International Journal of Tropical Agriculture
Volume 43, Numbers (1-2) • 2025 : pp. 111-117
ISSN: 0254-8755

An Application of a Univariate Model to Forecast Millet Production in India

Prema Borkar

Assistant Professor, Gokhale Institute of Politics and Economics (Deemed University), Pune, Maharashtra. India.
E-mail: prema.borkar@gipe.ac.in, premaborkar@gmail.com

Abstract: Millets are small-seeded crops classed as Major (Sorghum, Pearl millet, and Finger millet), Minor (Foxtail, Kodo, Banyard, Little, and Proso millets), and Pseudo (Amaranth and Buckwheat). Millets are referred to be nutri-cereals since they are highly nutritious and can greatly contribute to food and nutritional security. India is the world's largest millet grower, accounting for 43 percent of global production. India, Niger, China, Nigeria, and Mali produce 73 percent of the millet produced. India ranks 31st in the world in terms of yield per hectare. Labeling these crops as Nutri-cereals has been promoted in order to enhance demand and allow farmers to seek higher prices. According to the government, millets have a high potential for considerably contributing to the country's food and nutritional security, and hence they are not only a nutrient powerhouse, but also climate-adaptable crops with unique nutritional features.

The purpose of this study is to develop an ARIMA model for forecasting millets production in India. The annual data were obtained from 1966-67 to 2023-24 from the Directorate of Economics and Statistics, Ministry of Agriculture and Farmers' Welfare, Government of India. The data was fitted using the ARIMA model in order to project future millet production in India. The univariate Box-Jenkins (1970) ARIMA approach has been used for forecasting. The model parameters were estimated using the R programming language. The fitted model's performance was evaluated using multiple measures of goodness of fit. Projections for the years 2024-25 to 2029-30 are calculated using the chosen model. The numerical and graphical results of millet production are displayed graphically. These estimates are especially useful for altering marketing and policy campaigns to account for future changes.

Keywords: ACF – Autocorrelation Function, ARIMA – Autoregressive Integrated Moving Average, MAPE, Millets, PACF – Partial Autocorrelation Function, Residual Analysis.

INTRODUCTION

Millets are traditional crops with exceptional nutritional content and health advantages; nevertheless, millet products are becoming more popular with increasing lifestyle disorders, particularly in urban areas. Millets are robust crops grown in dry and semiarid regions, resistant to high temperatures and drought. Millets require 350 mm of water, whereas rice requires 1200 mm. They can be planted in intercropping (or possibly under mixed cropping with pulse and oilseeds) and provide food, fodder, fuel, and nutritional security. Millets play a significant role in sustainable agriculture and a healthy global community. They represent the regional food systems and cultures of Asia and Africa.

Millets offer financial stability since they are crops resistant to climate change and a reliable source of revenue for farmers. As we face dwindling natural resources and a

Article History: Received: 06 May 2025; Revised: 20 May 2025; Accepted: 02 June 2025; Published: 30 June 2025

changing environment in the twenty-first century, millets, which are hardy crops that can withstand severe weather conditions like drought and flooding, may be regarded as the perfect crop. Millets are required as the climate crisis worsens because they can be grown in the harshest conditions and help offset the demand that rice and wheat may not be able to meet due to the more pronounced consequences of climate change. India is the world's greatest producer and second-largest exporter of millets producing nine generally known millets.

According to APEDA, the world production of millets in 2022 was around 30.9 million tonnes from an area of 74.00 million hectares. In 2024, India remained the world's largest producer of millets, accounting for 43 percent of global production, with 13.5 million metric tons. Niger was the second largest producer, contributing 11 per cent with 3.4 million tons, followed by China with 9 per cent and 2.7 million tons. Mali was the fourth largest producer with 6 per cent and 1.8 million tons. The two most common millet crops farmed worldwide, sorghum and pearl millet, account for almost 90 percent of the world's millet production. Finger, Foxtail, Proso, Barnyard, Little, and Kodo millets are the next most popular millets.

Millets are grown in around 21 Indian states. Rajasthan, Maharashtra, Karnataka, Andhra Pradesh, Tamil Nadu, Kerala, Telangana, Uttarakhand, Jharkhand, Madhya Pradesh, Haryana, and Gujarat are leading the way. Millets are grown on an area of 12.19 million hectares in India, yielding 15.38 million tonnes at a yield of 1247 kg/ha. In terms of acreage (3.84 million hectares) and production (4.31 million MT), sorghum is India's fourth most significant foodgrain after rice, wheat, and maize. Bajra (7.05 million ha) contributes more than half of the country's millet acreage with a roughly equivalent share of production. It is worth noting that India is the world's leading producer of Barnyard (99.9 percent), Finger (53.3 percent), Kodo (100 percent), Little millet (100 percent), and pearl millet (44.5 percent), producing approximately 12.46 million metric tonnes from an area of 8.87 million ha.

The Autoregressive Integrated Moving Average (ARIMA) forecasting model (Kumar 1990, Hossain *et al*. 2006, Koutroumanidis *et al*. 2009), which is applied in this paper is one of the most popular and commonly used forecasting models for univariate time series data. Although it is used in a variety of functional areas, its application in agriculture is severely limited, owing to a lack of essential data as well as the fact that agricultural products are often dependent on monsoon and other elements that the model failed to incorporate. Few applications of the ARIMA model for forecasting agricultural products are worth highlighting in this area. Masuda and Goldsmith (2009) predicted global soybean production; Coorey (2006) used monthly data from January 1988 to September 2004 to project Sri Lanka's monthly total production of tea and paddy beyond September 1988. The purpose of this study is to develop an ARIMA model for forecasting millets production in India.

MATERIAL AND METHODS

The annual data were obtained from 1966-67 to 2023-24 from the Directorate of Economics and Statistics, Ministry of Agriculture and Farmers' Welfare, Government of India. The data was fitted using the ARIMA model to protect future millet production in India.

Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model analyses and forecasts time series data with equal spacing. A value in a response time series is predicted by an ARIMA model as a linear mixture of its prior values. Box and Jenkins (1976) popularized the ARIMA technique, and ARIMA models are frequently referred to as Box-Jenkins models.

This procedure is known as ARIMA (p, d, q), where p and q are the numbers of AR and MA terms, respectively, and d is the order of differencing required to make the series stable.

The moving average (MA) model, which shows a relationship between a value in the present (Zt) and residuals in the past with a non-stationary data pattern and d differencing order, and the autoregressive (AR) model, which shows a relationship between a value in the present (Z) and value in the past (Zt-k), are combined to form the ARIMA (p, d, q) model. The form of ARIMA (p, d, q) is:

$$\Phi_p (B) (1 - B)^d Z_t = \Theta_q (B) a_t$$

Where, p is the autoregressive model,

Q is the moving average model order,

D is differencing order

$$\Phi_p (B) = (1 - \Phi_1 B - \Phi_2 B^2 - \ldots\ldots\ldots. - \Phi_p B^p)$$

$$\Phi_p (B) = (1 - \Phi_1 B - \Phi_2 B^2 - \ldots\ldots\ldots. - \Phi_q B^q)$$

Model Identification

We looked at the ACF graph to see if the series was stationary or not. If an ACF graph abruptly ends or begins to decline, the time series value should be regarded as stationary. If the original series is stationary, d = 0, and the ARIMA models become ARMA models. The differences linear operator () denoted by:

$$\Delta Y_t = Y_t - Y_{t-1} = Y_t - B Y_t = (1 - B) Y_t$$

The stationary series:

$$W_t = \Delta^d Yt = (1 - B)^d Y_t = \mu + \Theta_q (B) \varepsilon_t$$

Model Estimation

Estimate the parameters for a preliminary model that has been chosen. The generated model must be evaluated for appropriateness by considering the residual attributes, such as whether the residuals from an ARIMA model are normal or randomly distributed. Ljung-Box Q statistics provide an overall review of the model's appropriateness. The test statistics Q is given in equation below:

$$Qm = n(n+2) \sum_{k- \ n-k} r2k(e)$$

Where, rk(e) = the residual autocorrelation at lag k

n = the number of residuals

m = the number of time-lags includes in the test

If the p-value associated with the Q statistics is low, the model is deemed insignificant. The analyst should evaluate a new or modified model and proceed with the study until a good model is found.

The Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Maximum Absolute Percentage Error (MaxAPE), and Maximum Absolute Error (MaxAE) are used to determine the accuracy of the fitted models.

RESULTS AND DISCUSSION

Identification

The Box-Jenkins ARIMA model methodology identifies the problem of deciding appropriate values of p, d, and q partially by following the steps described earlier. The initial step for fitting the ARIMA model starts with a stationarity test. The model made estimates for millet production after turning the variable being forecasted into a stationary series. An individual whose value only fluctuates around a constant mean and constant variance over time is said to have a stationary series. There are multiple methods for determining this. The most powerful method for determining stationarity is to look for non-stationary data on the chart or temporal plot of the information. Since the Xt graph in this instance was not stationary in the mean, there was a need to separate the data.

Model Identification

Model identification entails determining the order of the AR, MA parameters and differentiation value of the model. First, we determine if the series is stationary or not. In Figure 1, we found that the series of millet production is not stationary. If the series is not stationary, we must switch over for differences in the series. Figure 2 shows that the time series of second differences looks to be stationary in mean and variance, as the series level remains about constant over time and its variance remains roughly constant over time.

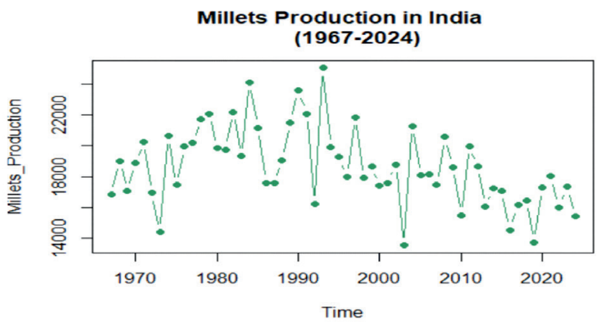

Figure 1: Original series for Millet Production

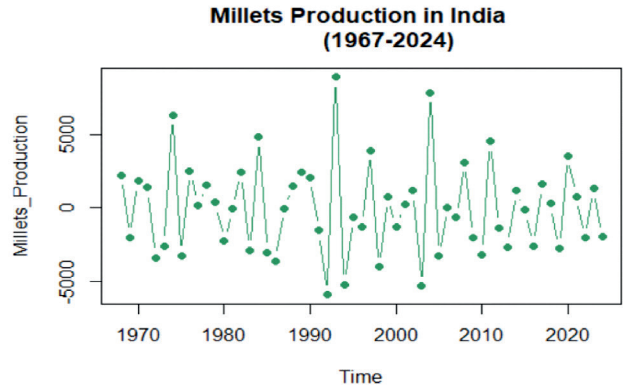

Figure 2: Differenced series for Millet Production

The stationarity of the time series is detected using the time plot Figure 2. The data fluctuate horizontally, and the pattern of fluctuation demonstrates that the time series has a constant in mean, and the plot also identifies a non-seasonality pattern. Some of the spikes veered greatly from the center.

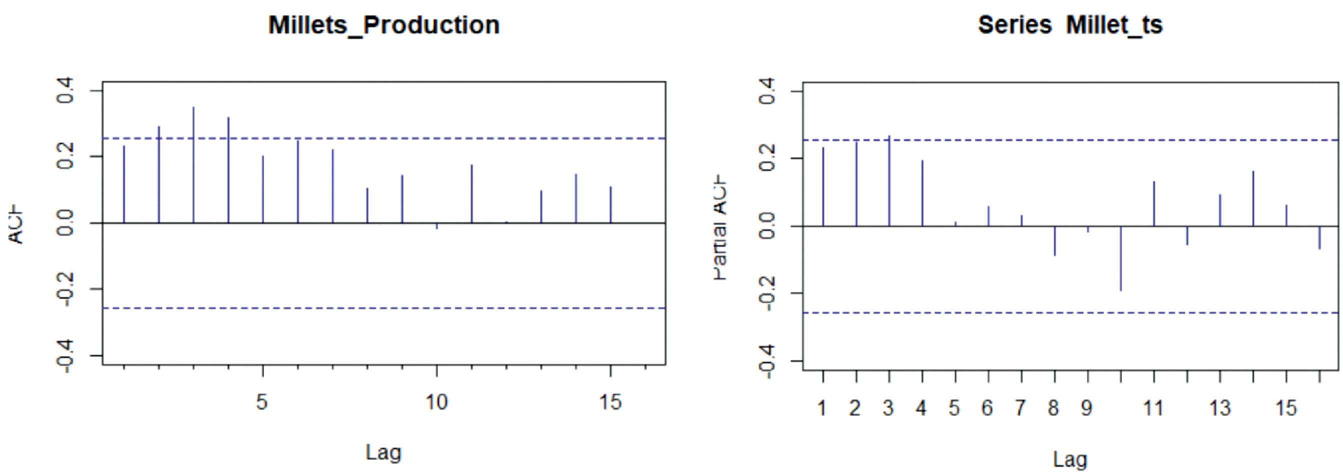

Figure 3: ACF and PACF plot fitted for original millet production

Figure 3 shows the auto-correlation function (ACF) and partial autocorrelation function (PACF) plot for lags 1 to 16 of the original series. The auto-correlation function plot for lags 1 to 16 of the first order differenced time series of India's millets production is shown in Figure 4 below.

The correlogram shows that the autocorrelation at lag 1 and lag 2 exceeds the significance thresholds, but all other autocorrelations between lags 1 to lag 16 do not. The partial correlogram shows that the partial autocorrelations at lags 1 and 2 exceed the significance bounds and are negative. Although here we have one outlier at lag 3 (coefficient at lag 3 is almost touching the significant limits), which we can assume it is an error and happened due to chance alone because all the other PACFs from lag 4 to 16 are within the significant limits. The partial autocorrelations tail off to zero after lag 4. The auto-correlation function assists in determining the best values for arranging moving average terms (MA) and partial auto-correlation functions (AR).

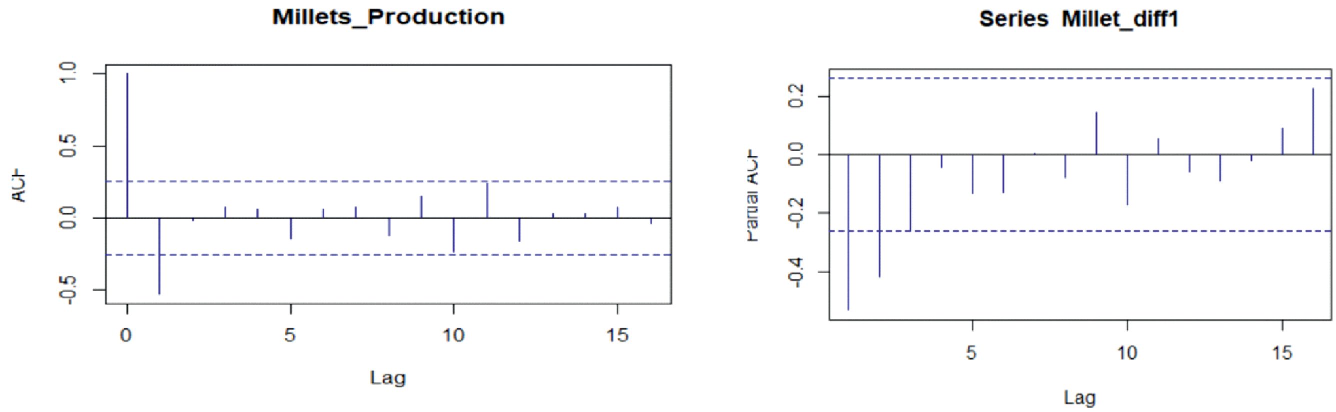

Figure 4: ACF and PACF plot fitted for first order differenced in millet production

The number of non-zero coefficients in ACF determines order of MA terms and the number of non-zero coefficients in PACF plots determines the order of AR terms. Based on the Akaike Information Coefficient (AIC) and Bayesian Information Criteria (BIC), the model ARIMA (0,1,1) was the best fitted model for millets production in India. The results of these coeficients are given in Table 1. ARIMA model was estimated after transforming the variables under study into stationary series through computation of either seasonal or non-seasonal or both, order of differencing. A careful examination of ACF and PACF up to 16 lags revealed the presence of non-seasonality in the data. However, the series was found to be stationary, since the coefficient dropped to zero after the fourth lag.

Table 1: Residual Analysis of the estimated model of the Millets Production in India

Model Fit Statistics	*Values*
AIC (Akaike Information Criterion)	1048.2
BIC (Bayesian Information Criterion)	1052.29
ME (Margin of Error)	-99.3495
RMSE (Root Mean Square Error)	2260.21
MAE (Mean Absolute Error)	1831.24
MPE (Mean Percent Error)	-1.9222
MAPE (Mean Absolute Percentage Error)	10.0282
MASE (Mean Absolute Scaled Error)	0.7508

Model Verification

Again, we should investigate whether the forecast errors seem to be correlated, and whether they are normally distributed with mean zero and constant variance. To check for correlations between successive forecast errors, we can make a correlogram and use the Ljung-Box test.

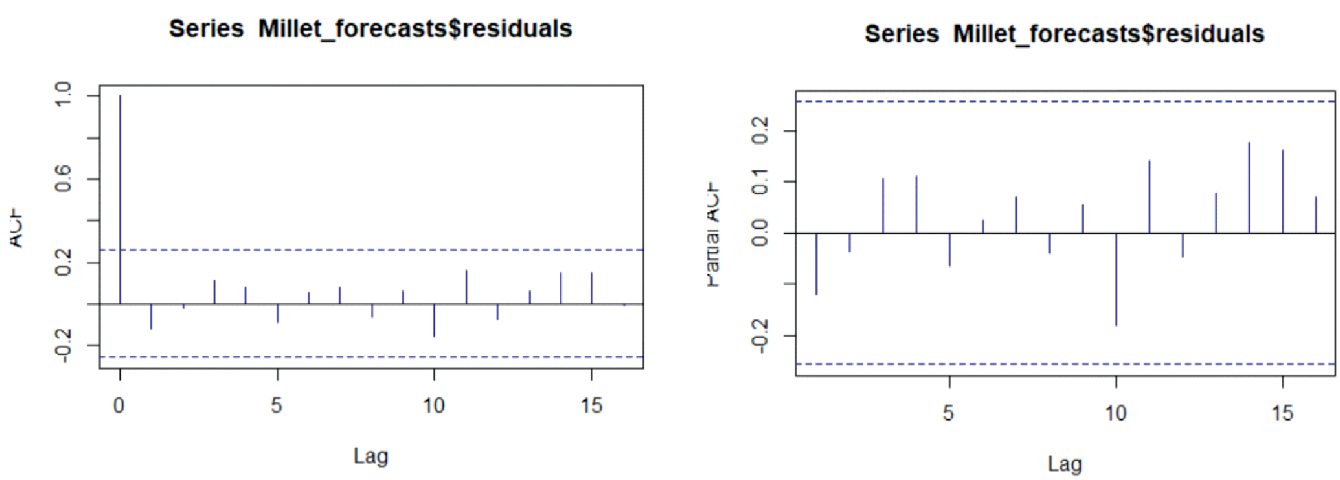

Figure 5: ACF and PACF plot of the residuals in millet production

The correlogram demonstrates that none of the sample autocorrelations for lag 1-16 exceed the significance thresholds, and the p-value for the Ljung–Box test is 0.9. We may conclude that there is very little evidence for non-zero autocorrelations in the forecast errors at lag 1-16. A time plot of the forecast errors and a histogram are created to determine whether the forecast errors are normally distributed with a mean zero and constant variance.

The forecast error time plot reveals that the forecast error appears to have a nearly constant variance across time. However, the forecast error time series appears to have a negative mean rather than a zero mean. We can validate this by computing the mean forecast error, which is around -0.22. The histogram of the forecast error illustrates that the variation of the forecast errors appears to be relatively consistent throughout time (albeit there may be slightly larger variance in the early half of the time series). The time series histogram demonstrates that the forecast errors are generally normally distributed and that the mean appears to be close to zero. As a result, forecast errors are likely to be regularly distributed with a mean of zero and a constant variance. Because successive forecast errors do not appear to be connected and forecast errors appears to be normally distributed with mean zero and constant variance, the ARIMA (0, 1, 1) model appears to be an adequate predictive model for millets production in India.

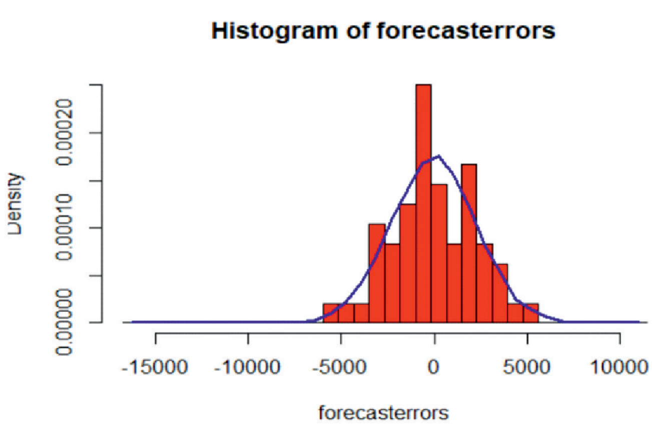

Figure 6: Histogram of the Forecasted Error

Forecasting the Millets Production

Once the three preceding steps of the ARIMA model are completed, we can obtain forecasted values by estimating the appropriate problem free model. Figure 7 indicates that the observed and forecasted millets production in India. The forecasted values were calculated from 2024-25 to 2029-30. The forecasts for millet production showed a linear trend. Figure 8 shows the forecasted millet production in India. The production of millets in India, for the forecasted year as obtained from ARIMA (0, 1, 1) was obtained as 16442 thousand tonnes with upper and lower confidence of 21397 thousand tonnes and 11487 thousand tonnes. The upper and lower confidence limits indicate that the millet production will not fall below the lower confidence limit or rise above the upper confidence limit in the next.

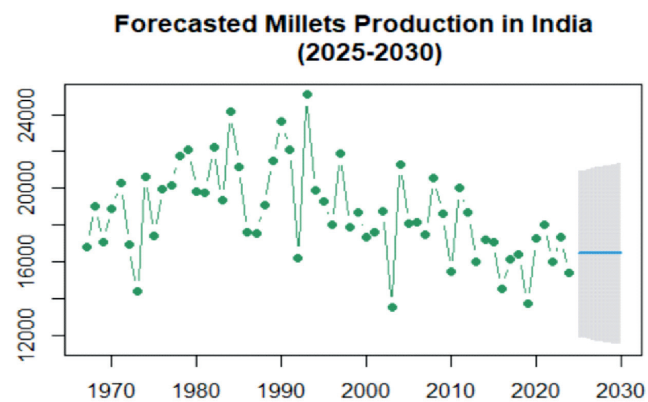

Figure 7: Observed and Forecasted Millet Production in India

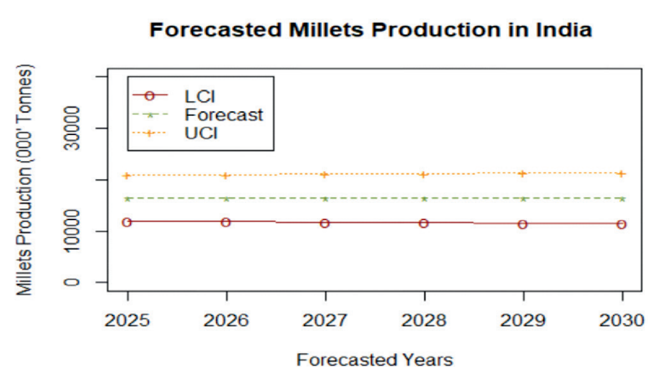

Figure 8: Forecasted Millet Production in India using ARIMA (0, 1, 1)

CONCLUSION

This paper aimed to forecast the millet production during 2025-25 to 2029-30 in India, by Autoregressive Integrated Moving Average (ARIMA) approach. On basis of results obtained it is concluded that ARIMA (0, 1, 1) model having minimum value of all measures of selection criteria was found to be the appropriate model amongst all for predicting the millets production in India. The model showed a good performance in case of explaining variability in the data series and also its predicting ability. Forecasting data points are used to offer a foundation for production control and planning, as well as to optimize industrial operations and economic planning.

REFERENCES

Box, G. E. P., & Jenkins, G. M. (1976). Time series Analysis, Forecasting and Control. San Francisco, Holden – Day, California, USA.

Cooray, T. M. J. A. (2006). Statistical analysis and forecasting of main agriculture output of Sri Lanka: rule-based approach. Appeared in 10th International Symposium, Sabaragamuwa University of Sri Lanka 221: 1-9.

Hossain, M. Z., Samad, Q. A., and Ali, M. Z. (2006) ARIMA model and forecasting with three types of pulse prices in Bangladesh: a case study. *International Journal of Social Economics* 33(4): 344-53.

Koutroumanidis T, Ioannou K, and Arabatzis G. (2009). Predicting fuelwood prices in Greece with the use of ARIMA models, artificial neural networks and hybrid ARIMA – ANN model. Energy Policy 37:3627-34.

Kumar Kuldeep. (1990). Some recent developments in time series analysis. *Singapore Journal of Statistics* 1:45-73.

Masuda, T. and Goldsmith, P. D. (2009). World soybean production: area harvested, yield, and long – term projections. *International Food Agribusiness Management Review* 12: 143-162