Forecast of Agricultural Crop Price using Time Series and Kalman Filter Method

Asia Pacific Journal of Multidisciplinary Research
Vol. 5 No.4, 15-21
November 2017 Part III
P-ISSN 2350-7756
E-ISSN 2350-8442
www.apjmr.com

Paulo V. Cenas
College of Computing, Pangasinan State University, Urdaneta City, Pangasinan, Philippines

Date Received: October 11, 2017; Date Revised: November 22, 2017

Abstract - *This study is an exploratory analysis of the possibility of improving the accuracy and precision of typical time series model in forecasting future prices of rice crop by combining the techniques of ARIMA and Kalman filter respectively. Using actual rice data collected over a period of five years, the performance of the typical ARIMA model was compared to the combined performance of ARIMA-Kalman filter using Mean Square Error(MSE) and Root Mean Square Error(RMSE) as the bases of comparison. Results of the analysis revealed that a more accurate and precise time series estimates of future price of rice can be achieved when the technique of Kalman filter is combined with typical ARIMA time series model. Further analysis showed that predicted values generated by the combined techniques from out- of -sample forecasts are fairly closer to the actual values. On the basis of the findings of this study, the development of a time series software that will work on combining the algorithms of the ARIMA and Kalman filter is recommended.*

Keyword: *ARIMA, Kalman Filter, Forecast, Time Series .*

INTRODUCTION

Agricultural price information is crucial to enable farmers and traders to make better decisions about how to grow, when to harvest, to which markets produce should be sent, and to whether store agricultural crop or not. Predictions of agricultural crop prices involve a large number of calculations based on different forecasting models.

Time series prediction refers to the process by which the future values of a system are forecasted based on the information obtained from the past and the current point data points. Time-series methods use time as independent variables to produce demand. In a time series, measurements are taken at successive points or over successive periods. The measurements may be taken every hour, day, week, month, or year, or at any other regular (or irregular) interval. Time series models are characterized by four components: trend component, cyclical component, seasonal component, and irregular component[1]. On one hand, Kalman filter is a device that separates a time series in two components, one called signal and the other is noise. Signals are contaminated by 'noise' from a variety of sources and it is only when estimated of true values are available that the next location may be estimated with any accuracy. The Kalman Filter is a very good tool and has been applied in a wide variety of estimation applications. The basic idea behind the filter is simple - to arrive at a conditional density function of the unobservable, the functional form of relationship with observable equation of motion and assumptions regarding the distribution of error terms. The filter uses the current observation to predict the next period's value of the unobservable and then uses the realization next period to update that forecast.

Recently, combinational tools and approaches have attracted researches in the field of forecasting. McIntosh and Bessler [2] conducted a study on forecasting agricultural prices using Bayesian approach utilizing matrix beta priors. The study explains the matrix beta approach and applies it to three individual forecasts namely ARIMA and ROLS approach using regression analysis and adaptive weighting technique. The results indicate that, given these data and a quadratic loss performance metric (MSFE), the analyst would have been better off than a composite forecast rather than attempting to identify a best individual forecast. Lately, Jha & Sinha [3] and Shahwan & Odening [4] has attracted their attention to Artificial Neural Network (ANN) as an alternative technique for estimation and forecasting. The superiority of ANN over ARIMA linear model methodology has been applied using monthly wholesale price series of agricultural products. The experimental analysis has shown that ANN models are able to attain a significant number of directions of monthly price change as compared to the ARIMA linear models. Further, it has also been observed that combining linear and nonlinear models lead to more accurate forecasts than the performances of these models independently.

A study by Moghaddam and Saeed Ebrahimijam [5] on a predictive model using extended Kalman filter which simultaneously fuses information and parameters of technical and fundamental analysis to predict short term stock prices. Then, as a real test, the model implemented for the shares of one of the industrial companies in Iran. Finally, the obtained results will be compared with other method results such as regression and neural networks which show its desirability in short-term predictions.

Therefore, due to its wide applicability and relevance, the importance of a thorough study on the performance of time series statistical techniques of fixed and open models is confirmed, specifically in the agricultural prices analysis context. Clements and Hendry [6] argue that a combination of forecasts may be superior to each of individual forecasts if it meets the following: the forecasting models do not draw from a common information pool, are of a different nature and type, or when models are differently sensitive to structural breaks and misspecification.

Choosing the right forecasting method to be used highly depends on the structure of time series, and as stated by Umar [7] to simplicity, accuracy and stability of the data. Choosing the best forecasting method for any particular situation is not a simple task Armstrong and Green [8] and sometimes more than one method may be appropriate based on accuracy measures such as MSE, MAD and other methods. This research attempts to formulate a new agricultural crop price forecasting model that has accurate predicting capability.

OBJECTIVES OF THE STUDY

This research attempted to forecast future prices of agricultural crop using time series and Kalman filter combined, more specifically tried to attain the following specific objectives: to characterize the data needed to produce accurate forecasts the prices of agricultural crops and to determine if there is an existing data to extract; and to improve the prediction accuracy of agricultural crop prices using a combination model of Time Series and Kalman filter prediction algorithms.

MATERIALS AND METHODS

Research Design

This research utilized exploratory research to determine the method of building a model of an accurate tool that will forecast the future agricultural crop prices. This study explored the possibility of combining two forecasting tools to generate a model that accurately forecasted the future agricultural crop prices using actual historical data.

Data

This study used purposive sampling to gather needed information because of specific purpose and the nature of the needed data in building a forecasting model. The agricultural crops price series was analyzed based on a series of monthly prices. These prices were nominal prices and were compiled from the National Statistics Office (NSO) and Philippine Agricultural Statistic Office.

Two periods were used to analyze the forecasting models, an in-sample and an out-of-sample period. The series from the in-sample period was used to generate the forecasting models, the out-of-sample period was utilized to check against the accuracy of the model based on the out –of –sample data. According to Dhuyvetter et al. [9] and Ye et al. [10] the main criterion for choosing this period is that its corresponding closeness to the period of desired time of forecasting. The latest data closely available to the desired period were from 1990 to 2014. Therefore, the in-sample period were series of agricultural crop prices from January 2007 to December 2011 and an out-of-sample period which was composed of series prices from January 2012 to December 2014.

Conceptual Framework

The relationship between previous agricultural crop prices and the process that will forecast the future prices of agricultural crops can be conceptualized at a fairly general level as depicted in the conceptual framework.

Phase 1. ARIMA Modeling Stage

Stage 1. Identification

On the identification stage, identify specific response agricultural crop price series and possible candidate data for ARIMA models to use. Stationarity tests can be performed to determine if differencing is necessary. In this analysis, there will be several outputs of ARIMA models that could be fit.

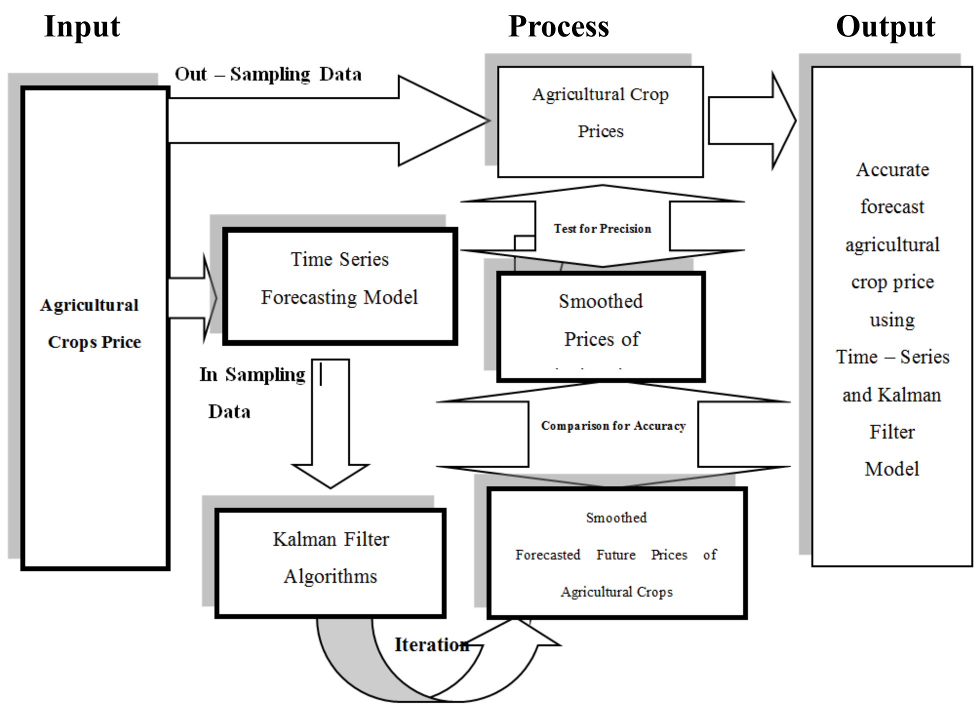

Figure 1

Stage 2. Estimation and diagnostic checking stage

In the estimation and diagnostic checking stage, the possible ARIMA model will be fit into the proposed agricultural crop price forecasting model and to estimate the parameters of that model. Under this stage, it will produce diagnostic statistics that help determine the adequacy of the model considering the properties of the residuals whether the residuals from an ARIMA model have the normal or random distribution.

Stage 3. Forecasting

In the forecasting stage, the initial output of the forecast model under time series will be further smoothed by using kalman filter algorithm.

Phase 2. Kalman Filter Algorithm

The initial agricultural crop price forecast will undergo a smoothing process using a Kalman Filter algorithm and will continue until a satisfactory model is arrived at. Undesirable and noisy data will be reduced/eradicated using an approximating function which attempts to capture important patterns to be able to produce a more accurate forecast model.

Phase 3. Check for Accuracy

The data that were treated by Kalman filter is compared to its accuracy with the data produced by the time series using Mean Square Errors (MSE) and Root Mean Squared Errors (RMSE). MSE is also called the average squared error of the forecast and is considered as a "non-parametric" statistic that indicates the size of the individual forecast errors from actual values deviations of forecasts from actual values are squared, ***larger errors*** reflect a significant decline in model performance. The square root of MSE, called the root mean squared error (RMSE) represents the mean size of forecast error, measured in the same units as the actual values. Like the standard deviation, the ***lower*** the RMSE the more accurate the forecast. The research examined accuracy standard using Mean Squared Errors (RMSE).

$$MSE = \sum(\hat{y}_1 - y_1)^2$$

$$RMSE = \sqrt{\sum_{t=T+1}^{T+h}(\hat{y}_t - y_t)^2}$$

Data Analysis

Time series autoregressive integrated moving average (ARIMA) and a Kalman filter algorithm is utilized to generate an agricultural crop price forecasting model. The time series ARIMA is based on the identification of standards (or components) in the historical series so that forecasts on a given variability could be given. The main components (or movements) of a time series comprise trend, seasonality, cycle and randomness. To analyze the historical development as well as the current situation, trend and seasonal indices are applied in this work. The ARIMA models have been used to make predictions, from the series of origin prices. Prices of agricultural crop will be forecasted using ARIMA Time Series and Time Series with Kalman Filter Algorithms models. These two tools are combined to come up with suitable model to forecast accurate future prices of agricultural crops.

The proposed model was constructed and tested following the four major phases:

- ARIMA Modeling Stage
- Kalman Filter algorithm implementation
- Accuracy check
- Precision check

RESULTS AND DISCUSSION

The detailed results of the simulated experiment via the two methods. It can be seen that the time series ARIMA with a total square mean of 4573615 and Kalman filter method with a total square error 3363455 outperforms time series ARIMA models in terms of MSE result. Focusing on the RMSE indicators, the overall value of the time series and the Kalman filter method is explicitly lower than Time Series ARIMA which indicated that the proposed model is more accurate in terms of price forecasting. Table 1 presents the data used in the simulation and the calculations for the two methods of analysis.

Table 1 shows the detailed results of the simulated experiment via the precision accuracy test between actual price and the forecasted price using time series and Kalman filter model.The large difference between the forecasting and actual values indicates that the model is not performing well. The result of the total error square difference is 1,053,086 which indicate that the precision accuracy is not performing well. There are possible reasons, first the proposed model is not precise due to limited experimental and second there is a need for more smoothing process using Kalman filter.

Table 1 Precision result of the Time Series ARIMA model and Time Series with Kalman filter Estimate

Date	Actual Price Per Sack	Time Series with Kalman Filter estimates	Time Series estimates	Time Series with Kalman Filter error square	Time series error square
Jan-07	757.19	795.53	865.128649	1470	11651
Feb-07	707.06	836.73	987.892853	16814	78867
Mar-07	1133.24	1331.42	889.892707	39275	59218
Apr-07	437.58	515.14	759.284167	6016	103494
May-07	440.51	515.51	683.933755	5625	59255
Jun-07	655.93	757.55	667.968613	10327	145
Jul-07	626.55	732.07	724.414494	11134	9577
Aug-07	693.04	728.77	758.55393	1277	4292
Sep-07	777.54	835.38	779.099138	3345	2
Oct-07	853.68	748.24	792.489611	11118	3744
Nov-07	909.08	819.67	816.870684	7994	8503
Dec-07	1523.27	848.82	812.600624	454883	505051
Jan-08	821.72	863.32	800.085323	1731	468
Feb-08	551.57	652.72	779.120931	10231	51779
Mar-08	670.75	788.05	794.016334	13759	15195
Apr-08	766.52	902.37	815.318381	18455	2381
May-08	690.88	808.51	804.73018	13837	12962
Jun-08	679.65	784.94	793.267711	11086	12909
Jul-08	560.08	654.4	796.585611	8896	55935
Aug-08	909.53	956.42	849.141318	2199	3647

Table 1 (Cont.) Precision result of the Time Series ARIMA model and Time Series with Kalman filter Estimate

Date	Actual Price Per Sack	Time Series with Kalman Filter estimates	Time Series estimates	Time Series with Kalman Filter error square	Time series error square
Sep-08	851.51	914.85	878.690971	4012	739
Oct-08	969.3	849.57	871.283747	14335	9607
Nov-08	995.16	897.28	839.403928	9580	24260
Dec-08	1278.01	712.15	804.492663	320198	224219
Jan-09	784.76	824.49	816.170264	1578	987
Feb-09	685.21	810.86	822.014032	15788	18715
Mar-09	790.51	928.75	834.693905	19110	1952
Apr-09	634.46	746.91	886.972937	12645	63763
May-09	671.62	785.96	973.23476	13074	90971
Jun-09	1300.3	1501.74	1041.6661	40578	66891
Jul-09	834.62	975.17	956.538675	19754	14864
Aug-09	571.15	600.59	812.728139	867	58360
Sep-09	644.75	692.71	726.355588	2300	6659
Oct-09	770.44	675.28	759.799753	9055	113
Nov-09	1033.84	932.16	838.842978	10339	38024
Dec-09	1726.42	962.02	928.924201	584307	636000
Jan-10	747.92	785.78	953.676342	1433	42336
Feb-10	1151.85	1363.08	975.737517	44618	31016
Mar-10	547.61	643.38	915.031843	9172	134999
Apr-10	740.78	872.07	918.777713	17237	31683
May-10	896.46	1049.09	934.640951	23296	1458
Jun-10	783.42	904.79	995.230096	14731	44864
Jul-10	913.56	1067.41	1021.40739	23670	11631
Aug-10	1081.85	1137.63	1014.35789	3111	4555
Sep-10	797.2	856.5	950.594209	3516	23530
Oct-10	1099.52	963.71	906.133318	18444	37398
Nov-10	793.82	715.75	862.153455	6095	4669
Dec-10	1762.16	981.94	879.755657	608743	778637
Jan-11	822.68	864.34	874.134561	1736	2648
Feb-11	716.68	848.11	885.953077	17274	28653
Mar-11	764.64	898.36	972.853327	17881	43353
Apr-11	783.13	921.92	1069.17136	19263	81820
May-11	1415.24	1656.2	1168.3025	58062	60978
Jun-11	778.88	899.55	1104.8099	14561	106230
Jul-11	860	1004.83	1064.14337	20976	41675
Aug-11	951.61	1000.67	1011.84465	2407	3628
Sep-11	1029.58	1106.17	1024.31041	5866	28
Oct-11	1121.57	983.04	1005.79144	19191	13405
Nov-11	1029.26	928.03	979.503502	10248	2476
Dec-11	1855.46	1033.93	968.472576	674912	786747
	Mean Square Error (MSE)			**3363435**	**4573615**
	Root Mean Squared Errors (RMSE)			1833.967	2138.601

Date	Actual price Per sack	Time Series with Kalman Filter error square	Error Square Difference
Jan-12	848.38	891.33	1844.7025
Feb-12	838.87	992.7	1844.7025
Mar-12	842.7	990.08	23663.6689
Apr-12	1029.27	1211.68	21720.8644
May-12	754.29	882.71	33273.4081
Jun-12	1063.78	1228.59	16491.6964
Jul-12	980.71	1145.86	27162.3361
Aug-12	1185.18	1246.29	27274.5225
Sep-12	1795.41	1928.97	3734.4321
Oct-12	1275.72	1118.14	17838.2736
Nov-12	1162	1047.72	24831.4564
Dec-12	2072.67	1154.96	13059.9184
Mean Square Error (MSE)			1,053,086.9

The graphical comparison between the actual price, Time Series ARIMA and Time Series with Kalman Filter model is reported in figure 2. The blue line graph shows the actual price of agricultural crops, the yellow line graph is the forecasted price and the green line graph represent the forecasted price produced by time series and Kalman filter model. As it shows, the Time Series with Kalman Filter model perform fairly well compared with the Time Series ARIMA model. Two main reasons can be used to explain the above observations. First, smoothing noise or unpredictable data were smoothed using a Kalman Filter algorithm which can effectively improve forecasting performance. Secondly, judgmental adjustment can also significantly improve forecasts and obtain more accurate forecasting results.

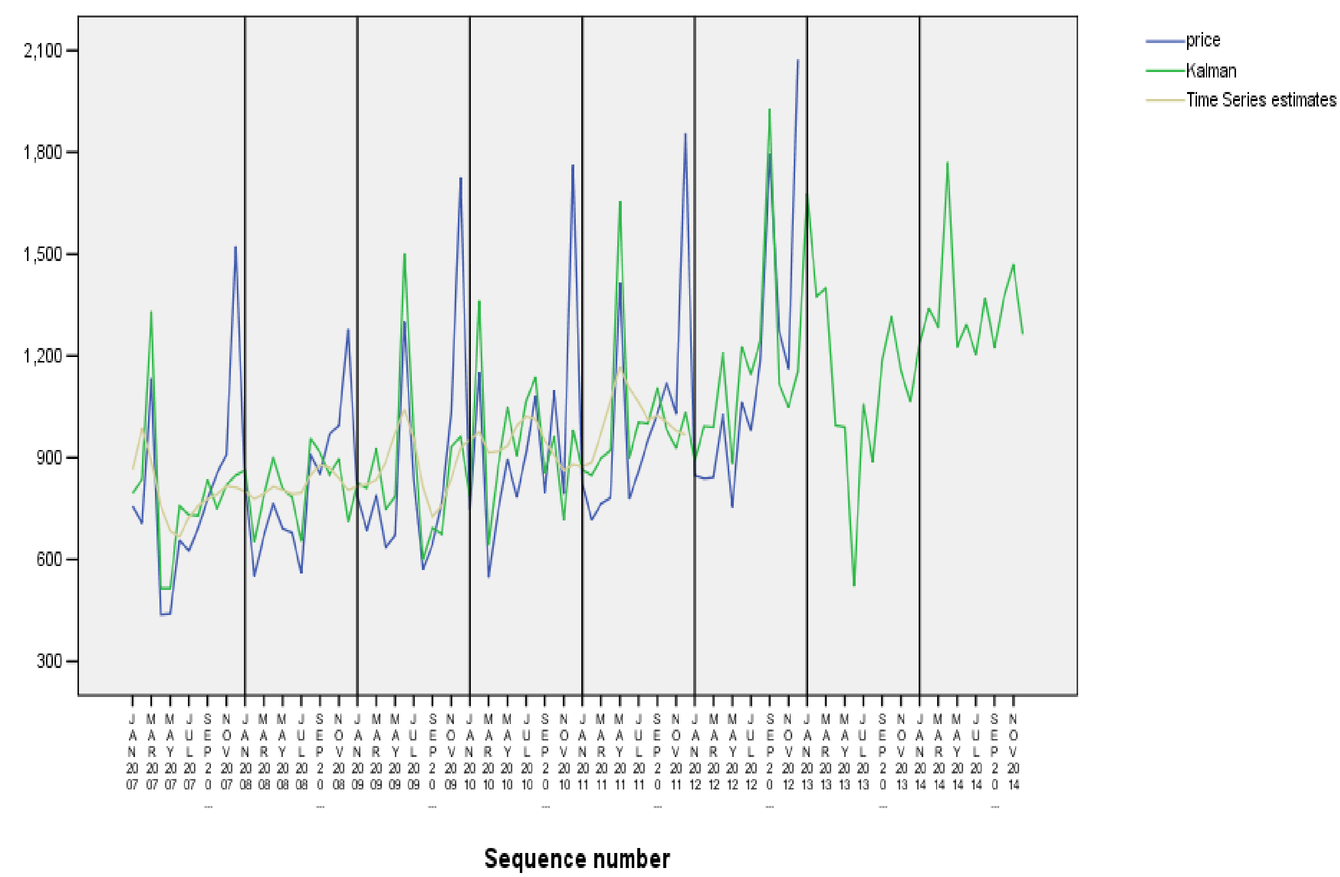

Figure 2 Actual Price, Time Series ARIMA and Time Series with Kalman Filter plot of Forecasted of Agricultural Crop Price from year 2007 to 2014.

CONCLUSION AND RECOMMENDATION

The results from the models Time Series ARIMA and Time Series with Kalman Filter Model provide efficient outcomes in terms of statistical results and predictable values for out-of-sample forecasting. The MSE and RMSE for the Time Series ARIMA model are significantly larger than the Time Series Kalman Filter Model which means the combination of two forecasting tools is more accurate. Furthermore, the predicted values generated by Time Series with Kalman Filter from out-of-sample forecasts are closer to the actual values.

The evaluation and analysis of forecasting performance of alternative forecasting models (or techniques) can be useful to the forecast user and the forecaster by providing some measure of confidence about the forecasts. Alternative performance criteria can be employed to assess the predictive ability of the forecasting models.

The need for accurate price forecasts is indeed overwhelming as decision makers to face commonly violent fluctuations in agricultural crop prices. Agricultural crop producers and market intermediaries require more accurate and useful forecasts in order to minimize the costs of wrong decisions.

For future work, the model can be applied to other agricultural crop products to test smoothing and accuracy performance. Further, other dimensions such as inflation, disaster and lead time will be included as other intervening variables to predict future prices.

REFERENCES

[1] Adhikari, R., & Agrawal, R. K. (2013). An introductory study on time series modeling and forecasting. *arXiv preprint arXiv:1302.6613*.

[2] McIntosh, C. S., & Bessler, D. A. (1988). Forecasting agricultural prices using a Bayesian composite approach. *Journal of Agricultural and Applied Economics*, *20*(2), 73-80.

[3] Jha, G. K., & Sinha, K. (2013). Agricultural Price Forecasting Using Neural Network Model: An Innovative Information Delivery System. *Agricultural Economics Research Review*, *26*(2).

[4] Shahwan, T., & Odening, M. (2007). Forecasting agricultural commodity prices using hybrid neural networks. *Computational Intelligence in Economics and Finance (Eds: Shu-Heng Chen, Paul P. Wang and Tzu-Wen Kuo), Springer, Berlin*, 63-74.

[5] Moghaddam, B. A., Haleh, H., & Ebrahimijam, S. (2011). Forecasting Trend and Stock Price with Adaptive Extended Kalman Filter Data Fusion. In *Proceedings of IEEE International Conference on Economics and Finance Research. Singapore:[sn]* (pp. 119-123).

[6] Clements, M., & Hendry, D. (1998). *Forecasting economic time series*. Cambridge University Press.

[7] Umar, M. A. (2007). Comparative Study of Holt-Winter, Double Exponential and the Linear Trend Regression Models, With Application to Exchange Rates of the Naira to the Dollar. *Research Journal of Applied Sciences*, *2*(5), 633-637.

[8] Armstrong, J. S., & Green, K. C. (2005). *Demand forecasting: evidence-based methods* (No. 24/05). Monash University, Department of Econometrics and Business Statistics.

[9] Dhuyvetter, K.C., E. Dean, and J.L. Parcell. (2003). "Using The CrudeOil And Heating Oil Markets For Diesel Fuel Purchasing Decisions." Journal of Agribusiness,21: 213-229.

[10] Ye, M., J. Zyren, and J. Shore. (2005). "A Monthly Crude Oil Spot Price Forecasting Model Using Relative Inventories." International Journal of Forecasting, 21: 491-501