International Journal of Advanced Science and Technology
Vol.87 (2016), pp.31-36
http://dx.doi.org/10.14257/ijast.2016.87.04

Empirical Study on Agricultural Products Price Forecasting based on Internet-based Timely Price Information

Wang Xin[1] and Wang Can[2]

[1]*Notional Engineering Research Center for Agricultural Product Logistics, 250103, Ji'nan, China*
[1,2]*Shandong Institute of Commerce and Technology, 250103, Ji'nan, China*
[1]*ghostur@163.com,*[2]*wangcan24@126.com*

Abstract

The Web crawler technology enables timely and accurate collection of agricultural products price among 214 large-scale wholesale markets in China, which brings enormous convenient for a better observation of agricultural products future price trend. This article, which is based on market prices of various agricultural products collected by web crawler technology, develops quantitative analysis predictive models of wholesale prices. The predictive model of cucumber prices presents regularity of ARIMA (3,1,2) model. By comparing the relative errors between predicted and actual values, we find that ARIMA (3,1,2) model can provide high accuracy of short-term prediction for cucumber prices in Shandong Shouguang wholesale market.

Keywords: *ARIMA model; web crawler technology; time series; ultra-short term forecasting*

1. Introduction

In China, the price of agriculture produce has risen considerably in recent years. Much research related to the fluctuation of the agricultural products price has been conducted by domestic and overseas scholars. Methods of qualitative and quantitative analysis are employed by most scholars in order to describe the current price fluctuations and significant findings have been obtained. However, there are relative few studies on a specific market and certain produce. This article, based on actual data collected by crawler technology, attempts to carries on trend estimate of certain produce price on a specific market by using time series model. Through scientific analysis and forecast to historical prices and future price trend of agriculture produce, it can help farmers, peasant brokers, enterprises as well as government comprehensively.

2. Price Collection Timely by using Internet Technology

Since the launch of China Agricultural Information and China Agricultural Science and Technology websites in 1994, trading models (*e.g.*, online trading, futures trading) have appeared in succession. However, the massive amount of prices information generated on Internet has not been effectively integrated to form decisions that have directive meanings. It is found that focusing crawler technology is used by Xinxin Trading Platform of Agricultural Product for targeted data collection, which forms a humanized interface for data presentation. Therefore, we continually updated the data by crawler on the platform, providing solid data for further research.

ISSN: 2005-4238 IJAST

3. Empirical Analysis and Modeling Price of Agriculture Produce

3.1. Model Specification & Data Sources

ARIMA model is established by using past values and random perturbed variable. The model is generally referred to as ARIMA (p,d,q) where p, d, and p refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. ARIMA (p,d,q) is given by:

$$Xt = \varphi 1Xt-1 + \varphi 2Xt-2 + ... + \varphi pXt-p + \varepsilon t + \theta 1\varepsilon t-1 + ... + \theta q\varepsilon t-q \qquad (t=1,2,……,T)$$

The paper constructs ARIMA models of cucumber prices in Shandong Shouguang wholesale market in order to forecast short-term price.

Date used in this study is from Xinxin Agricultural Products Service Platform (www.xinxinjiage.com). The 955 samples are daily cucumber prices between 4 Jan 2011 and 15 Aug 2013. In the sections below, ARIMA model of cucumber prices is established to forecast short-term price between 16 Aug 2013 and 3 Sep 2013. The prediction error is obtained by comparing actual price with forecast price in order to estimate the plausibility of model and refine the model.

3.2. Model Building

3.2.1. Data Steadiness Testing

Figure 1 exhibits that cucumber price action has remained relatively steady in Shandong Shouguang wholesale market. Although the increase differs annually, the price peaked at both sides and it was low in the middle of each year.

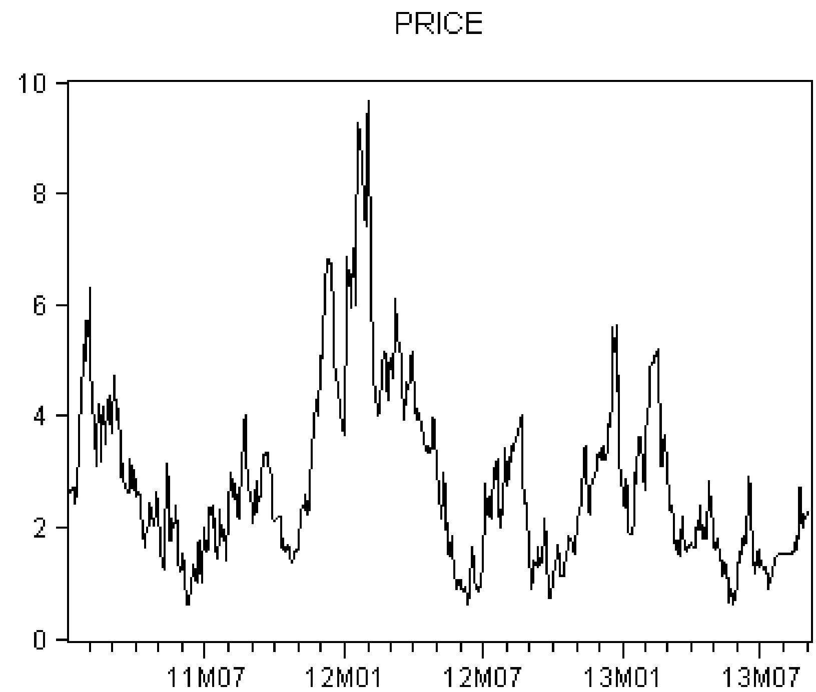


Figure 1. Time Series Plots of Cucumber Prices in Shandong Shouguang Wholesale Market

3.2.2. Parameter Estimation and Model Selection

Figure 2 shows the ADF test result of the original time series of prices.

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.315435	0.0145
Test critical values:	1% level	-3.436991	
	5% level	-2.864361	
	10% level	-2.568324	

*MacKinnon (1996) one-sided p-values.

Figure 2. ADF Test Result of the Original Time Series

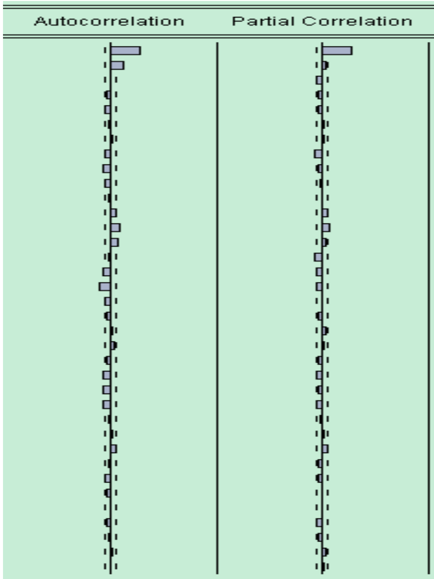
In Figure 2, ADF=-3.315435 and hypothesis of the existence of unit root is accepted, which indicates the original time series is non-stationary. The ADF test result after the first order difference is shown in Figure 3.

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-23.04397	0.0000
Test critical values:	1% level	-3.436991	
	5% level	-2.864361	
	10% level	-2.568324	

*MacKinnon (1996) one-sided p-values.

Figure 3. The ADF Test Result after the First Order Difference

In Figure 3, ADF=-23.14397 and hypothesis of the existence of unit root is rejected, which indicates the original time series after the first order difference is stationary. Therefore, we got d=1 for ARIMA (p,d,q) model of the original time series. The graphs of autocorrelation as well as partial correlation after the first order difference are shown in Figure 4.



	AC	PAC	Q-Stat	Prob
1	0.283	0.283	76.740	0.000
2	0.117	0.040	89.914	0.000
3	-0.006	-0.053	89.947	0.000
4	-0.044	-0.037	91.801	0.000
5	-0.051	-0.026	94.350	0.000
6	-0.009	0.020	94.429	0.000
7	0.010	0.013	94.528	0.000
8	-0.062	-0.080	98.195	0.000
9	-0.067	-0.039	102.51	0.000
10	-0.055	-0.017	105.47	0.000
11	-0.022	0.005	105.95	0.000
12	0.046	0.055	107.98	0.000
13	0.098	0.067	117.20	0.000
14	0.082	0.025	123.67	0.000
15	-0.024	-0.071	124.23	0.000
16	-0.076	-0.063	129.89	0.000
17	-0.102	-0.055	139.93	0.000
18	-0.050	0.006	142.37	0.000
19	-0.045	-0.031	144.31	0.000
20	0.018	0.030	144.62	0.000
21	0.028	0.023	145.37	0.000
22	-0.034	-0.045	146.48	0.000
23	-0.072	-0.056	151.56	0.000
24	-0.073	-0.043	156.72	0.000
25	-0.067	-0.052	161.07	0.000
26	-0.023	-0.011	161.58	0.000
27	0.023	0.015	162.08	0.000
28	0.049	0.042	164.48	0.000
29	-0.020	-0.032	164.87	0.000
30	-0.061	-0.048	168.50	0.000
31	-0.034	-0.002	169.67	0.000
32	-0.003	0.003	169.67	0.000
33	-0.026	-0.058	170.36	0.000
34	-0.021	-0.041	170.82	0.000
35	0.014	0.030	171.01	0.000
36	0.005	0.024	171.04	0.000

Figure 4. Autocorrelation Coefficients and Partial Correlation Coefficients after the First Order Difference

Analysis of Figure 4 reveals that the autocorrelation coefficients and partial correlation coefficients decrease. ARMA model is considered after the first order difference. Values of AC and PAC are significantly lower when q ≥3 and p ≥6 respectively, so we have q=2 and p=5. Therefore, ARMA (5,2) model is preliminarily established. Through applying AIC criterion, we can figure out the optimum model among models with 1≤ p ≤ 5. The result is shown in Figure 5.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.002975	0.006470	-0.459842	0.6457
AR(1)	1.055350	0.035867	29.42434	0.0000
AR(3)	-0.114586	0.025857	-4.431432	0.0000
MA(1)	-0.790379	0.045768	-17.26941	0.0000
MA(2)	-0.162649	0.037762	-4.307220	0.0000

Figure 5. Estimated Coefficients Testing of First Order Differences Models

As you can see from Figure 5, the estimates of explanatory variables are significant at 1% level. Therefore, ARIMA (3,1,2) is the optimum model.

3.3. Time Series Forecast and Test

The residuals sequence can be automatically generated after coefficients estimation. The residuals are white noise, which shows that the model provides a good fit to the real data.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.001	0.001	0.0009	
		2	0.005	0.005	0.0207	
		3	-0.011	-0.011	0.1333	
		4	-0.019	-0.019	0.4786	
		5	-0.024	-0.024	1.0253	0.311
		6	0.027	0.027	1.7391	0.419
		7	0.058	0.057	4.9194	0.178
		8	-0.035	-0.036	6.0927	0.192
		9	-0.027	-0.028	6.7933	0.236
		10	-0.021	-0.019	7.2280	0.300
		11	-0.012	-0.009	7.3719	0.391
		12	0.040	0.041	8.9042	0.350

Figure 6. Correlation Plots of Residuals

Through analyzing cucumber prices between 4 Jan 2011 and 15 Aug 2013, the optimal regression model was achieved. ARIMA (3,1,2) model is established to forecast short-term price between 16 Aug 2013 and 3 Sep 2013. The forecast effect is shown in Figure 7.

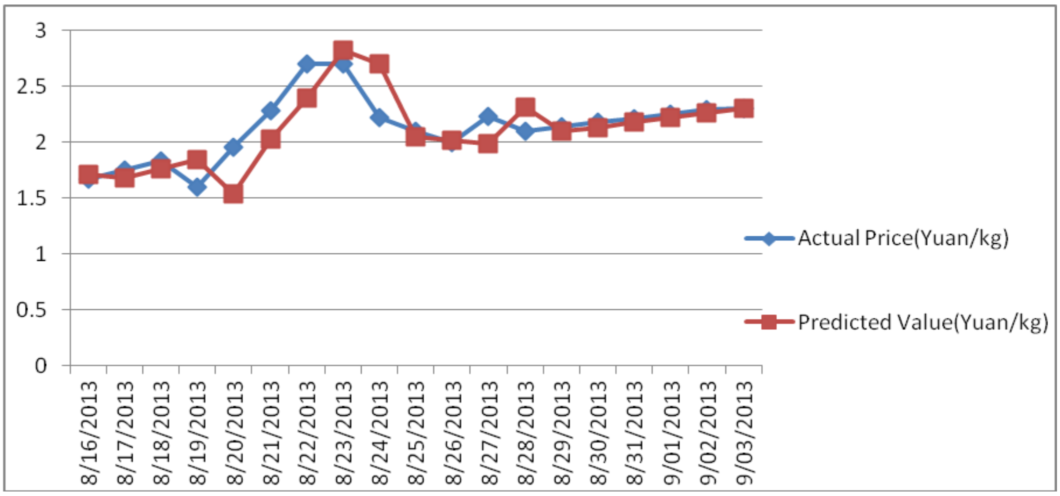


Figure 7. Forecasting Effect Diagram

The relative errors between actual and predicted prices are shown in Table 1.

Table 1. Model-predicted Data and Corresponding Forecasting Precision

Date	Actual Price (Yuan/kg)	ARIMA(3,1,2)	
		Predicted Value (Yuan/kg)	Error/%
8/16/2013	1.67	1.71	2.61
8/17/2013	1.75	1.68	-4.28
8/18/2013	1.83	1.76	-3.63
8/19/2013	1.6	1.85	15.68
8/20/2013	1.96	1.54	-21.50
8/21/2013	2.28	2.03	-10.77
8/22/2013	2.7	2.40	-11.23
8/23/2013	2.7	2.83	4.64
8/24/2013	2.22	2.70	21.51
8/25/2013	2.1	2.05	-2.31
8/26/2013	2	2.02	0.86
8/27/2013	2.23	1.99	-10.69
8/28/2013	2.1	2.32	10.40
8/29/2013	2.14	2.10	-2.09
8/30/2013	2.18	2.13	-2.33
8/31/2013	2.21	2.18	-1.19
9/01/2013	2.25	2.22	-1.33
9/02/2013	2.29	2.26	-1.23
9/03/2013	2.3	2.30	0.01

The precision of prediction model ARIMA (3,1,2) is satisfactory as a whole. The forecast model can control more than 50 percent of the predicted value relative error in 3%, more than 80% of predicted value relative error in 10%.

4. Conclusions and Suggestions

Due to the variety of influence factors and randomness of agricultural product price fluctuation, modeling the farm produce market price can be challenging. Therefore, this article carries on ultra-short term forecasting of agricultural product price by using time series model. The article establishes ARIMA(3,1,2) model to forecast short-term price between 16 Aug 2013 and 3 Sep 2013 based on daily cucumber prices between 4 Jan 2011 and 15 Aug 2013. Conclusions are as follows:

(1) From empirical analysis we can identify that the cucumber daily price fluctuates due to randomness and seasonal factors. The cucumber price is low in the middle of each year, which is probably a connection with the harvest season.

(2) The time series method based on ARIMA model has overcome the limitation of traditionally used casual model. The model without dependent variables can give a good prediction by using the historical data. More than 50% of the predicted value relative error can be controlled in 3% and more than 80% of the relative error can be controlled in 10%. Therefore, the practical application value of ultra-short term forecasting can be guaranteed. The ARIMA model can provide high accuracy of cucumber price prediction and can be extended to other agricultural products in other markets.

(3) The short-term fluctuations characteristics of farm produce price shows there are other factors of fluctuation besides seasonal factors. Further information transparency is necessary in order to further maintain agricultural market stability. This article applies scientific forecasting models to analyze farm produce price so that market players can be guided according to market changes, and therefore farmers, peasant brokers, enterprises as well as government can make decisions in relation to price forecast and market adjustment in a more rational way.

References

[1] G. Xu, “Statistical Forecasting and Decision-making”, Shanghai University of Finance and Economics Press, **(1998)**.
[2] G. E. P. Box and G. M. Jenkins, “Time Series Analysis Forecasting and Control”, Beijing: China Statistics Press, **(1997)**.
[3] X. Zhang, “Users Guide of Econometric Software Reviews (Version 2)”, Nankai University Press, **(2004)**.
[4] Z. Li and A. Ye, “Advanced Econometrics”, Tsinghua University Press, **(2000)**.
[5] Z. Wang and W. Li, “A Brief Analysis of Chicken Market and Market Forecasting”, Guide to Chinese Poultry, **(2004)**.
[6] G. Li, S. Xu, Z. Li and X. Dong, “Research on Ultra Short-term Market Price Forecasting of Agricultural Produce - Modern Time Series Method Modeling Based on Daily Wholesale Price of Tomatoes”, Journal of Huazhong Agricultural University, **(2010)**.
[7] F. Liu, R. Wang and C. Li, “Application of ARIMA model on Agricultural Produce Price Forecasting”, Computer Engineering and Applications, **(2009)**.
[8] X. Yao, H. Peng and Y. Zhu, “Modeling and Application of ARIMA Time Series Model on Farm-fresh Produce Price Forecasting”, System Sciences and Comprehensive Studies in Agriculture, **(2007)**.

Author

Wang Xin, Minister of Platform Operations of the Notional Engineering Research Center for Agricultural Product Logistics, Dean of Business Administration Academy of Shandong Institute of Commerce and Technology. Master's degree, Associate Professor. Professional Research field: E-commerce, Project management, Agricultural price analysis and forecasting research.