Time Series Forecasting of Honey Production in Turkey

Burcu Erdal¹, Tolga Tipi ²

¹Bursa Uludağ University, Faculty of Agriculture, Department of Agricultural Economics, Gorukle, Bursa, Turkey (ORCID: 0000-0002-6839-913X), berdal@uludag.edu.tr
²Bursa Uludağ University, Faculty of Agriculture, Department of Agricultural Economics, Gorukle, Bursa, Turkey (ORCID: 0000-0002-1090-3639), ttipi@uludag.edu.tr

(First received 3 February 2022 and in final form 29 April 2022)

DOI: 10.31590/ejosat.1066665

Abstract

Beekeeping is an essential agricultural activity in socio-economic terms, as it can be done with a small capital, it is an activity where the family workforce can be used well, it can be done without being dependent on the soil, and it is a production branch that increases the income and living standards of the rural population. Considering Turkey’s economic conditions in rural areas as well as its ecological richness, sustainable, organized, and conscious beekeeping is seen as an agricultural activity that must be expanded and developed. In this context, considering the data of the years 1966-2019 covering 54 years of honey production, forecasting the future is quite essential in developing policies in this field. This study aims to forecast honey production in Turkey between the years 2020-2029 using the Autoregressive Integrated Moving Average (ARIMA) model. In the forecasting of future honey production, 54 years of honey production data covering the years 1966-2019 and the Autoregressive Integrated Moving Average (ARIMA) model were used.

Keywords: Honey, Forecast, Time series, Beekeeping, Turkey.

Godement: Bal Üretiminin Zaman Serileri Tahmini

Oz


Anahtar Kelimeler: Bal, Tahmin, Zaman serileri, Arıçılık, Türkiye

* Corresponding Author: berdal@uludag.edu.tr

http://dergipark.gov.tr/ejosat 417
1. Introduction

In addition to producing products such as honey, beeswax, royal jelly, bee venom, pollen, and propolis, which are extremely valuable in terms of human health and nutrition, beekeeping has vital importance in natural life and plant production with the pollination. The facts that beekeeping can be done with a small capital, it is an activity where the family workforce can be used well, it can be done without being dependent on the land, and it is a production branch that increases the income and living standards of rural residents both add importance to beekeeping socio-economically and ensure that it takes an integral part in rural development programs in developing countries. Also, beekeeping is an important agricultural activity that should be supported, given its importance in environmental sustainability through pollination.

Beekeeping mainly stands out as a small family business. It is also carried out as a second source of income alongside other agricultural activities. Therefore, beekeeping, which has an important contribution to the national economy and the environment, is carried out extensively in most countries of the world.

Due to the facts that beekeeping is an additional source of income for families who still live in the countryside, is an easy and enjoyable pursuit, honey, which is derived from bees, is used as a valuable nutrient and health resource, can be stored for a long time without spoiling compared to other agricultural products, and can be marketed easily have caused an increased interest in beekeeping in the world and Turkey thus increasing the number of colonies and honey production. The importance of honey and byproducts obtained from beekeeping such as royal jelly, propolis, and pollen in terms of health has become more apparent, and applications of “apitherapy,” which is used in the meaning of treatment with bee products, have become more popular.

Generally, with different climatic and natural conditions, numerous beekeepers with millions of bees, land structure, very rich vegetation, and genetic diversity in honey bee populations, Turkey is a country with considerable potential in beekeeping. Apart from natural bee pastures in Turkey, agricultural lands consist of clover, soya bean, oilseed crops such as sunflower, fruit trees such as apple, citrus, and almonds further increase the chances of Turkey in beekeeping [1].

Although Turkey is in a crucial position among the world countries on the production of honey, beekeeping byproducts production is at a reasonably low level, and there is not enough demand for the production of byproducts.

Considering Turkey’s economic conditions in rural areas as well as its ecological richness, sustainable, organized, and conscious beekeeping is seen as an agricultural activity that must be expanded and developed.

The bee products produced in Turkey are good in terms of quality and health, and other goals should be to reduce costs by increasing yield per colony and thus improve competitive power.

Environmental effects of climate change, input costs, insufficient supports, foreign trade policies, and fluctuations in domestic market prices can adversely affect honey production. Foreseeing the future is extremely important in terms of honey production, especially against the supply-demand and price developments experienced today. Knowing the amount of production will contribute to agricultural planning and development of agricultural policies in line with the realities of the country and development goals. For this reason, there is a need for countrywide studies to estimate the production of agricultural products. There are many studies that estimate the production amount, planting area, price and export of agricultural products using the ARIMA model [2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19].

In case policies and strategies are developed integrating the production and marketing dimensions of beekeeping in Turkey, and they are implemented effectively, Turkey may be able to benefit from its current beekeeping potential. Research on the future of honey production in Turkey will be important in terms of directing the production and determining the measures to be taken to ensure sustainable production. Although studies based on honey production estimation were done before, updating these studies is essential in making new forecasts and seeing the implemented policies’ results.

This study aims to determine Turkey's current status in world honey production and estimate the amount of honey production for the next ten years using production quantities in the period 1966-2019. In this study, the Autoregressive Integrated Moving Average (ARIMA) modeling approach has been adapted to forecast honey production in Turkey. The time series modeling is a valuable tool for production planning and decision-making of support policy by the government in agriculture.

1.1. The Importance of Beekeeping in Turkey

Today, beekeeping is one of the most common agricultural activities in the world. As of 2019, 1,852,598 tons of honey has been produced from 90 million bee hives in the world (Table 1). Although India has the highest number of bee hives, the leading country in honey production is China, which makes 24% of the total production. Total honey production after China is in Turkey, with a share of about 6% [20].

In 2019, the world average honey production per hive was 20.6 kg. The quantity of production per bee hive in Turkey was relatively low at 13.5 kg. China, the world’s largest honey producer, was the country with the highest honey production per hive with 48.4 kg.

The best-known bee product after the honey is beeswax. In the production of beeswax, which is used in beekeeping as well as the metal industry, cosmetics industry, textile industry, medicine making, candle making, and many more areas, Turkey’s position is within the first five countries. Despite having the greatest number of bee hives, India, which has a share of 5.5% in honey production, produces approximately 39% of the total world beeswax production. In 2019, a total of 66,099 tons of beeswax was produced in the world [20].

In 2019, 675,845 tons of honey was imported for 2,029.31 million dollars in the world, whereas 628,202 tons of honey was exported for 1,964.69 million dollars. The countries that import the greatest amount of honey in the world are the USA, Germany, the United Kingdom, Japan, and France. The countries that export the greatest amount of honey are China, Argentina, India, and Ukraine, which also have a say in the production [21].

Turkey, which ranks second in terms of production, could not have a sufficient market in the world honey exports. Turkey
exported only 5,543 tons of honey in 2019 and generated an income of 24,773 thousand dollars [21]. The main reason for this state is that increased honey production costs due to low productivity reduced Turkey’s chance of competing in the international market.

Table 1. Number of hives and honey production by country (2019)

<table>
<thead>
<tr>
<th>Country</th>
<th>Beehives (Number)</th>
<th>Honey Production (tons)</th>
<th>% of World Honey Production</th>
<th>Yield (kg/beehive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>9,230,940</td>
<td>447,007</td>
<td>24.13</td>
<td>48.4</td>
</tr>
<tr>
<td>Turkey</td>
<td>8,128,360</td>
<td>109,330</td>
<td>5.90</td>
<td>13.5</td>
</tr>
<tr>
<td>Argentina</td>
<td>2,985,026</td>
<td>78,927</td>
<td>4.26</td>
<td>26.4</td>
</tr>
<tr>
<td>Iran</td>
<td>7,516,720</td>
<td>75,463</td>
<td>4.07</td>
<td>10.0</td>
</tr>
<tr>
<td>USA</td>
<td>2,812,000</td>
<td>71,179</td>
<td>3.84</td>
<td>25.3</td>
</tr>
<tr>
<td>Ukraine</td>
<td>69,937</td>
<td>69,937</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>India</td>
<td>12,247,332</td>
<td>67,141</td>
<td>3.62</td>
<td>5.5</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>3,093,859</td>
<td>63,526</td>
<td>3.43</td>
<td>20.5</td>
</tr>
<tr>
<td>Mexico</td>
<td>2,157,866</td>
<td>61,986</td>
<td>3.35</td>
<td>28.7</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>6,220,182</td>
<td>53,782</td>
<td>2.90</td>
<td>8.6</td>
</tr>
<tr>
<td>Tanzania</td>
<td>2,984,290</td>
<td>30,937</td>
<td>1.67</td>
<td>10.4</td>
</tr>
<tr>
<td>Republic of Korea</td>
<td>2,141,015</td>
<td>29,518</td>
<td>1.59</td>
<td>13.8</td>
</tr>
<tr>
<td>Romania</td>
<td>1,689,500</td>
<td>25,269</td>
<td>1.36</td>
<td>15.0</td>
</tr>
<tr>
<td>World</td>
<td>90,116,413</td>
<td>1,852,598</td>
<td>100.00</td>
<td>20.6</td>
</tr>
</tbody>
</table>

Note: “—” means no available data.

Source: Food and Agriculture Organization (FAO) [20]

Beeckeeping in Turkey was carried out by 80,675 producers (families) as of 2019. Due to Turkey’s ecological and socio-economic structure, beeckeeping can be made in every region in Turkey. The Aegean, Black Sea, and Mediterranean regions are the most important production regions in terms of both the existence of beehives and production share. In 2019, Turkey’s total number of beehives was 8,128,360 units, honey production was 109,330 tons, while beeswax production was 3,971 tons [22]. There is a decrease in honey yield per beehive over the years (Table 2). Productivity per beehive in the world is the most important indicator for the competitiveness of the beekeeping industry.

2. Material and Method

In this research, 54 years of honey production data obtained from TURKSTAT (Turkey Statistical Institute) covering the years 1966-2019 were used.

In the time series analysis, the forecasts obtained by using the Autoregressive Integrated Moving Average (ARIMA) method were evaluated. For the analysis of time series, the SPSS 23 and Eviews 11 programs were used.

Time series analysis comprises methods for analyzing time-series data to extract meaningful statistics and other characteristics of the data and to forecast future events based on known past events to predict data points before these are measured [23].

Table 2. Number of beehives, honey and beeswax production in Turkey

<table>
<thead>
<tr>
<th>Years</th>
<th>Number of Beehives (units)</th>
<th>Honey Production (tons)</th>
<th>Honey Yield (kg/beehive)</th>
<th>Beeswax Production (tons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>3,428,442</td>
<td>54,655</td>
<td>15.9</td>
<td>2,863</td>
</tr>
<tr>
<td>1995</td>
<td>3,916,038</td>
<td>68,620</td>
<td>17.5</td>
<td>3,735</td>
</tr>
<tr>
<td>2000</td>
<td>4,267,123</td>
<td>61,091</td>
<td>14.3</td>
<td>4,527</td>
</tr>
<tr>
<td>2005</td>
<td>4,590,013</td>
<td>82,336</td>
<td>17.9</td>
<td>4,178</td>
</tr>
<tr>
<td>2010</td>
<td>5,602,669</td>
<td>81,115</td>
<td>14.5</td>
<td>4,148</td>
</tr>
<tr>
<td>2014</td>
<td>7,082,732</td>
<td>103,525</td>
<td>14.6</td>
<td>4,053</td>
</tr>
<tr>
<td>2015</td>
<td>7,748,287</td>
<td>108,128</td>
<td>14.0</td>
<td>4,756</td>
</tr>
<tr>
<td>2016</td>
<td>7,900,364</td>
<td>105,727</td>
<td>13.4</td>
<td>4,440</td>
</tr>
<tr>
<td>2017</td>
<td>7,991,072</td>
<td>114,471</td>
<td>14.3</td>
<td>4,488</td>
</tr>
<tr>
<td>2018</td>
<td>8,108,424</td>
<td>107,920</td>
<td>13.3</td>
<td>3,987</td>
</tr>
<tr>
<td>2019</td>
<td>8,128,360</td>
<td>109,330</td>
<td>13.5</td>
<td>3,971</td>
</tr>
</tbody>
</table>

Source: Turkish Statistical Institute (TURKSTAT) [22]

Working on future predictions, Box and Jenkins developed the Autoregressive Integrated Moving Average (ARIMA) method in 1970 that is used in the analysis of univariate time series. ARIMA model put forward by Box and Jenkins has been one of the most used time series models for predictions for the future [24-25].

ARIMA models form an important part of the Box-Jenkins method to time series modeling. The general methodology of the Box–Jenkins approach includes the steps of (i) identification of the model, (ii) parameter estimation, and (iii) model diagnostic checking and forecast ARIMA models are homogeneous non-stationary processes. The process of selecting the most suitable (p, d, q) structure in the ARIMA model is called model identification.

The ARIMA model is usually denoted as ARIMA (p, d, q), in which p is the order of the autoregression (AR) component, d is the order of the differencing process to form a stationary times series (I), and q is the order of the moving average (MA) process. In an ARIMA model, the value of Y at time t is estimated as the equation

\[ Y_t = \mu + \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \]

where \( Y_t \) is the value at time \( t \), \( \varphi \) is the AR parameter, and \( \theta \) is the MA parameter.

The major problem in the ARIMA modeling technique is to choose the most appropriate values for p, d, and q. This problem can be partially resolved by looking at the Auto Correlation Function (ACF) and Partial Auto Correlation Functions (PACF) for the series (Iqbal et al., 2005). The order of AR(p) is identified by partial autocorrelation function (PACF) while the order of MA(q) is identified by autocorrelation function (ACF) [3].

In the estimates made by the ARIMA method, the value of the series at any time is determined by a linear equation consisting of values belonging to the previous period and error terms. It is accepted that the variance of the series used in the model is constant and its mean is zero, that is, the series is stationary.
To analyze the ARIMA model, first of all, the reasons that distort the stationarity should be eliminated and the data series should be made stationary. The process of making the time series stationary is done by taking the first and second differences of the series. Whether the series has become stationary with the process of taking differences is determined by Augmented Dickey-Fuller Test or autocorrelation (ACF) and partial autocorrelation (PACF) plots.

Autocorrelation (ACF) and partial autocorrelation (PACF) functions were taken into consideration to define the most suitable ARIMA model, and the significance of the most suitable p, d, and q parameters were checked. If the predictive power of a model is found to be insignificant, it cannot be used in forward-looking predictions.

Some statistical tests are used to determine the adequacy and predictive power of the model. The Normalized Bayesian information criterion (NBIC), the $R^2$, stationary $R^2$ Root Mean Square Error (RMSE), the Mean Absolute Percentage Error (MAPE), Mean absolute error (MAE), and the Ljung – Box Q statistic are used to test for the adequacy and statistical appropriateness of the model [26].

The stationary $R^2$ is a measure that compares the stationary part of the model to a simple mean model. This measure is preferable to ordinary $R^2$ when there is a trend or seasonal pattern. Similarly, $R^2$ estimates the proportion of the total variation in the series explained by the model [27]. This measure is most useful when the series is stationary. The RMSE shows low values indicating accuracy in model performance. Normalized BIC is a general measure of the overall fit of a model and has been widely used for model identification in time series and linear regression analysis. It is closely related to the Akaike information criterion (AIC). Lower values of the NBIC and RMSE and high values of $R^2$ were preferable. The low RMSE indicates a good fit for the model [27]. Also, the higher value of the $R^2$ indicates a perfect prediction over the mean.

The following equations were used to test the model:

\[ MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Y_t - F_t}{Y_t} \right| \times 100 \]

\[ RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (Y_t - F_t)^2} \]

\[ MAE = \frac{1}{n} \sum_{t=1}^{n} |Y_t - F_t| \]

\[ NBIC = \ln(\sigma^2) + k \ln(n) \]

\[ R^2 = 1 - \frac{\text{Error sum of square}}{\text{Total sum of square}} \]

where $n$ is the number of observations, $r_t$ is the estimated autocorrelation of the series at lag $k = 1, 2, \ldots, m$, and $m$ is the number of lags being considered.

The hypothesis of Ljung - Box test is:

$H_0$: Residual is white noise

$H_1$: Residual is not white noise

If the sample value of $Q$ exceeds the critical value of a $\chi^2$ distribution with $h$ degrees of freedom, then at least one value of $r$ is statistically different from zero at the specified significance level, and the null hypothesis is rejected. The model should have been considered inadequate. Here the degrees of freedom, $h = (m - p - q)$; $p$ and $q$ are the numbers of AR and MA terms, respectively.

3. Results and Discussion

In this study, the production amounts for the next 10-year period are estimated by using the ARIMA model with 54-year honey production series. The autoregressive integrated moving average (ARIMA) model is considered to be one of the best models when the data consists of at least 50 observations.

To apply the ARIMA model to time series data, firstly, the stationarity analysis should be done. The fluctuation of the production indicates that the dataset is non-stationary. It can also be visualized from the plot of ACF and PACF (Figure 1).

![Figure 1. ACF and PACF of the honey production data series](image)

Also, Augmented Dickey-Fuller (ADF) stationarity test was performed to test the stationarity of the series (Table 3). According to the results of the stationarity analysis test results, no stationarity was found in the data of the time series, and the series was made stationary by taking the first differences of the series (Table 3).

<table>
<thead>
<tr>
<th>Original series</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-3.133940</td>
<td>0.1101</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-4.156734</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-3.504330</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-3.181826</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>First differentiated series</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-7.925234</td>
<td>0.0000</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-4.148465</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-3.500495</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-3.179617</td>
<td></td>
</tr>
</tbody>
</table>

When only the first-order differences of the series were taken, the degree of difference was determined as $d = 1$, since the stationarity was provided.
The values of p for AR and q for MA in ARIMA estimation were determined from the ACF and PACF graphs. PACF gives p value for AR, and ACF gives q value for MA. In determining the most suitable model for the data, the ACF and PACF charts of the series's first difference were examined (Figure 2).

The PACF graph reveals that the first two lag is significant. Then it is seen that the size of the relationship decreases and approaches zero. When the ACF graph was examined, it was seen that many lags were significant (Figure 2).

Figure 2. ACF and PACF of the first differentiated honey production data series

The SPSS 23 Expert Modeler was used to fit the model to the data. To decided on the most suitable model, the significance test results of the coefficients in the models and the concordance between the prediction series and the original series were examined. According to the first difference data, the fixed term AR (2) MA (0) model was determined as the best model. Thus, ARIMA (2, 1, 0) model, including AR (2), I (1), and MA (0), was used for future honey production estimation (Table 4).

The results of the analysis show that the parameters of the ARIMA (2,1,0) model are statistically significant (p <0.05). The AR (1) coefficient \( \phi_1 \) was estimated to be \(-0.539\) with a standard error of 0.136 and a t-statistic of \(-3.963\) while the AR (2) coefficient \( \phi_2 \) was estimated to be \(-0.296\), with a standard error of 0.140 and a t-statistic of \(-2.122\). The constant was estimated to be 1887.322 with a standard error of 352.976 and a t-statistic of 5.347.

<table>
<thead>
<tr>
<th>Model</th>
<th>Ljung-box Q Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA (2,1,0)</td>
<td>18.802</td>
</tr>
</tbody>
</table>

For this model Q = 18.802. The 10% and 5% points of chi-square with 16 degrees of freedom are 23.50 and 26.30, respectively. Therefore, since Q is not unduly large and the evidence does not contradict the hypothesis of White Noise behavior in the residuals, the model is very much adequate and significantly appropriate.

One of the adequately fitted model indicators is that of scattered residuals in a rectangular shape around the zero at a horizontal level. The ARIMA model (2,1,0) is adequate because the residual ACF plots and PACF in figure 3 show a random variation. Thus, from the origin zero (0), the points below and above are all uneven. Hence the model fitted is adequate.

According to the calculations made, the MAPE value was found to be 7.945. “Mean Absolute Percentage Error (MAPE)” statistic value is used to make the prediction accuracies and to determine the future predictive power of the model [29]. Models with a MAPE value of less than 10% are classified as very good, models between 10-20% as good, models between 20-50% acceptable, and models with a MAPE value above 50% as wrong and incorrect [15].

The low value of RMSE indicates a good fit for the model. Also, the high value of the R-Squared indicates a perfect prediction over the mean. The adequacy and significant appropriateness of the model were confirmed by exploring the Normalized Bayesian Information Criterion (NBIC). The ARIMA (2,1,0) model had the least BIC value of 17.118.

The model can also be checked for adequacy by doing a chi-square test, known as the Ljung-Box Q statistic, on the autocorrelations of the residuals (Table 5).

Table 5. Diagnostic checking of residual autocorrelations of the selected ARIMA model

Observed and fit values are shown in figure 4. Fit values estimated by ARIMA (2,1,0) model and observed values show a good harmony.
Considering Turkey’s rural economic conditions and its ecological richness, sustainable, organized, and conscious beekeeping is an agricultural activity that must be supported. The social, environmental, and economic contributions of beekeeping to countries’ economies should be considered.

Honey production estimates were made for ten years, covering the years 2020-2029 with the ARIMA (2,1,0) model, which was obtained by considering the honey production data of 1966-2019. Under current conditions, it is estimated that honey production in Turkey will increase by approximately 19% in 2029 compared to 2019. Although the estimated increase in honey production is expected to meet the domestic demand, the demand for other honey byproducts will likely increase more rapidly with changing consumer preferences. Compared with the current situation in the world, Turkey’s honey production per beehive is low, and the country has almost no international trade of honey products.

Although there is government support in Turkey related to beekeeping, there is a need for new support policies regarding production quantity, yield per beehive, and the production of bee products other than honey. Encouraging the production of these byproducts and honey will increase the income of producers and by making beekeeping more attractive, it will increase the production of both these byproducts and honey. Fulfilling beekeeping products' export potential after meeting the domestic market demand is especially important for the country’s economy.

The major factor affecting honey export is reported to be its price, increasing concerns, including over food reliability and residues, currently, cause trade restrictions to be imposed on many countries in international trade. It is noted that consumers are willing to pay higher prices for quality honey [17]. Therefore, high-quality honey produced by conscious and organized producers can be exported especially to European Union countries at high prices. Beekeeping supports should be arranged in a way to take into account both the quality characteristics and the amount of production of the product, and it should encourage producer organization and branding in honey. The organization is important for the marketing, promotion, and competitiveness of honey that can be consumed safely.

5. Acknowledge

This study was not supported by any institution.

References


