A Comparative Study of ARIMA and Holt-Winters Exponential Smoothing Models for Rice Price Forecasting in Tanzania

Yohana James Mgale[1,2], Yunxian Yan[1], Shauri Timothy[2]

[1]College of Economics and Management, Jilin Agricultural University, Changchun, China
[2]Department of Rural Development and Regional Planning, Institute of Rural Development Planning, Dodoma, Tanzania
Email: mgaleyj@gmail.com, ymgale@irdp.ac.tz

Abstract

Time series models are promising tools for forecasting commodity prices. However, their applications to guide producers in agricultural investments and marketing decisions are still limited. This article compares the ARIMA and Holt-Winters Exponential Smoothing models in terms of forecasting the monthly wholesale rice price in Tanzania. Even with very little difference, the Holt-Winters additive model showed the best results for forecasting rice prices compared to the ARIMA model. Thus, both models can be used to forecast the prices of agricultural products.

Subject Areas

Business Finance and Investment, Economics

Keywords

Price Forecasting, ARIMA, Holt-Winters Exponential Smoothing, Rice, Tanzania

1. Introduction

One of the main challenges facing smallholder farmers in developing countries is fluctuating prices. Agricultural prices tend to vary and fluctuate rapidly over time, as prices always adapt to various changes [1]. High price fluctuations can influence production decisions as it becomes difficult for the producer to predict the exact level of future prices or to adjust the timing of the sale [2]. Price fluctuations also offer traders the opportunity to manipulate price information at the farm level [3]. A risk-averse producer will produce less output when faced

with price uncertainty and vice versa, ultimately affecting food supplies availability. Price uncertainty also hurts poor consumers as they have to reduce their food consumption or, in extreme cases, simply starve. In addition to its impact on individuals, price volatility can significantly impact economic and social resilience, and hamper economic growth [4] [5] [6].

Tanzania is primarily an agricultural country, and rice being among the main crop produced; it's associated with risks and uncertainties [7]. Production risks include losses of drought, pests, and disease, which can be minimized by the appropriate use of technology and agricultural insurance [8]. On the other hand, price risks are more difficult to avoid because prices vary according to the supply and demand behavior of commodities in domestic and foreign markets [2]. In general, smallholder rice farmers are exposed to risks of price fluctuations that can compromise business profitability and, therefore, need to invest in innovative planting techniques and make predictions of the future price situation to plan the investments and costs [7] [9]. One of the alternatives to reduce uncertainty in the economic decision-making process is to use time series forecasting models [10].

Given the economic importance of rice in Tanzania, which is the second most important crop after maize, the most traded crop than any other food crop, employs over 1.6 million people and is highly produced by small farmers (about 90 percent) [11] [12]. This article aims to use time-series statistical models to forecast rice prices in Mbeya region (one of the major rice-producing regions in Tanzania) and identify the most suitable model. The study is justified by the fact that there is little research with time series models to predict commodity prices. Such analysis can provide tools that help farmers and policymakers in their decision-making. In this study, the models to be tested are the Holt-Winters Exponential Smoothing and the ARIMA model, models that can predict trends and seasonality. The rest of the paper is structured as follows: Part 2 describes the data and methods, which constitute the data types and econometric techniques used in this article. Part 3 illustrates the results, while Part 4 offers conclusions.

2. Data and Methods

2.1. Data

The data used for empirical analysis are monthly rice prices (quoted in Tanzanian shillings per 100 kg) marketed at wholesale markets in the Mbeya region from January 2004 to September 2019, totaling 189 observations. The data set was obtained from the Ministry of Industry and Trade, Tanzania[1]. The free software R was used to develop the ARIMA and Holt-Winters Exponential Smoothing models. For the choice of the best model, we reserved the last nine observations of the series for comparison with the predicted values. The descriptive results for the wholesale rice prices including mean, maximum, and standard deviations are presented in Table 1.

---

[1]See http://www.mit.go.tz/.

Table 1. Descriptive results of the nominal wholesale rice prices[a].

Measure	Values
Mean	116,371.23
Median	108,541.57
Standard Deviation	47,876.44
Coefficient of variation	0.41
Minimum	44,545.62
Maximum	229,375.35
Observations	189

Note: [a]Prices are in (1000) Tanzanian shillings.

2.2. Method

The use of the ARIMA and Holt-Winters exponential smoothing models to forecast rice prices relied on their ability to predict time series data showing both trend and seasonal variation. The objective is to choose the best model for predicting the prices of agricultural products.

2.2.1. ARIMA Model

Among the statistical models developed to predict future values based on a historical database is the Autoregressive Integrated and Moving Averages (ARIMA) model. Proposed by Box & Jenkins [13], the model is suitable for describing stationary and non-stationary series. The methodology consists of fitting ARIMA models (p, d, q) to a data set. Where p is the number of terms of the autoregressive part (AR), d is the number of successive differences, and q is the number of terms of the moving average part (MA).

The procedure for searching a stochastic model that represents a time series is an iterative process consisting of three phases: Identification phase, where the condition of stationarity, the behavioral structures such as trend, seasonality, periodicities, and the issues of data correlation (autocorrelation) are verified; Estimation phase through which suitability of the considered model is validated based on estimated parameters; and Diagnosis phase where the model adjustment through residual analysis, statistical tests, and model selection/adequacy criteria is verified [13].

When a time series has a time-dependent mean and variance, it is not stationary, and it will need to be transformed. The most common transformation is to take successive differences from the original series to a stationary series [14].

The first difference of Z_t is defined by:

$$\Delta Z_t = Z_t - Z_{t-1} \tag{1}$$

The second is given by:

$$\Delta^2 Z_t = Z_t - 2Z_{t-1} + Z_{t-2} \tag{2}$$

In typical situations, according to the authors, it will be sufficient to take one

or two differences for the series to become stationary. The number of differences needed to make the series stationary is called the order of integration *d*. The inclusion of the integration order term allows the use of the ARIMA models (*p*, *d*, *q*) given by the equation:

$$\Delta^d Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{3}$$

where $\phi_1, \cdots, \phi_p$, are the parameters for the autoregressive part and $\theta_1, \cdots, \theta_2$, are the moving average, ε_t is an error that cannot be estimated from the model, *d* represents the number of differences applied in the series.

As the main objective of the ARIMA model is to generate future forecast values, the paper uses the Akaike Information Criterion (AIC) proposed by Akaike [15] to assess the degree of accuracy of the forecasts.

$$\text{AIC} = \ln\left(\hat{\sigma}_e^2\right) + \frac{2(p+q)}{n} \tag{4}$$

where $\hat{\sigma}_e^2$ is estimated error variance; *n* is the sample size, and *p*, *q* are the parameter values.

From the above criteria values, it is possible to choose the most suitable model to make predictions *h* steps ahead; the lower the criteria values, the more suitable the model for making predictions. In addition to the selection criteria, we analyzed the white noise condition of the waste through the Ljung-Box test [16].

2.2.2. Holt-Winters Exponential Smoothing Model

Holt-Winters exponential smoothing model can be divided into a multiplicative model and additive model. The choice of the model depends on the seasonal component of the series.

The model equations are given by:

$$Z_t = (\beta_0 + \beta_1 t) + SN_t + IR_t \rightarrow \text{Additive} \tag{5}$$

$$Z_t = (\beta_0 + \beta_1 t) \times SN_t \times IR_t \rightarrow \text{Multiplicative} \tag{6}$$

With three smoothing equations (Equations (7) to (9)):

$$l_t = \alpha\left(y_t / sn_{t-l}\right) + (1-\alpha)\left(l_{t-1} + b_{t-1}\right) \rightarrow \text{Level} \tag{7}$$

$$b_t = \gamma\left(l_t - l_{t-1}\right) + (1-\gamma)b_{t-1} \rightarrow \text{Trend} \tag{8}$$

$$sn_t = \delta\left(y_t / l_t\right) + (1-\delta)sn_{t-l} \rightarrow \text{Seasonal} \tag{9}$$

where (γ, and δ) are smoothing constants, *l* is number of seasons in a year, SN_t is a seasonal pattern and IR_t refers to irregular components.

3. Empirical Result

3.1. Estimation of ARIMA Model

Before starting a time series analysis, it is necessary to identify the stationarity properties of the data series. The nature of the series (**Figure 1(a)**) leads us to differentiate the series to make it stationary (**Figure 1(b)**).

The autocorrelation (ACF) and partial autocorrelation (PACF) graphs (**Figure**

2) indicates which model to use. The differentiated ACF graph gives us an indication of a significant autoregressive parameter (p), and the PACF graph demonstrates one parameter for moving averages (q).

In order to identify the best ARIMA model, we tested several competing models. Table 2 shows the best significant ARIMA models for the rice price series studied.

According to the Akaike information criterion (AIC) [15], the best model to predict the rice prices studied is ARIMA (0, 1, 3) (2, 1, 0) [12] (see Table 2). To confirm the model's fit, we checked whether it meets the white noise assumptions. Figure 3(a) shows the distribution of the residuals and Figure 3(b) verifies that the errors are concentrated in the range of (−2, 2), which means that there is no variance, *i.e.*, the errors are around the mean. Looking at the ACF
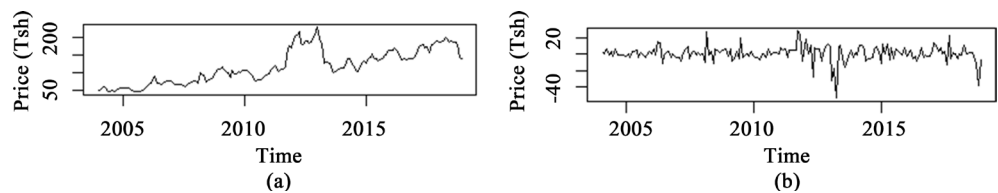

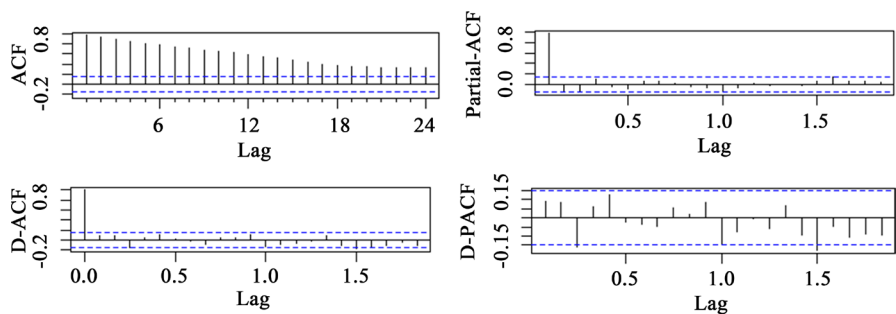**Figure 1.** (a) Original price series; (b) Price series with the first difference.

Figure 2. Autocorrelation (ACF) and Partial Autocorrelation (PACF).

Table 2. Competing ARIMA models.

ARIMA	AIC
(0, 1, 3) (2, 1, 0) [12]	1279.418
(3, 1, 0) (2, 1, 0) [12]	1281.175
(1, 1, 2) (2, 1, 0) [12]	1282.015
(0, 1, 0) (2, 1, 0) [12]	1284.557
(2, 1, 1) (2, 1, 0) [12]	1284.995
(0, 1, 2) (2, 1, 0) [12]	1285.58
(1, 1, 0) (2, 1, 0) [12]	1286.215
(0, 1, 1) (2, 1, 0) [12]	1286.297
(2, 1, 0) (2, 1, 0) [12]	1286.623
(1, 1, 1) (2, 1, 0) [12]	1288.122

Note: A good model is the one that has minimum AIC among all the other models (Akaike, 1974).

and PACF plots (Figure 4), we can see that there is no significantly non-zero autocorrelation in any lag, so the data are independent and demonstrates a good fit for the model.

After identifying the best fit model, the next step was to forecast rice prices for the next 24 months. The results of the fitted values (180 months) (Figure 5(a)) and the predicted values (Figure 5(b)) show that the forecast prices for both periods are within the allowable range and can therefore be justified in predicting future rice prices.

To further justify the results, we compared the forecast results of the ARIMA model with nine steps ahead and the actual rice prices for the same period. The results in Table 3 show that the predicted price values for the specified period are very close to the actual price values. In addition, the actual values are within the low and high limits at both the 80 and 95 percent significance level. Based on these results, it is clear that the ARIMA model can be used to forecast the prices of agricultural products.

3.2. Estimation of Holt-Winters Model

The Holt-Winters model falls into two categories, additive and multiplicative. In order to choose the best model, we first find the fitted values for the additive and multiplicative models. In the second step, we compared the fitted values of the models to the original series. The model that best corresponds to the price series studied was then chosen to compare with the ARIMA model. According to Figure 6(a) and Figure 6(b), the best model using the selection criteria was the
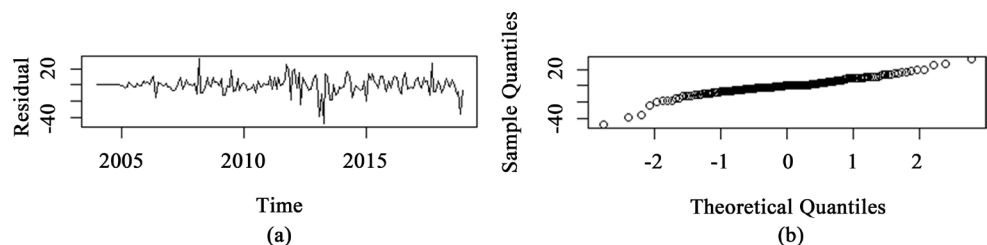

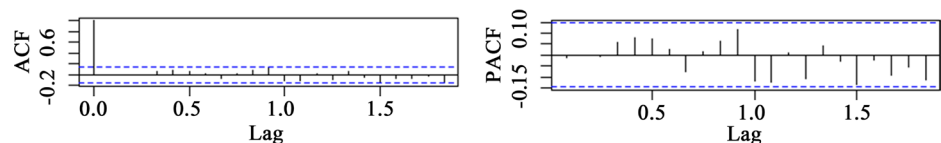**Figure 3.** (a) Graphs of residuals; (b) QQ-Plot.

Figure 4. Autocorrelation (ACF) and partial autocorrelation (PACF) function of Residual.
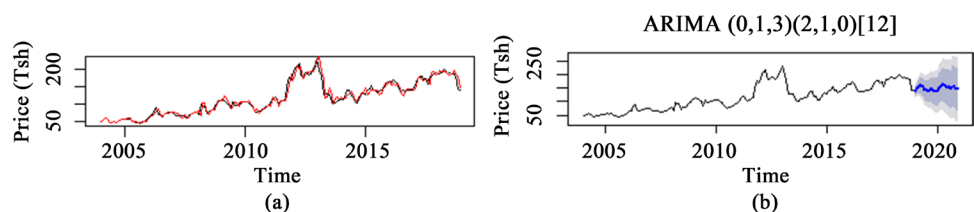

Figure 5. (a) Fitted values within the observed period; (b) Forecasts with 24 months ahead.

Holt-Winters additive model.

3.3. Model Comparison

In order to compare the two models, we used the MAPE (Mean Absolute Percent Error) selection criteria. As proposed in the methodology, the last nine observations of the nominal rice price series were reserved for the comparison of the proposed models. The results in Table 4 show that both models could be

Table 3. Predicted against real values (h = 9).

Period	Predicted Price Values					Actual Price values
	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95	
Jan 2019	137.676	124.102	151.249	116.917	158.435	140.120
Feb 2019	149.031	128.863	169.199	118.187	179.876	143.630
Mar 2019	159.297	132.958	185.636	119.015	199.579	150.450
Apr 2019	158.986	129.149	188.822	113.356	204.616	153.675
May 2019	153.901	120.937	186.865	103.487	204.315	145.678
Jun 2019	140.985	105.165	176.804	86.203	195.766	133.750
Jul 2019	142.618	104.154	181.082	83.792	201.443	130.950
Aug 2019	138.452	97.514	179.389	75.842	201.061	140.120
Sep 2019	144.117	100.846	187.388	77.939	210.294	143.630

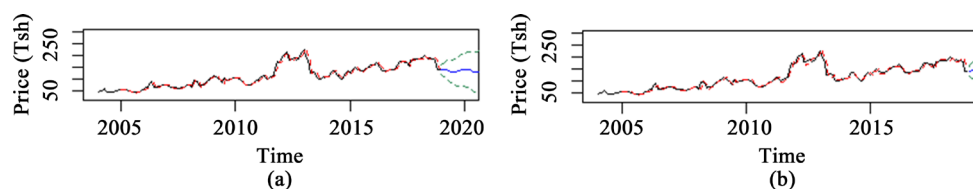Note: Prices are in (1000) Tanzanian shillings.

Figure 6. (a) Additive model; (b) Multiplicative model.

Table 4. Forecast results of the models under analysis (h = 9).

Period	Holt-Winters (*MAPE* = 5.3314)	ARIMA (0, 1, 3) (2, 1, 0) [12] (*MAPE* = 5.7512)	Actual price values
Jan 2019	140.393	137.675	140.120
Feb 2019	141.033	149.031	143.630
Mar 2019	147.561	159.297	150.450
Apr 2019	148.129	158.986	153.675
May 2019	141.136	153.901	145.678
Jun 2019	128.449	140.985	133.750
Jul 2019	128.574	142.618	130.950
Aug 2019	123.266	138.452	125.500
Sep 2019	123.330	144.117	127.496

Note: Prices are in (1,000) Tanzanian shillings.

used to forecast rice prices studied. However, according to the MAP selection criteria, the price values predicted by the Holt-Winters additive model are closer to the original rice prices than those predicted by the ARIMA model (see Table 4). Thus, on the basis of these results, the Holt-Winters additive model is more efficient in predicting the prices of agricultural products than the ARIMA model.

4. Conclusions

In this study, we looked at forecasting agricultural commodity prices using the time series model. The data used are the wholesale rice prices in Tanzania for the period January 2004 to September 2019. The study found that both of the studied models (Holt-Winters and ARIMA) performed well in predicting rice price values. However, in the model comparison, although by very little difference, the Holt-Winters additive model was closer to the actual data. Thus, it is clear from this article that time series can be used to forecast the prices of agricultural products, not only rice but also other crops, thus helping producers, policymakers and other actors to make informed production and marketing decisions.

Furthermore, this study is limited to rice prices in Tanzania. Therefore, similar studies on other agricultural products are also recommended.

Authors' Contributions

YJM designed and analyzed the statistical data for the study. ST contributed to the data collection and analysis. YY supervised the study. All authors have read and approved the final and revised version of the manuscript.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Food and Agriculture Organization of the United Nations, Organization for Economic Cooperation and Development, International Fund for Agricultural Development, International Food Policy Research Institute, International Monetary Fund, United Nations Conference on Trade and Development, *et al.* (2011) Price Volatility in Food and Agricultural Markets: Policy Responses. Technical Report, Food and Agriculture Organization of the United Nations; Organization for Economic Cooperation and Development, Rome. https://doi.org/10.1596/27379

[2] Sirisupluxana, P. and Bunyasiri, I.N. (2018) Risk Assessment and Risk Management Decisions: A Case Study of Thai Rice Farmers. *The Business & Management Review*, **9**, 200-207.

[3] Seck, P.A., Tollens, E., Wopereis, M.C., Diagne, A. and Bamba, I. (2010) Rising Trends and Variability of Rice Prices: Threats and opportunities for Sub-Saharan Africa. *Food Policy*, **35**, 403-411. https://doi.org/10.1016/j.foodpol.2010.05.003

[4]   Headey, D.D. (2011) Rethinking the Global Food Crisis: The Role of Trade Shocks. *Food Policy*, **36**, 136-146. https://doi.org/10.1016/j.foodpol.2010.10.003

[5]   Worako, T.K., Jordaan, H. and van Schalkwyk, H.D. (2011) Investigating Volatility in Coffee Prices along the Ethiopian Coffee Value Chain. *Agrekon*, **50**, 90-108. https://doi.org/10.1080/03031853.2011.617865

[6]   Ismail, A., Ihsan, H., Khan, S.A. and Jabeen, M. (2017) Price Volatility of Food and Agricultural Commodities: A Case Study of Pakistan. *Journal of Economic Cooperation and Development*, **38**, 77-120.

[7]   Nkuba, J., Ndunguru, A., Madulu, R., Lwezaura, D., Kajiru, G., Babu, A., *et al.* (2016) Rice Value Chain Analysis in Tanzania: Identification of Constraints, Opportunities and Upgrading Strategies. *African Crop Science Journal*, **24**, 73-87. https://doi.org/10.4314/acsj.v24i1.8S

[8]   Pochara, F. (2012) Commodities Exchange: Options for Addressing Price Risk and Price Volatility in Rice. African Development Bank, Abidjan.

[9]   Achandi, E.L. and Mujawamariya, G. (2016) Market Participation by Smallholder Rice Farmers in Tanzania: A Double Hurdle Analysis. *Studies in Agricultural Economics*, **118**, 112-115. https://doi.org/10.7896/j.1528

[10]   Kantanantha, N., Serban, N. and Griffin, P. (2010) Yield and Price Forecasting for Stochastic Crop Decision Planning. *Journal of Agricultural, Biological, and Environmental Statistics*, **15**, 362-380. https://doi.org/10.1007/s13253-010-0025-7

[11]   National Bureau of Statistics (Tanzania) (2017) 2016/17 Annual Agriculture Sample Survey Crop and Livestock Report. National Bureau of Statistics (Tanzania), Dodoma. http://www.nbs.go.tz

[12]   Wilson, R.T. and Lewis, I. (2015) The Rice Value Chain in Tanzania. A Report from the Southern Highlands Food Systems Programme. The Food and Agriculture Organization of the United Nations, Rome.

[13]   Box, G.E., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2015) Time Series Analysis: Forecasting and Control. John Wiley & Sons, Hoboken.

[14]   Montgomery, D.C., Jennings, C.L. and Kulahci, M. (2015) Introduction to Time Series Analysis and Forecasting. John Wiley & Sons, Hoboken.

[15]   Akaike, H. (1974) A New Look at the Statistical Model Identification. In: Parzen, E., Tanabe, K. and Kitagawa, G., Eds., *Selected Papers of Hirotugu Akaike*, Springer, New York, 215-222. https://doi.org/10.1007/978-1-4612-1694-0_16

[16]   Box, G.E. and Jenkins, G.M. (1976) Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.