Full Length Research Paper

Prediction of added value of agricultural subsections using artificial neural networks: Box-Jenkins and Holt-Winters methods

Elham Kahforoushan\textsuperscript{1*}, Masoumeh Zarif\textsuperscript{1} and Ebrahim Badali Mashahir\textsuperscript{2}

\textsuperscript{1}Department of Agricultural Economics, Faculty of Agricultural Engineering, University of Zabol, Zabol, Iran.
\textsuperscript{2}Tarh Ab Araz Consulting Engineers Company, Iran.

Accepted 11 January, 2010

Added value of agricultural sub sectors is affected by many factors such as quantity production per agricultural sub sectors and selling price of producers and is related to some factors such as government investment and monetary and financial policies. This study examines the performance of artificial neural network, Box-Jenkins and Holt-Winters-no-seasonal models in forecasting added value of agricultural sub sectors in Iran. It compares error criterions for determining the best model. Results showed that Box-Jenkins and artificial neural network are appropriate and artificial neural network indicated good result relatively in learn stage, but Box-Jenkins model gave better results in forecasting of unseen data. Holt-Winters model had the lowest mean absolute percent error in both of model fitting and model validation stages.

Key words: Artificial neural network, Box-Jenkins, Holt-Winters, added value of agricultural sub sectors.

INTRODUCTION

Agricultural sector is regarded as one of the most important economical parts of Iran. It holds a special position considering production, employment, food security, foreign exchanges and relative advantages and plays a key role in economical development process. Added value of agricultural section includes added value of each of agronomy, animal husbandry, fishery and forestry subsections. The percentage portion of added value of agricultural section is about 13% in combination of national economy (Database of Central Bank of Iran). Determining added value in agricultural section is considered as one of the necessities of economical calculations which is required for different decision makings and specifying important policies in investment, employment, income and principally development planning and the likes and also in the evaluation of five-year program for agricultural development. So it is calculated to satisfy above mentioned requirements through study of amount of production of agricultural subsections, sale prices of producers of agricultural section, gross value of products of each of agricultural subsections, middleman costs of each of subsections, added value of each of subsections and finally added value of all of agricultural section at constant and current costs.

Several studies refer to relations between state investment and added value and also long-term relations between fiscal and monetary policies and added value of agricultural section. Having exact information from future events is necessary for appropriate short and long-term planning on added value since choosing appropriate policies is required for formation of agricultural section as an effective section in economical development. Nowadays, prediction of future events has attracted attention of researchers in different aspects and a variety of methods has been innovated, among them, we can refer to non-regression methods such as simple average, moving average, exponential adjustment and regression methods including GARCH, ARCH, ARIMA and artificial neural network methods (Najafi and Tarazkar, 2006). Added value of agricultural subsections has been forecasted through using three Holt-Winters, ARIMA and artificial neural network methods at this study, aim of which is

\*Corresponding author. E-mail: e_kahforoushan@yahoo.com. Fax: +984115248613.
which is comparing accuracy of these methods with error criteria.

Background

Several studies have been done on different economic variables alongside many predictions, and different methods have been used in each of them with different outcomes resulting. We refer to some samples in order to compare and have a more complete understanding of the subject. Henry et al. (2007) compared efficiency of artificial neural network with seasonal Holt-Winters and ARIMA models for prediction of rice exporting of Thailand and concluded that Holt-Winters model is an appropriate and satisfactory one in prediction of rice monthly export at modeling stage. But they weakly performed in prediction of future unspecified amounts.

In contrast, ANN demonstrated the seasonal and nonlinear effects relatively well. Deetae (1991) applied analysis and Box-Jenkins methods for evaluation of rice price in the farm and the efficiency of Box-Jenkins model became obvious. Kerdsomboon (1999) used preliminary statistic prediction models in order to study rice production and demonstrated the better performance of Box-Jenkins method. Sangpattaranate (2005) applied four additive Holt-Winters's prediction techniques, Box-Jenkins and regression analysis for Thailand rice prices and made it clear that despite of relatively well performance of analysis model, the Box-Jenkins model was the best. Baki Billah et al (2006), in a study under title of choosing exponential smoothing model for prediction, introduced MAPE criterion as an appropriate method for model choosing at model evaluation stage. The method with procedure is appropriate for annual data and seasonal method is appropriate for time series less than annual from among three single exponential smoothing, with time and seasonal procedure methods. Hyndman (2001) used single exponential smoothing method in his study and demonstrated that its performance is better than ARIMA. Nadjafi and Tarazkar (2006) applied artificial neural network and ARIMA procedure for prediction of pistachio export of Iran and compared the results. The study results demonstrated that feed-back neural network has better performance comparing with other neural networks and ARIMA procedures and is able to anticipate the amount of pistachio exports more exactly. Sadrolashrafi and Nassabian (2002) in a study on effects of difference of optimized and compiled added value of agricultural subsections on optimized GDP through using input-output table of 1991 concluded that the government can achieve 57 and 71% from the complied added values of animal husbandry and forestry subsection and also 37 and 45% of the complied added values of agronomy and fishery subsections considering population annual growth rate and domestic gross production annual growth rate scenario in the third development program. Kopahi and Kiani (2000) forecasted the amount of required investment in agricultural section from 2000 - 2004 through finding a relation between state investment and added value of agricultural section.

An investment function with a time interval has been used in this regard. Akbari et al. (2003), in a study on the effects of the government's expenditures on added value of agricultural subsection, introduced learning and research charges of the government as the most effective variable on added value of agricultural section. Civil charges occupy the next place. Current costs and subsides paid to the producers of this section has statistically meaningful effect on added value of agricultural section. Torkamani and Bagheri (2002), through study of the relation between private and state investment and growth of added value in agricultural section demonstrated that variables of the ratio of private and state investment to added value and growth of export in agricultural section have a positive effect on the growth of this section and variables of growth of agricultural import and growth of employment have a negative effect. Growth of added value variable of agricultural section just has a bilateral relationship with private and state investment and a unilateral relationship with other variables. Co-integration relations have been estimated through Johansson' co-accumulation test and use of VAR model.

METHODOLOGY

In this study, added value of agricultural subsections has been forecasted through using three Holt-Winters, ARIMA and artificial neural network methods. Its goal is to compare the accuracy of these methods with error criteria such as MAPE (Mean Absolute Percent Error), MSE (Mean Square Error), MAE (Mean Absolute Error) and ME (Mean Error). Time series of 1936 - 2005 years of four agricultural subsections is the data used in this study. Data of 1959 - 2005 years has been extracted from database of Central Bank of Iran and the statistics of 1936 - 1958 years has been extracted from research plan of domestic gross production estimation for 1936 - 1958 years done by bank and monetary research institute of Central Bank (Khavarinezhad, 2001) and real prices of 1997 have been used in order to omit inflation effects from statistics.

Exponential smoothing

This method began with the work of Holt-Winters in 1950. Afterwards, the main model exposed to many changes and several studies has expressed its exceptional performance. Bowerman and O'connell (1979) stated that this model is on the basis of an autoregressive statistical model in which the information in the relation with forecasted series is just used and in contrary to regression models predictions using constant factors, the predictions of this model adjusted according to the previous predictions errors (Hyndman, Koehler, Snyder and Grose, 2002) classified exponential smoothing.

Each method consisted of one of five types of trend (none, additive, damped additive, multiplicative and damped multiplicative) and one of three types of seasonality (none, additive and
multiplicative). Among the 15 different exponential smoothing methods, the best known are simple exponential smoothing (no trend, no seasonality), Holt’s linear method (additive trend, no seasonality), Holt-Winters’ additive method (additive trend, additive seasonality) and Holt-Winters’ multiplicative method (additive trend, multiplicative seasonality) (De Gooijer et al., 2006). Gardner and McKenzie (1988) provide some simple rules based on the variances of differenced time series for choosing an appropriate exponential smoothing method. Hyndman et al. (2002) also proposed an information criterion approach, but using the underlying state space models.

Different methods of smoothing require primary estimation of parameters related to ($\alpha$, level), ($\beta$, trend) and ($\gamma$, seasonality) (Hyndman, 2002). In Eviews software, parameters are estimated by minimizing sum of error squares and in QSB software, error criteria is selected by the researcher and the software chooses the best parameter considering error criteria. These parameters have been defined between zero and one. Theory of the method used in this study will be discussed later.

**Single exponential smoothing**

This is a single-parameter method and is appropriate for series randomly move around a constant mean. If the parameters are estimated close to one, means that this procedure is a random step. Parameter $\alpha$ is related to level (mean), $\beta$ to the trend and $\gamma$ to the seasonality. These parameters have been defined between zero and one.

If the smoothed series is $\hat{y}_t$ and the main series is $y_t$, $\hat{y}_t$ is randomly obtained from the following relation:

$$\hat{y}_t = \alpha y_t + (1-\alpha) \hat{y}_{t-1}$$

(1)

In which, 0 ≤ $\alpha$ ≤ 1 is the smoothing factor, $\hat{y}_t$ series are smoothing. We can rewrite the relation (1) through repetitive replacements:

$$\hat{y}_t = \alpha \sum_{s=0}^{t-1} (1-\alpha)^s y_{t-s} \quad (2)$$

This relation states the reason of calling this method as exponential smoothing. $y_t$ prediction is a weight mean of previous $y_i$ values in which weights reduces exponentially with time. These predictions have been proved for all future observations. This constant is obtained through following relation:

$$y_{t+k} = y_t \quad \text{for } k > 0$$

Where $t = \text{end of the sample}$. We require a value for $\alpha$ and an initial value for $y_t$ in order to begin repeat procedure. Bowerman and O’connell (1979) suggested that $\alpha$ values around 0.01 to 0.3 perform relatively well.

**No-seasonal Holt-Winters**

No-seasonal Holt-Winters used in this research is a double-parameter method and is appropriate for series with time trend and without seasonal pattern. In this method, predictions are made with trend and without seasonal fluctuations.

$$y_{t+k} = a(t) + b(t)k$$

located on trend with $a(t)$ intercept and $b(t)$ slope.

**Box-Jenkins**

Stationary time series can be modeled in different ways. If a time series become stationary after $d$ times of difference then modeled by ARIMA procedure, the main time series will be called integrated moving average autoregressive time series in which $p$ is the order of autoregressive and $q$ is the rank of moving average process:

ARIMA ($p, d, q$): $y_t = \theta + \alpha_1 y_{t-1} + \beta_1 t + \beta_2 t + \beta_3 t + \epsilon_t$

Box-Jenkins has four stages; the first stage is the identification stage in which the real values of $p, d, q$ is determined by a single root test or correlogram. At the second stage the parameters of the model are estimated. Third stage is called recognition control sought after choosing of a special model of ARIMA and estimation of its parameters in order to determine that whether it appropriately fits the data since it is possible another ARIMA model fits better. For this purpose, we can use function of Box-Pieres (Q) and Lijang-Box (LB) tests. At the fourth stage the estimation is done with the best selected model. The predictions resulted from this model is especially for short-term predictions and in most cases, is more reliable than traditional modeling method of econometrics. Of course, it is necessary to separately judge about each special case (Abrishami, 2006).

**Artificial neural network**

The structure of artificial neural networks is like human’s brain including a set of connected neurons. Each set is called a layer. For example, an ANN consists three neural layers (node):

- **Input layer**, which receives outside information and acts as an independent variable. So, the number of neurons of input layer is determined on the basis of nature of the problem and depends on the number of independent variables. The last layer is the output layer, from which the solution of the problem results.

Several numbers of middle layers have separated input and output layers which is called hidden layers and are merely an intermediate result in calculation process of output value. The nodes have been connected in adjacent of the layers from lower layers to upper ones by a rotative arch. There is no theoretical principle for determining appropriate number of hidden units or layers in a network. Researchers have used different relations including $n/2, n, 2n, 2n+1$ in order to determine number of hidden neurons. In these relations, $n$ is number of input neurons. Trial and error is the best way for determining number of hidden nodes (Zhang et al., 1999). Figure 1 shows the simplest node in which sum of $N$ harmonious inputs enters the network. In this research we used feed-back network in which F is output function, $\beta_0$ is bias unit (equivalent to 1), G is output function of j units of hidden layers, $\theta_{kj}$ refers to input weight of k of j neuron, $\beta_j$ is output weight from hidden layers in output layers unit and X is the input vector.

$$F = F \left[ \beta_0 + \sum_{i=1}^{j} \beta_j G \left[ \sum_{k=1}^{n_j} \theta_{kj} X_k \right] \right]$$

Activation function determines the relationship between inputs and outputs of a node and a network. It is possible that a network have different activation functions for different nodes at the same or different layers, but nowadays almost all researchers use similar activation functions for nodes at the same layers (Zhang et al., 1998). Neural network helps the smallest error of prediction of learning set through using criteria such as mean square error and...
the data should be normalized for presentation by using following formula:

\[ N_i = 0.8 \times \left[ \frac{X_i - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \right] + 0.1 \]

\( N_i \) and \( X_i \) represent normalized and main data; \( X_{\text{min}} \) and \( X_{\text{max}} \) represent minimum and maximum of data (Haykin, 1994). Data normalization at the meaning of preprocessing and post processing improves the performance of the network. Data are usually preprocessed before learning of the network meaning exercising some conversions on inputs and outputs of the network in order to extract the properties from the inputs and change of the output to a more understandable form for the network. Inputs of the network are changed to their primary form, referred to post processing, after learning and extraction of results from the network.

**Compare of prediction methods**

Criteria such as MAPE and MAE are used for examining of prediction power.

**Mean Absolute Error**

\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |e_i| \]

**Mean Square Error**

\[ MSE = \frac{1}{n} \sum_{i=1}^{n} e_i^2 \]

**Mean Absolute Percent Error**

\[ MAPE = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{|e_i|}{y_i} \right) \]

In these relations, \( n \) is the number of predictions, \( e_i \) is the prediction error obtained from the difference between predicted and real values. A MAPE criterion is one of the percent error criterions which is more desired and is one of the most applied unit-less criteria. A single error criterion such as MAPE is mostly used for comparing several time series with different measures (Najafi and Tarazkar, 2006).

**RESULTS AND DISCUSSION**

**Exponential smoothing model**

Single exponential smoothing with linear trend method has been used in this study which is the same as Holt-Winters' non-seasonal method. This method has been selected considering nature of data which is annually with a linear trend and also MAPE criterion at the evaluation stage of the model by WinQSB. Modeling has been done considering data of the period 1936 to 1996; smoothing parameters and error criteria have been listed in Table 1.

The goal of model validation is to show that the constructed model has the capability for producing data acceptable for future. Data used in this stage relate to
Table 1. Smoothing parameters and error criteria of modeling and model validation stages.

<table>
<thead>
<tr>
<th>Sub sector</th>
<th>α (Average)</th>
<th>β (Trend)</th>
<th>MAP</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agronomy</td>
<td>0.92</td>
<td>0.08</td>
<td>8.9</td>
<td>547149.6</td>
<td>544.33</td>
</tr>
<tr>
<td>Animal husbandry</td>
<td>0.85</td>
<td>0.61</td>
<td>5.01</td>
<td>113588.1</td>
<td>225.012</td>
</tr>
<tr>
<td>Fishery</td>
<td>1</td>
<td>0.19</td>
<td>10.6</td>
<td>11707</td>
<td>57.13</td>
</tr>
<tr>
<td>Forestry</td>
<td>1</td>
<td>0.01</td>
<td>14.9</td>
<td>1585.24</td>
<td>22.49</td>
</tr>
</tbody>
</table>

Model fitting

<table>
<thead>
<tr>
<th>Sub sector</th>
<th>β (Trend)</th>
<th>MAP</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agronomy</td>
<td>0.92</td>
<td>6.3</td>
<td>4385428</td>
<td>1761.22</td>
</tr>
<tr>
<td>Animal husbandry</td>
<td>0.85</td>
<td>2.4</td>
<td>222237.7</td>
<td>368.82</td>
</tr>
<tr>
<td>Fishery</td>
<td>1</td>
<td>15.3</td>
<td>38277.46</td>
<td>182.45</td>
</tr>
<tr>
<td>Forestry</td>
<td>1</td>
<td>13.1</td>
<td>11214.39</td>
<td>83.63</td>
</tr>
</tbody>
</table>

1997 - 2005 period which was regarded as out of sample data at modeling stage. Considering the less amount of mean absolute value percent error, it is evident that exponential smoothing method has presented very good results in both stages.

Box-Jenkins model

Augmented Dickey-Fuller test statistic was used in order to survey the static mode of time series. In this test, the equation \( \Delta y_t = \alpha + \beta t + \delta y_{t-1} + V_t \) is tested through adding \( \Delta y \) breaks in order to omit correlation of estimation disorder sentences and being zero of \( \alpha, \beta, \delta \). Results of the test demonstrate that none of the variables are at static level but they reach static level after one time of differentiation (Noferesti, 1999).

We determine the number of autoregressive sentences (p) and number of motion mean sentences (q) in order to predict with ARIMA procedure after determination of static level of the variables which requires trial and error and examining of different models. Schouarts-Bazin values is one of the criteria that helped us in choosing the model, the best model is selected considering the smaller values and also remainders test by Q Box-Pieres test function. Also, it is necessary to take into account the meaningfulness of AR and MA correlations and also the value of \( R^2 \). After estimation of ARIMA models for each of the agricultural subsections by using sample period of 1936 - 1996, prediction was made for out of sample data until 2005 and error criteria of modeling stage were estimated. Then, error criteria of model validation stage were calculated again through using real available data for 1997 - 2005 periods. The results have been shown in Table 2. Considering MAPE in these two stages, it can be observed that modeling of out of sample data predictions are more accurate.

Artificial neural network

The result were tested for feed-back neural network having 2, 3, 4 and 5 layers and finally the best evaluation with a higher criterion of \( R^2 \) was specified for each subsection and error criteria were calculated in both network learning and its test in both stages which have been shown in Table 3. As it can be observed, mean absolute value percent error in learning stage was higher than network testing stage in all subsections and this demonstrates the better performance of neural network in prediction of future data.

SUMMARY AND CONCLUSION

In this study, artificial neural network, Box-Jenkins and Holt-Winters methods compared with each other in order to predict the added value of each of the agricultural subsections. The form of primary time series and single root tests demonstrated that time series of all four sections have procedure and non-seasonal Holt-Winters (single exponential smoothing with time trend) and Box-Jenkins models were selected to be compared with artificial neural network because of their capability in description of time series data trend.

The non-seasonal Holt-Winters model requires estimation of parameters of level and trend solved with trial and error and minimizing criterion of mean absolute value percent error. It is possible that a prediction model shows better results at modeling stage but has not a good performance at prediction stage. Table 4 demonstrates the error criteria of models at modeling and validation stages. The results manifest that non-seasonal Holt-Winters model has performed very well in both stages and mean absolute value percent error was less than 15%, while artificial neural network performed better than Box-Jenkins at modeling stage. The mean absolute value percent error was less than 42% for all groups through using artificial neural network at this stage, while this criterion was less than 62% for Box-Jenkins. The performance of Box-Jenkins was better than artificial neural network at testing and validation stage of models and mean absolute value percent error was less than 23% while it has been calculated as less than 36% for artificial neural network.

Although, according to MAP criteria, artificial neural network has more errors then ARIMA model considering test section, this can not lead to conclusions with regard to weakness of artificial neural network. Remember that
neural networks need more data for learning and examination. So, changes in number of learning and examinational data may have different results. Other criteria (Tables 2 and 3) demonstrate that artificial neural network operates better than ARIMA at learning stage, that is, identification and description of relations in
Table 5. Predicted value using best model until 2011.

<table>
<thead>
<tr>
<th>Year</th>
<th>Agronomy</th>
<th>Animal husbandry</th>
<th>Fishery</th>
<th>Forestry</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>35613.35</td>
<td>17368.90</td>
<td>1375.09</td>
<td>725.62</td>
</tr>
<tr>
<td>2007</td>
<td>36534.39</td>
<td>17770.16</td>
<td>1405.05</td>
<td>742.51</td>
</tr>
<tr>
<td>2008</td>
<td>37455.43</td>
<td>18171.42</td>
<td>1435.00</td>
<td>759.40</td>
</tr>
<tr>
<td>2009</td>
<td>38376.46</td>
<td>18572.68</td>
<td>1464.96</td>
<td>776.30</td>
</tr>
<tr>
<td>2010</td>
<td>39292.50</td>
<td>18973.93</td>
<td>1994.91</td>
<td>793.19</td>
</tr>
<tr>
<td>2011</td>
<td>40218.54</td>
<td>19375.19</td>
<td>1524.87</td>
<td>810.08</td>
</tr>
</tbody>
</table>

agricultural subsections. The outcomes demonstrate that use of several study criteria can lead to obtaining more reliable results because in some cases, results of prediction criteria are different. Also, normalizing of data plays a significant role in improving of function of artificial neural network. Predictions made for the next year have been demonstrated in Table 5.

The government should take actions for increasing of added value growth through appropriate planning in order to maintain annual growth rate of Iran economy which is about 6% and considering almost 2.5% growth predicted for added value of agriculture section in spite of continuing of ascending process of added value of agricultural subsections. Considering that each of the economical sections should obtain part of their added value through promoting of productivity, it is necessary to provide opportunity for growing of added value through required planning by the country's planning and management authorities in order to promote productivity of this section.

REFERENCES

Nesabian Sh, Sadrolashrafi M (2002). The impact of difference between optimal value added and codified for Agricultural subsectors on optimal GDP in third programming of development. Iranian J. Agric. Sci. 8: 35-42.