Forecasting Pistachio Production in Turkey: A Comparison of ARIMA, Grey, and Exponential Smoothing Models

Görkem ÖRÜK

Department of Agricultural Economics, Faculty of Agriculture, Siirt University, P.O. Box: 56100, Siirt, Turkey

Author for correspondence; Email: gorkem.ozturk@siirt.edu.tr ORCID No: 0000-0003-3767-0537

Received: March 03, 2022/ Revised: May 10, 2022/ Accepted: May 17, 2022

This research aims to forecast pistachio production in Turkey using the Auto-Regressive Integrated Moving Average (ARIMA), the grey model (G.M. (1,1)), and Exponential Smoothing (Holt's and Brown's) models. The Food and Agriculture Organization provided the annual statistics on pistachio production in Turkey from 1961 to 2019, which was used in this study. Pistachio production is expected to increase between 2020 and 2030. The mean absolute error (MAE), root mean square error (RMSE), and the mean absolute percentage error (MAPE) statistics were used to assess the performance of the models. The G.M. (1,1) model was found to be significantly better than other models, and it was used to make future predictions. Results showed that the forecasted production value for 2020 was 157.80 thousand tons, rising to 302.86 thousand tons in 2030. Pistachio production is expected to increase during the next ten years.

Keywords: pistachio, ARIMA, grey forecasting, exponential smoothing

Abbreviations: ACF—autocorrelation function, ARIMA—auto-regressive integrated moving average, MAE—mean absolute error, RMSE—root mean square error, MAPE—mean absolute percentage error, PACF—partial autocorrelation function

INTRODUCTION

The agricultural sector is important in Turkey, but its proportion has been diminishing in recent decades due to the growth of the industrial and services sectors (Tipi et al. 2009; Baran 2017). Turkey provides favorable environmental conditions for pistachio farming. Pistachio production dates back to ancient times, and Turkey is one of the primary gene centers for this crop (Satil 2003). The Food and Agriculture Organization of the United Nation (FAO) estimates that about 1,125,305 tons of pistachios are produced globally on 8.30 thousand hectares (FAO 2020). The countries having with the most pistachios area largest areas of pistachio production are Turkey (45.96%), Iran (19.61%), the United States of America (18.12%), and Syria (7.27%). On the other hand, the United States of America is the first in terms of production (42.12%), followed by Turkey (26.34%), Iran (16.88%), China (7.13%), and Syria (6.17%). Turkey’s total pistachio production was 85,000 tons in 2019 (FAO 2020). According to FAO data, Turkey ranks eighth in world pistachio exports with a share of 3.61% (FAO 2022). While pistachios grown in Turkey are smaller and more costly, they are disadvantageous in the international markets. However, their more intense flavor provides an advantage. 247,947 tons of pistachio were consumed in 2020 in Turkey and pistachio consumption per person is 3 kg per year.

However, in the face of supply-demand and price developments, forecasting is essential in agricultural products, especially for policy-making (Berk and Ucum 2019; Caner and Engindeniz 2020). It is also crucial for advanced planning, formulation, and implementation of food procurement, distribution, and import-export decision (Ullah et al. 2018).

The ARIMA model has been widely used for predicting the production of some agricultural products (Khan et al. 2008; Koc and Tonkaz 2010; Celik 2013; Biswas et al. 2014; Ucar et al. 2016; Ucum 2016; Celik et al. 2017; Ucar et al. 2016; Darekar and Reddy 2017; Celik et al. 2017; Kaygisiz and Sezgin 2017; Bars et al. 2018; Ullah et al. 2018; Berk and Ucum 2019; Paidipati and Banik
The grey estimation model is another method for predicting the production of agricultural products. Zhang et al. (2007) and Li et al. (2017) used GM (1,1) model to estimate maize yield. Xiu (2011) examined forecasting grain production based on the grey system model. Ghali et al. (2018) conducted a study to forecast the total production and harvested area of paddy. Li and Zhu (2018) have done predictions about the agriculture production price using the grey model. Ghalih et al. (2018) conducted a study to forecast the total production and harvested area of paddy. The study aims to forecast the production of pistachio in Turkey and to find an accurate forecasting models that are able to estimate the best result for pistachio production.

MATERIAL AND METHODS

The survey material included information on pistachio production from 1961 and 2019. The statistics used in the study were obtained from FAOSTAT (FAO 2019).

In the current study, the Autoregressive Integrated Moving Average (ARIMA) model, the G.M. (1,1) model, and two exponentials smoothing (Holt’s and Brown’s with a linear trend) models were used to forecast pistachio production. Data from 1961 to 2019 were used for the ARIMA model, while the data from 2010 to 2019 were used for the G.M. (1,1) model. B, since the Grey model is only suitable for estimating data from small samples (Liu et al. 2021). Pistachio production was forecast for the years 2020 to 2030.

The Grey model theory was first developed in 1989 by Julong Deng to provide theories, concepts, ideas, and techniques for analyzing latent and complex systems. In the Grey model theory, a dynamic model has been developed that includes a group of differential equations called the grey differential model (GM) (Deng 1989). Most of the previous researchers preferred the GM (1,1) model for their predictions because of its computational efficiency (Yang et al. 2018).

The GM (1,1) model is presented as below (Deng 1989; Li and Zhu 2018; Bayrakci and Aksoy 2019).

\[
X^{(0)} \text{ is the original non-negative data series, and the raw data series are created as below: } \\
X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\} \quad (1)
\]

Where \( n \) is the number of the data.

The sequence of \( X^{(1)} \) is created by making an accumulation of:

\[
X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\} \quad (2)
\]

where \( X^{(1)}(k) = \sum_{i=1}^{k} X^{(0)}(i), k=1,2,\ldots,n \).

By creating the \( X^{(1)} \) series, equation \( x^{(0)}(k) + ax^{(1)}(k) = b \) is obtained, and it is called the original G.M. (1,1) model.

\[
Z^{(1)}(k) = \frac{1}{2} \left[ x^{(1)}(k)+x^{(1)}(k-1) \right], \quad k = 2,3,\ldots,n \quad (4)
\]

The basic format of the G.M. (1,1) model is obtained with the \( Z^{(1)} \) series.

The equation is as follows:

\[
x^{(0)}(k) + az^{(1)}(k) = b \quad (5)
\]

The whitening equation of \( x^{(0)}(k) + az^{(1)}(k) = b \):

\[
\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (6)
\]

\( \hat{a} = (a, b)^T \) series parameters are found as below:

\[
\hat{a} = (B^T B)^{-1} B^T Y
\]

Y, B, of them, are:

\[
Y = \begin{bmatrix}
x^{(0)}(2) \\
x^{(0)}(3) \\
\vdots \\
x^{(0)}(n)
\end{bmatrix}, \quad B = \begin{bmatrix}
\frac{-x^{(1)}(2)}{2} & 1 \\
\frac{-x^{(1)}(3)}{2} & 1 \\
\vdots & \vdots \\
\frac{-x^{(1)}(n)}{2} & 1
\end{bmatrix} \quad (7)
\]

Based on the estimated coefficients \( a \) and \( b \), the grey estimation equation is calculated by solving the equation (6):

\[
x^{(1)}(k+1) = \left( x^{(1)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a} \quad (8)
\]

After the inverse cumulative operation of the whitening equation of the model, the predicted values are obtained with the model shown as below:

\[
x^{(0)}(k+1) = a^x x^{(1)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k) = (1 - e^a) \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} \quad (9)
\]

The most commonly used prediction model is the Autoregressive Integrated Moving Average (ARIMA) model, also referred to as the Box-Jenkins model. The
ARIMA model is defined as an ARIMA (p, d, q). Here, d is the degree of differencing, p is the autoregressive order, and q is the order of the moving average process (Saeed et al. 2000).

The development of an ARIMA model is based on the following steps: (1) Defining the Model, (2) Parameter Estimation and Selection, (3) Model Verification or Diagnostic Checking, and (4) Use of the Model (Nath et al. 2019). The main assumptions of ARIMA models are serial dependency, normality, stationarity, and invertibility (Amin et al. 2012).

Exponential smoothing is a procedure to continually revise a forecast (Touama 2014). Simple exponential smoothing has been expanded to enable data prediction with a trend by Holt (1957). The method includes one forecast equation and two smoothing equations (Deppa 2018). The Brown’s model is specific to the Holt’s model. The forecasting formula is based on extrapolating a line through the two centers (Roy et al. 2018). Brown’s model is suitable for linear trend series without seasonality (Brown and Meyer 1961).

The ARIMA method’s first step is to determine whether the time series is stationary or not. ARIMA methods have been applied to stationary series. First, a correlogram test was performed. Figure 2 depicts the autocorrelation function (ACF) and partial autocorrelation function (PACF) for pistachio production. The ACF and PACF results show that data is not stationary, so a differencing process is required.

The ACF and PACF graphs of the first difference series are shown in Figure 3. The series of pistachio production is stationary at the first difference, according to the results of the ARIMA forecasts.

The results of the Augmented Dickey-Fuller test are shown in Table 2. The data is also not steady, according to the Augmented Dickey-Fuller (ADF) test ($p < 0.05$). As a result, the first difference was used to attain stability ($p < 0.05$).

**RESULTS AND DISCUSSION**

**ARIMA Model Results**

Descriptive statistics of pistachio production are presented in Table 1. In 2018, the maximum production of pistachio was 240,000 thousand tons.

The trend in the data is exhibited by the nonstationary shape of the time series as is seen in Figure 1.

Fig. 1. Trends of pistachio production.

Fig. 2. Correlogram for pistachio production.
For various values of p and q (0 to 3), ARIMA models for pistachio production were fitted and d was chosen as one based on the ADF test. The optimal model for pistachio production was chosen based on the least Akaike information criterion (AIC), Hannan-Quinn criterion (H.Q.), and Schwarz-Bayesian Information Criteria (SBC). ARIMA (1,1,1) was chosen as the optimum model for pistachio production (Table 3).

Table 4 contains information on the ARIMA (1,1,1) model as well as its outcomes. In the previous studies, similar findings have been reported. The ARIMA (1,1,1) model has been proven to be appropriate for sugarcane area and productivity (Suresh and Priya 2011); wheat area (Prabakaran et al. 2013); chestnut production (Baser et al. 2018); and cherry production (Akin et al. 2021). However, Celik (2013) found that the ARIMA (2,1,0) model was the best for projecting pistachio production in Turkey between 1936 and 2011.

The Kolmogorov-Smirnov, Jarque-Bera, and Doornik-Hansen tests were used for the diagnostic checking of normality. According to the data, residuals are normally distributed (p > 0.05) (Table 5).

Comparison and Evaluation of Multiple Models

ARIMA, G.M. (1,1), Exponential Smoothing (Holt and Brown) models were employed for estimating. The findings of the mean absolute error (MAE), root mean square error (RMSE), and the mean absolute percentage error (MAPE) approaches are shown in Table 6. The results show that G.M. (1,1) outperforms the other models. Grey forecasting models can be used to forecast other agricultural commodities. Quartey-Papafio et al. (2021) discovered that grey models fell below 10% of the MAPE value, and forecasting performance is superior to the ARIMA for cocoa production.

The forecast of pistachio production from 2020 to 2029 is shown in Table 7. For the year 2020, the estimated value of production was 157.80 thousand tons. The total pistachio production forecast for 2030 is 302.86 thousand tons. According to the forecast, pistachio production will increase in the coming years. There will be an increasing amount of pistachios in the country for domestic consumption and export in the future. The increase in production will contribute to the improvement of farmers’ welfare. The results of this study can provide a reference for policy makers on the development of pistachio production policy strategies. Moreover, it is also important for ensuring the supply-demand balance and preventing price fluctuations in the pistachio market.
CONCLUSION

Pistachio farming in Turkey adds added value by ensuring the use of idle lands that are unlikely to have economic value, reducing unemployment, and preventing desertification by controlling erosion in the barren areas (Anonymous 2017). This research was carried out with the goal of forecasting pistachio production in Turkey. The models compared in this study are the ARIMA, G.M. (1,1), and Exponential Smoothing (Holt and Brown) models. The RMSE, MAE, and MAPE average of GM (1,1) is smaller than the other models and indicates that the GM (1,1) model has better forecasting performance. Therefore, the G.M. (1,1) model was selected as the best model for the time series data and was used to forecast production up to the year 2030.

According to projections, pistachio production will rise over the next ten years. Pistachio production is expected to reach 157,800.76 thousand tons in 2020 and 302,862.02 thousand tons in 2030. The outcomes of this projection will assist Turkey’s agricultural planning in the future. The conclusions of this study can be used as a reference for future forecasting research. Furthermore, this research has demonstrated that the G.M. (1,1) model can be utilized to anticipate future agricultural product values.

The following measures are suggested to increase the pistachio production in Turkey: (1) Pistachio production can be increased by planting high-quality improved cultivars. Selection and breeding studies should be initiated to improve cultivars; (2) The juvenile period of pistachio trees is quite long compared to other fruit types. This long juvenile period can be shortened by providing awareness to farmers about irrigation and cultural practices; (3) One of the problems in pistachio production is insufficient pollination due to the low number of male trees in the orchards. This situation increases the blank nut rate. Artificial pollination method should be applied in existing gardens where there is no pollinator variety, and one male tree should be planted for every 10 female trees for new gardens (Erturk et al. 2015); (4) Provincial/District Directorates of the Ministry of Agriculture and Forestry should inform farmers about harvesting and also should establish inspection and control units in order to prevent incorrect harvesting practices; (5) Growers are facing serious problems such as a weak market system and sales uncertainty resulting in price fluctuations due to excess supply. Grower unions and cooperatives are important in solving these problems (Ucar et al. 2016).

REFERENCES CITED


ROY A, DUTTA G, DAS PK. 2018. Identifying the most effective model for understanding the growth rate of


