%
	QUARTER, period 4	0.461	85.50%

Table 9 The average relative error rate of the model prediction

	season decomposition method	MLP method
Quarter	7.85%	2.01%
Month	12.46%	6.14%

7. Summary

In this paper, we take Guizhou 2014 cigarette sales forecasting as an example, applying monthly, quarterly cycle seasonal decomposition forecasting models and MLP forecasting models to forecast cigarette sales, comparative analysis the accuracy of the prediction results. The results have showed that the quarterly MLP model is more suitable Guizhou cigarette market.

Acknowledgement

This research is sponsored by the science and technology foundation of Guizhou province (Guizhou J word [2013]2085). 2013 annual social science planning project of Guizhou province (13GZYB29).

References

[1]Xi Hao. Nationwide cigarette sales related factors and long-term prediction [J]. Chinese Journal of Tobacco, 1996, 02:67 -70.

[2]Wang Weimin, Wang Yun, Zhang Guoan. Research on national cigarette demand forecast based on gray Markov model [J]. Chinese Journal of Tobacco, 2009, 06:66-69.

[3]Zou Liang. Analysis on Huaihua cigarette demand predictive based on combination forecasting model [J]. Chinese Journal of Tobacco, 2009, 02:80-85.

[4]Shi Yanfeng. Prediction thinking and application for combination of two cigarettes predict models of Ankang City [J]. Economist, 2010, 09:231-233.

[5]Zhong Dongting, Zhang Yue. BP neural network prediction method to improve the sales of tobacco research [J]. Industrial Technology & Economy, 2007 (9):115-118.

[6]Cai Pingping, Wu Zheng, Wang Guangwu. Cigarette festival demand forecasting model research [J]. Industrial Control Computer,2012,09:112-113.