Advance Journal of Food Science and Technology 5(7): 855-858, 2013
ISSN: 2042-4868; e-ISSN: 2042-4876

Submitted: February 23, 2013 Accepted: April 04, 2013 Published: July 05, 2013

Forecasting of Fresh Agricultural Products Demand Based on the ARIMA Model

[1]Haoxiong Yang and [2]Jing Hu
[1]School of Business,
[2]School of Computer and Information Engineering, Beijing Technology and Business University,
Beijing 100048, P.R. China

Abstract: The price of fresh agricultural products changes up and down recently. In order to accurately forecast the agricultural precuts demand, a forecasting model based on ARIMA is provided in this study. It can be found that asymmetric information and unbalance about supply and demand exist in the market through analyzing the reasons. The ARIMA model for fresh agricultural products can forecast the demand in order to providing some guides for farmers. The results show that the predictive value are in good condition when compare with the actual data. Then this model is available.

Keywords: ARIMA, fresh agricultural products, unsalable

INTRODUCTION

After the price soaring of vegetables some time ago, the news of poor sales of fresh agricultural products began to appear around the country last month. Unlike the previous years, much more varieties seem to be unmarketable ones this year. In addition to vegetables, fruits suffered from this as well and the amount seems larger, counting by tons. Accompanied with the slow selling of fruits and vegetables, media have called on people to buy "love vegetables" and "love fruits". Some officials also take action to help local farmers promote their products. It is not unfamiliar to confront this situation every year when the slow selling of vegetables and fruits happens. But it is regretful that the only thing we can do is just to jump to rescue rather than prevent the new trend of slow selling. The vegetable prices have fallen 20% nowadays than the same time last year and the price of some, like cabbages and celeries, have fallen 50%, according to the investigation in Shouguang of Shangdong Province, the vegetable distributing center.

The phenomenon that the price of fresh agricultural products ranged from time to time. Such price fluctuations exist in cabbage and celery which respectively has the price of only eight fens and one mao. This wildly fresh agricultural price fluctuation partly because of the asymmetry of fresh agricultural market information, concentration of producing, production varies with the seasons and the imbalance of supplying and demanding, etc. It goes without saying that the price of a product is inversely proportional to the amount of the production. Such as the farmers of

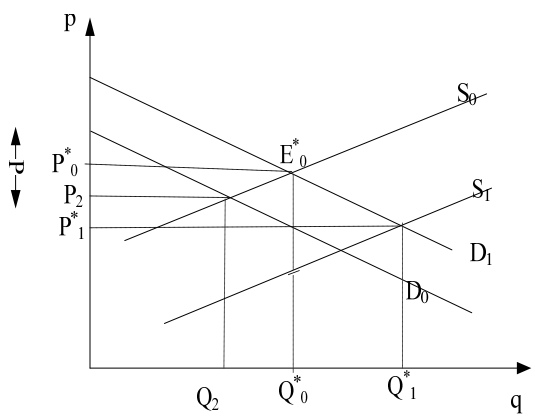

Fig. 1: supply and demand curve

South Korea, they are required to join in the peasant association (Choi *et al*., 2006). After growing vegetables, fruit or any other grains, those farmers never have any worries about agricultural products. And after their products becoming ripe enough, the following thing the farmers should do is only to transport the products to the food processing factories (Zia *et al*., 2011). With the help of Peasant Association, a lot of farmers markets have been set which can help determine the amount of agricultural products that should be produced and the following benefits is that agricultural products would never been over produced.

A high level of the price of the fresh agricultural product has bad effect on common person; on the contrary a low level of the price has bad effect on those poor farmers. The unstable price of fresh agricultural products may bring a fatal hurt to farmers, beside

Corresponding Author: Hoaxing Yang, School of Business, Beijing Technology and Business University, Beijing 100048, P. R. China

consumers may also be related. As is well known, the price of a product is varies with the amount of the products (Chang *et al*., 2006), Fig. 1. So the main idea of this study is to build a time sequence model to predict the demand of fresh agricultural products. And basing on this model we can find out a way to instruct the currency of fresh agricultural products and farmers are also included too.

MATERIALS AND METHODS

The basic idea of the ARIMA models is that it regards data objects which are formed by forecasting objects over time as a random sequence (Fan *et al*., 2010). It uses a mathematical model to approximately describe this sequence. This model can forecast the future from past value and the present value after it has been identified.

Let {} be a stationary sequence. Actual observed value , is present value which forecast and denoted by. $\hat{y}_t(1)$ Is prediction length. Prediction hope Variance of the Forecast Error for ARIMA Processes to minimize as much as possible, that is:

$$E(y_{t+1} - \hat{y}_t(1))^2 = \min$$

Therefore minimum variance forecasting is defined the best prediction. So the best predicted conditions of occurs in the condition of the happening of actual observed value, the conditional expectation of $\hat{y}_t(1)$ is:

$$\hat{y}_t(1) = E(Y_{t+1} / y_1, y_2, y_3、、、y_t)$$

Model building: IMA model is also called Autoregressive moving average model, which translates non-stationary time series into a stationary time series. Then the building of model is only to the variables' Hysteretic values, as well as stochastic error's current values and Hysteretic values. It takes sequences which are formed by the forecasting index over time as random sequences. Random variables which have a dependency reflect the continuity of the data at the time. They are not only under the influence of external factors, but also have their own Change Law. It regards data objects which are formed by forecasting objects over time as a random sequence. It uses a mathematical model to approximately describe this sequence. This model can forecast the future from past value and the present value after it has been identified.

The mathematical expression of ARIMA (p, d, q) model is:

$$\hat{y}_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \delta_j \varepsilon_{t-j}$$

If $\alpha_i = 0$, it will turn into

$$\hat{y}_t = \sum_{j=1}^{q} \delta_j \varepsilon_{t-j}$$

which is called moving average processes model (MA (q)).

If = 0, it will turn into

$$\hat{y}_t = \sum_{i=1}^{p} \alpha_i y_{t-i}$$

which is called autoregressive process model (AR (p)).

Among them AR is auto regression, p is autoregressive term; MA is moving average is the number of moving Average items and d is the difference when the time series are steady.

Methods: Annual fresh agriculture products' sales volume (as a demand) can be seen as a random time series over time, on which we can describe the demand with mathematical models through the analysis of certain factors including randomness, smoothness as well as seasonal effects so as to predict the demands in the future.

To form a stable random sequence adopting logarithms and differential processes

ARIMA modeling is based on a stationary column which requires that self value fluctuate around the sequence mean, namely, self value can't rise or drop sharply. Otherwise, we need to add differential smoothing treatment to the original data. Meanwhile, data should be technically processed unless the autocorrelation function values and partial autocorrelation function value close to zero. Parameter d in the model ARIMA (p, d, q) (d generally does not exceed 2) is the number of times of difference operation to turn the original stationary sequence into a non stationary sequence.

According to the rules of the time series model identification, the significance of selecting an appropriate model among lies in minimizing the parameter risks of the model. AR model is appropriate for the stationary sequence when its partial correlation function is truncated and its autocorrelation function is tailing; MA model suit the stationary sequence whose partial correlation function is trailing and the autocorrelation function is truncated; if the partial correlation function and autocorrelation function are both tailing, ARIMA model is supposed to be the best choice.

Specific algorithm is:

- Regarding sequential process as an ARIMA (p, d, q) model and making sure the order q. If the q is too large, then refuse MA model. Otherwise the MA model can be accepted. Then it can go on the following operation
- Regarding sequential process as a ARIMA (p, d, q) model. It can estimate order p and coefficient AR. Then it can go on the following operation

- Making sure the order of ARIMA (p, d, q) model. If q≠0, the original sequence is ARIMA (p, d, q) model, otherwise is AP (p) model
- Estimating model parameter and examining statistical significance
- Diagnosing residual series are white noise or not through hypothesis testing
- Using the model which has passed test to proceed forecast

Example analysis: This study applies ARIMA model to analyzing and forecasting cabbages' demand. According to the Table 1, the study will build ARIMA model. The tab lists the demand of cabbages between 2005 and 2010. The research designs to forecast the demand about 2010 using the method of random time series. According the predicted value, the farmers can decide how much plants they can plant.
(Data source: Statistics Bureau of P.R.C)

RESULTS AND DISCUSSION

According to the Table 2, it can be seen that demand is a non stationary sequence. Then the study can get a now sequence through differential transform. Now the study does some preprocessing to. Then it can get statistic, such as Z = 1.79<1.96. Compared with a significance level of 0.05, is steady. Afterwards the study will do difference to this sequence.

The order selection of model: According to Table 3, it can be seen that partial autocorrelation function is clipped at two-step department. Then when order is p = 3 and q = 2, the study sets up ARIMA (3, 1, 2) model. The characteristic is that Partial Autocorrelation Function is tailed at two-step department. First, this study process data fitting between ARIMA (3, 1, 1)×(1, 1, 0) model and ARIMA (3, 1, 2)×(1, 1, 0) model. Second compared with ARIMA (1, 1, 0)×(1, 1, 0) model, the value of AIC and SBC is minimum in ARIMA (3, 1, 2)×(1, 1, 0) model. The type and scale of model is decided through partial correlation functions (Jin and Zhen, 2007). The ARIMA (3, 1, 2)×(1, 1, 0) model is founded, which lays a foundation for further study of plants' demand, hence the precision is improved.

The test and forecast of data model: According to the consequence presenting in the Table 4, parameter estimations all get through the significance test. The residual series in this model generally are stationary sequences whose values are o. Their autocorrelation coefficients are located in the inner of confidence interval. In addition, there is no obvious show about statistics in almost all of the time point. Through testing with related data, the ARIMA (3, 1 and 2)×(1, 1 and 0) model has high fitting degree and applicability. By autocorrelation test about residual, it gets through white noise significance test (Luiz and Carlos, 2010).

The Table 5 shows mean values, minimum values, maximum values and percentile about eight fitting optimization indexes in this model. According to two R^2, the ARIMA (3, 1 and 2) model has high fitting degree. The smooth R^2 is 0.472, while R^2 is 0,296. Because variable data is seasonal data, stable R^2 are more representative.

The Table 6 shows parameter estimation values of ARIMA (3, 1, 2) model. There have two parts in ARIMA (3, 1, 2) modular and MA. The significance levels of AR are 0.000, 0.000 and 0.074. All of these items are very significant except AR (3). Therefore ARIMA (3, 1 and 2) is suitable.

Table 1: cabbage demand changes between 2005 and 2010

Year	2005	2006	2007	2008	2009	2010
Demand	2385	3065	2687	3297	3886	3765

Table 2 Partial auto covariance about Yt

Autocorrelation	PartialCorrelation	AC	PAC	Q-Stat	Porb
		0.429	0.429	6.32	0.009

Table 3: Partial auto covariance about yt

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Porb
		0.749	0.749	16.69	0.0000

Table 4: Correlation function of residual series char

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Porb
		-0.032	-0.032	0.049	0.009

Table 5: Model fitting

Model Fitting								
					Percentile			
Fitting statistics	Mean value	SE	Min. value	Max. value	5	15	50	90
Stationary R-square	0.472	.	0.472	0.472	0.47	0.47	0.47	0.47
R-square	0.296	.	0.296	0.296	0.296	0.296	0.296	0.296
RMSE	1.36	.	1.36	1.36	1.36	1.36	1.36	1.36
MAPE	215.3	.	215.3	215.3	215.3	215.3	215.3	215.3
Max. APE	13730.54	.	13730.54	13730.54	13730.5	13730.5	13730.5	13730.5
MAE	1.379	.	1.379	1.379	1.379	1.379	1.379	1.379
Max AE	5.824	.	5.824	5.824	5.824	5.824	5.824	5.824
Normalization of BIC	0s.898	.	0.898	0.898	0.898	0.898	0.898	0.898

Table 6: ARIMA model parameters

Model parameters		Estimation	S.E.	t	Siq
US spread model 1 AR, seasonality	Lag 1	0.902	0.129	7.142	0.000
Lag 2		-0.387	-0.091	-4.855	0.000
Lag 3		-0.134	-0.052	-1.788	0.076
Seasonal difference		1			
MA seasonality	Lag 1	1.442	0.121	11.538	0.000
Lag 2		-0.534	0.123	-4.528	0.000

On the basis of ARIMA (3, 1 and 2) model it can forecast the demand of cabbages about 2010 by using SPSS. The predicted value is 3342. Compared with actual value, the discrepancy is 11.3%. There have some deviations between predicted value and actual value, but they consistent with each other and keep the accuracy of 89%.

At present, the forecast is over. Producers can choose appropriate values in the interval combine with actual conditions. At last, they can make sure the most appropriate production.

CONCLUSION

The analysis above shows that it is feasible to predict the demand of agricultural products by ARIMA model in a short time. Besides, we can come to a conclusion that the longer predict time is and the larger the numerical prediction of the variance is. As a result, while comparing the result which we predict to the actual value it turns out to be that the deviation is larger. Because of the limit of the sample data, the model may unperfect. It requires appropriate revise to prediction equations. Only doing like this, the model can achieve best effect.

In general, ARIMA is a fairly practical model. In the forecasting process, it cannot separate the trends and seasonal part which are used in the original sequence. But the model can deal with the seasonal fluctuation forecasting which are caused by the temporal variation season fluctuation prediction. The prediction of time series model is very accuracy in the short term. As times increasing, the predicted error will gradually increase. But the model is relatively simple. So the requirement of information is much little. It has a very wide use in the actual situation.

ACKNOWLEDGMENT

The authors thank the National Social Science foundation of China (11CGL105), Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (PHR201108077) for support.

REFERENCES

Chang, H.C., L.Y. Ouyang, K.S. Wu and C.H. Ho, 2006. Integrated vendor-buyer cooperative inventory models with controllable lead time and ordering cost reduction. Eur. Oper. Res., 170(2): 481-495.

Choi, T.M., L. Duane and Y. Houmin, 2006. Quick response policy with Bayesian information updates. Eur. J. Oper. Res., 170(3): 788-808.

Fan, W., X. Ci, B. Fu and Y. Zhang, 2010. Flight demand forecasting model based international conference on services science. Manage. Eng., 6: 188-192.

Jin, Y. and S. Zhen, 2007. Regional agricultural logistics demand characteristics of the flow space. Logist. Econ., 19: 43.

Luiz, C.M.M. and A.S.L. Carlos, 2010. A new methodology for the logistic analysis of evolutionary S-shaped processes: Application to historical time series and forecasting. Technol. Forecasting Soc. Change, 77(2):175-192.

Zia, W., S. Himadri, M.D. Deyb, K. Ashfanoor and I.K. Shahidul, 2011. Modeling and forecasting natural gas demand in Bangladesh. Energ. Policy, 39(11): 7372-7310.