International Journal of Research in Business Studies and Management
Volume 2, Issue 6, June 2015, PP 15-19
ISSN 2394-5923 (Print) & ISSN 2394-5931 (Online)

Forecasting Major Vegetable Crops Productions in Tunisia

Anis Khayati

[1]*Department of Economics and Finance, University of Bahrain*

ABSTRACT

This study aims at forecasting the short- run productivity per hectare for major vegetable crops in Tunisia. The study uses statistical analysis of forecasting based on annual data for the period 1989-2013. Two types of time-series models are used: smoothing models and stochastic models. The preferable model is selected for each crop to forecast productivity. In addition, statistical indicators are used to evaluate forecasting accuracy and model efficiency.

Results show that productivity per hectare for the crops under study tend to increase during the mentioned period. The runs test indicates that changes in that productivity are trended. Also, results indicate that the Holt model seems to be the best model to predict the productivity of potatoes, artichoke and pepper, and the ARIMA model is the best model to predict the productivity of tomatoes, while the Winters model is the best model to predict the productivity of onions. Based on the results of those models, we observe that the productivity of potatoes and tomatoes is expected to increase during the 2014-2020 period, while the productivity of pepper remains relatively stable. Meanwhile, the productivity of artichoke and onions fluctuate between an increase and a decrease. Also, results show that the forecasts of productivity related to the crops under study are efficient and unbiased.

Keywords: Forecast, smoothing models, stochastic models, Tunisia.

INTRODUCTION

Forecast is an approach that can help decision-makers, whether from the economic or non-economic fields in making their future decisions with greater accuracy. In fact, forecast is needed by national governments for the establishment of various policy decisions related to storage, distribution, pricing, marketing, import-export, etc.

The productivity of agricultural crops is generally characterized by continuous changes due to many factors such as the precipitation fluctuations and the economic, technical and agricultural conditions. The study of the nature and direction of changes in that productivity is helpful in the evaluation of the efforts made to increase agricultural production. Also, the forecast of productivity of various agricultural crops allows making accurate predictions about productivity levels during the coming years. Therefore, forecast represents one of main tools of making efficient development policies and successful economic plans in the field of agricultural production.

In this paper, a methodology for crop yield estimation is developed for major vegetable crops in Tunisia. Many previous studies were related to yield forecasting models (Hebel et al., 1993; Hall and Clutter, 2004; Lenny et al., 2006). Models developed by Mehta et al. (2000) and Agarwal et al. (2001) were also used to forecast yields of various crops, and studies done by Bazgeer et al. (2007), Andarzian et al. (2008), Esfandiary et al. (2009), Xingjie et al. (2010), Ahmad and Kathuria (2010), Mehta et al. (2010), Adrian (2012) and Verma et al. (2011) have developed and used different indicators in the context of crop yield prediction.

In the same framework, this study seeks first to determine the general trend and the annual growth rate in productivity per hectare for the main vegetable crops in Tunisia, namely potatoes, tomatoes, artichoke, pepper and onions. Those crops represent the most important vegetable crops in Tunisia and about 58% of the total cultivated area in Tunisia in 2012. The main purpose of the study is to forecast the productivity per hectare for the mentioned agricultural crops in the short run during the period 2014-2020.

Address for correspondence:

aelkhayati@uob.edu.bh

The paper is divided as follows. Section 2 presents the methodology used in this study and specifies the main models used in the empirical work. Results from empirical analysis are presented and discussed in section 3. Finally, section 4 presents the main conclusions.

METHODOLOGY

There are two main types of quantitative techniques that can be used in the forecast analysis: the causal or regression method and the time-series method. The latter is used in this study. The first method assumes that there is a causal relationship between the forecasted variable or the dependent variable and one or more independent variables. In the second method, the forecast process is based on previous observations, which means that it is related to the past in respect to variable and error values. Time-series analysis is divided into four main forecasting models, namely the deterministic models, the smoothing models, the analytical models and the stochastic models. In this study, we use three types of smoothing models, which are the Simple Exponential Smoothing Models, the Double Exponential Smoothing (Holt) Models and the Triple Exponential Smoothing (winters) Models. In addition, we use one type of stochastic models which is the Autoregressive Integrated Moving Average (ARIMA) model.

In the Simple Exponential Smoothing method, let the time-series data be denoted by Y_1, Y_2, ……, Y_t. Suppose we wish to forecast the next value of our time series Y_{t+1} that is yet to be observed with forecast for Y_t denoted by F_t. Then the forecast F_{t+1} is based on weighting the most recent observation Y_t with a weight value α and weighting the most recent forecast F_t with a weight of (1- α), where α is a smoothing constant between 0 and 1. Therefore, the forecast for the period t+1 is giving by:

$$F_{t+1} = F_t + \alpha\,(Y_t - F_t) \qquad (1)$$

The Double Exponential Smoothing method uses simple exponential smoothing in order to forecast. The forecast is obtained as a weighted average of past observed values where the weights decline exponentially so that the values of recent observations contribute to the forecast more than the values of earlier observations. Most time-series have three components: trend, seasonal and irregular. The irregular component is the residual after trend and seasonality have been removed. The Holt method accounts for only the trend and irregular components. The Winters method builds on this by allowing for seasonality.

Autoregressive Integrated Moving Average models, or ARIMA models, are a class of models that can be used to produce forecasts. The general ARIMA models have three parts: the auto-regression part (AR), the integration part (I) and the moving average part (MA). The main assumption surrounding the AR part of a time-series is that the observed value depends on some linear combination of previous observed values up to a defined maximum lag (denoted p), plus a random error term ε_t. The main assumption surrounding the MA part of a time-series is that the observed value is a random error term plus some linear combination of previous random error terms up to a defined maximum lag (denoted q). To make the analysis, we require that all of the observations are independently identifiable. There should be no autocorrelation in the series and the series should have zero mean. This process of differencing is known as integration and the order of differencing is denoted d.

The general seasonal model is denoted ARIMA(pdq)(PDQ), where p, d and q refer respectively to the orders of the non seasonal AR, I and MA parts of the model and P, D and Q refer respectively to the orders of the seasonal AR, I and MA parts of the model.

It is also important to evaluate and test the results obtained from the forecasting in order to identify the degree of efficiency of the models, and in order to choose the best model among them. There are different indicators that can be used in the evaluation and judgment of the degree of efficiency and accuracy of the model. The main indicators are the Mean Absolute Error (MAE), the Mean Absolute Deviation (MAD), the Root Mean Squared Error (RMSE), the Mean Absolute Percentage Error (MAPE), the Thiel test (U) and the correlation coefficient between actual and forecasted values.

This study used time-series data for the period 1989-2013 obtained from the Statistical Yearbook of the Ministry of Regional Development and Planning. The study is based on econometric and statistical analysis. The annual growth rates of productivity of the crops under study were estimated based on the exponential model. A Runs Test was conducted in order to know whether the changes or fluctuations in the time-series of the crops under study are trended or random.

RESULTS

Table 1 shows a summary of some statistical measures and indicators of the time-series data related to the productivity of the agricultural crops under study which are: potatoes, tomatoes, artichoke, pepper and onions during the period 1989-2013. We notice that the average productivity per hectare for the crops under study has been respectively around 13.57; 31.11; 7.23; 11.32 and 18.45 tons per hectare. This productivity has tended to increase during the period of study, as the average annual rate of growth in the productivity has reached respectively 11.36%, 18.63%; 7.40%, 7.44 % and 14.11%. This rate was statistically significant at the 1% level for all crops. The productivity per hectare for the crops under study has also followed an upward trend during the mentioned period with the exception of artichoke, as the amount of the annual change in productivity was about 0.15, 0.97, - 0.035, 0.28 and 0.32 respectively. This change is statistically significant at the 1% level except for artichoke.

Table 2 presents the results of the runs test. We notice that the number of runs during the period of study was 11 for potatoes and artichoke, 10 for tomatoes and onions, and 9 for pepper. The critical values for a number of runs n = 25 and α = 0.05 show that the minimum number of runs = 12 and the maximum number of runs = 21. Therefore, it is clear that the numbers of runs for all crops under study are outside the border, as they are below the minimum frontier. This indicates that the changes occurring in the productivity of crops during the study period are not random but follow a particular pattern and as such they are trended. Therefore, we can accept the hypothesis that the changes or fluctuations in the time-series data are not random and take a general trend following the interaction of the agricultural, economic and weather factors on productivity.

Table1. *Indicators of major vegetable crops (1989-2013)*

Statistical Indicators	Potatoes	Tomatoes	Artichoke	Pepper	Onions
Mean (Tons/Ha)	13.579	31.112	7.234	11.324	18.453
Standard Error (SE)	0.314	1.650	0.290	0.482	0.618
Standard Deviation	1.724	9.038	1.593	2.644	3.386
Sample Variance	2.973	81.695	2.538	6.993	11.467
Value of annual variation β	0.151*	0.974*	-0.035	0.283*	0.323*
Average annual growth (%)	11.368*	18.635*	7.407*	7.444*	14.110*

Table2. *Results of the models used in forecasting and their statistical indicators*

Statistical Indicators	Potatoes	Tomatoes	Artichoke	Pepper	Onions
Number of Runs	11	10	11	9	10
Model	Holt	ARIMA(1,0,1)	Holt	Holt	Winters
MAPE	0.142	0.107	0.132	0.153	0.094
MAE	0.188	0.151	0.175	0.329	0.341
RMSE	0.223	0.207	0.221	0.168	0.179
MAD	0.237	0.242	0.141	0.158	0.168
Theil's U	0.128	0.119	0.096	0.136	0.147
Forecast Error	0.136	0.117	0.094	0.109	0.123
Correlation coefficient between actual and forecast values	0.887	0.919	0.937	0.942	0.895

The results presented in Table 2 also demonstrate that according to the correlation coefficients between the actual and the forecast values, the Holt model seems to be the best model to predict the productivity of potatoes, artichoke and pepper, while the ARIMA model (1,0,1) is the best model to predict the productivity of tomatoes, and the model Winters is the best model to predict the productivity of onions. The correlation coefficients between actual and forecast values were respectively 0.877, 0.919, 0.937, 0.942 and 0.895 for the crops of potatoes, tomatoes, artichoke, pepper and onions. Those values represent the highest correlation coefficients among all tested models.

Table 3 presents the expected values of the productivity of the crops under study for the period 2014-2020, based on the results of the models used in the forecast of each crop. We observe that the productivity of potatoes and tomatoes is expected to increase during the 2014-20 period, while we notice the relative stability of the productivity of pepper. Meanwhile, the productivity of artichoke and onions fluctuates between an increase and a decrease during the same period.

The best forecasted results are those that are unbiased and efficient (Genberg and Martinez, 2014). The assessment of the forecast results is generally done through the regression of actual and forecast values in accordance with the following equation.

$$Fy_t = \alpha + \beta_0 Ay_t \quad (2)$$

Where: Fy_t indicates the forecast values and Ay_t refers to the actual values of the variable at period *t*. Based on the works of Granger and Newbold (1986), we can conduct a test of unbiased forecasting based on the forecast error (e_t), with:

$$e_t = Ay_t – Fy_t \quad (3)$$

Table3. *Forecast values of the productivity of main vegetable crops (2014-2020)*

Years	Potatoes	Tomatoes	Artichoke	Pepper	Onions
2014	14.76	32.11	7.92	15.432	22.69
2015	14.78	32.51	8.42	15.436	23.59
2016	14.80	33.12	7.92	15.436	22.69
2017	14.82	33.53	8.52	15.438	23.59
2018	14.84	33.92	7.83	15.439	23.92
2019	14.86	34.35	8.63	15.442	22.42
2020	14.88	34.76	8.03	15.442	24.02

Pons (2000) considers that we can also test unbiased forecasting using an OLS regression, with:

$$e_t = (Ay_t – Fy_t) = \gamma + \mu_t \quad (4)$$

In this case, the null hypothesis for unbiased forecasting is that $\gamma = 0$, which can be analyzed using the t-test. Results of this test are shown in Table 4. We notice that the estimation bias of forecasting for all crops under study are statistically different from zero at the 5% level, which means that the estimation is not biased.

Also, forecasting is considered to be efficient if the forecast errors (e_t) are statistically independent in the available data during the study period. The efficiency of forecasting was tested using the following regression equations.

$$e_t = \alpha_1 + \beta\, Fy_t \quad (5)$$

$$e_t = \alpha_2 + \rho\, e_{t-1} \quad (6)$$

The condition of efficiency is that $\beta = 0$ in equation (5) and $\rho = 0$ in equation (6). The results of this test are shown in Table 4. We notice that forecasting of productivity for all crops is efficient since the null hypothesis of $\beta = 0$, $\rho = 0$ is accepted at the 5% level.

Table4. *Results of Forecast Accuracy (Unbiasedness and Efficiency) for Vegetable Crops*

Years	Potatoes	Tomatoes	Artichoke	Pepper	Onions
Unbiasedness γ	0.0040 (3.27)	0.023 (6.05)	0.0019 (6.18)	0.0044 (3.52)	0.0048 (2.89)
Efficiency: β	-0.155 (-4.50)	-0.114 (-5.16)	-0.0058 (-6.62)	-0.271 (-3.92)	-0.342 (-3.26)
ρ	0.129 (5.58)	0.034 (6.31)	0.023 (7.20)	0.152 (4.72)	0.197 (4.07)

CONCLUSION

Forecasting is considered to be an approach that helps decision makers to take future decisions with a high degree of precision. This study aims to identify the general trend and the rate of annual growth for the hectare productivity for the following main vegetables in Tunisia: potatoes, tomatoes, artichoke, pepper and onions, based on annual time-series data covering the period 1989-2013. It aims also to forecast the short-run productivity per hectare during the period 2014-2020 for the mentioned crops.

The study was based on statistical and econometric analysis. Annual growth rates in productivity per hectare of the agricultural crops under study have been estimated and a Runs Test has been used to identify whether changes or fluctuations in the time-series data are random or trended.

Two types of time-series models were used, namely the smoothing the stochastic models. The best model of forecast of the productivity of each crop was identified. Also, we used a certain number of indicators to evaluate the accuracy and the efficiency of forecasting models such as the MAE, the

MAD, the RMSE, the MAPE and test Thiel models (U) and the standard correlation coefficient between the actual and forecast values.

The results showed that the productivity per hectare for all crops under study tended to increase as the annual growth rate of productivity of these crops has reached respectively 11.36%, 18.63%; 7.40%, 7.44 % and 14.11%. The runs test showed that changes or fluctuations occurring in the productivity of crops are not random but rather trended. In accordance with the correlation coefficient between actual and forecast values, results show that the Holt model is the best model to predict the productivity of potatoes, artichoke and pepper, while the ARIMA (1,0,1) is the best model to predict the productivity of tomatoes and the Winters model is preferred to explain the productivity of onions.

In addition, results showed that the productivity of potatoes and tomatoes is expected to increase during the 2014-2020 period, while we notice the relative stability of the productivity of pepper. Meanwhile, the productivity of artichoke and onions fluctuates between an increase and a decrease. The forecasting of productivity for all crops produced efficient and unbiased results.

REFERENCES

[1] Adrian, D. (2012). "A Model-based Approach to Forecasting corn and Soybean Yields", USDA, National Agricultural Statistics Service, Research & Development Division.

[2] Ahmad, T.; and Kathuria, O.P. (2010). "Estimation of Crop Yield at Block Level", *Advances in Applied Research*, 2(2), pp. 164-172.

[3] Agarwal, R.; Jain, R.C and Mehta, S.C. (2001). "Yield Forecast Based on Weather Variables and Agricultural Inputs on Agroclimatic Zone Basis", Indian Journal of Agricultural Sciences, 71(7), pp. 487-490.

[4] Andarzian, B.; Bakhshandeh, A.M.; Bannayan, M.; Emam, Y.; Fathi, G. and Alami, S. "Wheat Pot : A Simple Model for Spring Wheat Yield Using Monthly Weather Data", Biosystems Engineering., 99, pp. 487–495.

[5] Bazgeer, S.; Kamali, G. and Mortazavi, A. (2007). "Wheat Yield Prediction through Agrometeorological Indices for Hamedan, Iran", Biaban, 12, pp. 33-38.

[6] Esfandiary, F.; Aghaie, G. and Mehr, A.D. (2009). "Wheat Yield Prediction through Agro meteorological Indices for Ardebil District", *World Academy of Science, Engineering and Technology*, 49, pp. 32-35.

[7] Granger, C.W. and Newbold, P. (1986). *Forecasting Economic Time Series*, 2nd Edition, Academic Press, Orlando, FL.

[8] Genberg, H. and Martinez, A. (2014). "On the Accuracy and Efficiency of IMF Forecasts: A Survey and Some Extensions", IEO Background Paper, BP/14/04, IMF.

[9] Hall, D.B. and Clutter, M. (2004). "Multivariate Multilevel Non Linear Mixed Effects Models for Timber Yield Predictions", Biometrics, 60, pp. 16-24.

[10] Hebel, P.; Faivre, R.; Goffinet, B. and Wallach,D. (1993). "Shrinkage Estimators Applied to Prediction of French Winter Wheat Yield", Biometrics, 49, pp. 281-293.

[11] Holden, L.P. and Peel, D.A (1990). *On Testing for Unbiasedness and Efficiency of Forecasts*, the Manchester School.

[12] Mehta, S.C.; Agarwal, Ranjana and Singh, V.P. (2000) Strategies for composite forecast, Journal of Indian Society of Agricultural Statistics, 53(3), pp. 262-272.

[13] Ministry of Regional Development and Planning, Department of Statistics, Statistical Yearbook, Tunisia

[14] Pons, J. (2000). "The Accuracy of IMF and OECD Forecasts for G7 Countries." *Journal of Forecasting*, 19-1, pp. 53-63.

[15] Xingjie, Y.; Yongqiang, Y. and Zhang, W. (2010). "The Harvest Index Model of Winter Wheat in China Based on Meteorological Data", *Scientia Agricultura Sinica*, 43(20), pp. 4158-68.