

Selection of an Appropriate Model out of VAR, ARIMA and ANN Models for Prediction of Rice Production in Iran

Mansoureh Aligholi¹, Mahboobeh Mohammadi²

^{1,2} Department of Business Management, Faculty of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran

Received: June 26, 2014
Accepted: September 4, 2014

ABSTRACT

The main target of the present research is to select an appropriate model for predicting Iranian rice production level out of ANN (artificial neural network), VAR and ARIMA (econometrics models) models. Determination of error values of each of the methods and comparison of the models are subsidiary targets of the research. Statistical population of the research included amount of produced rice in provinces: Guilan, Mazandaran, Golestan, Fars and Khuzestan. Statistical population and sample are the same. Matlab7 and Eviews software were used in this research. Data were collected from website of Chamber of Commerce of Iranian Ministry of industry, Mining and Commerce and were compared with Ministry of Agriculture information bank statistics. In order to analyze data, data from recent 10 years were collected for prediction of rice production from Ministry of Agriculture and the three models were used. The models were compared using prediction power evaluation indices like MAD index, MSE index and so on. Results showed that artificial neural network model was better than ARIMA and VAR models in terms of 4 error criteria and is an appropriate model for prediction of rice production.

KEYWORDS: prediction, decision-making, artificial neural networks, neuron or node, biological neural network

INTRODUCTION

Prediction of future has been a challenging issue in economics and capital market for economic researchers and finance experts. Exact predictions are necessary especially in planning for basic goods market adjustments. Rice is one of the strategic products and one of the main cereals for human. Rice is the second highly-consumed agricultural product in Iran after wheat. Therefore, it is necessary to have enough information about its future consumption and production in order to be able to do short-term and long-term planning. Prediction is an important element in decision-making and its main goal is to reduce risk in decision-making. Prediction is of great importance for many individuals like managers, present and potential stockholders, creditors and employees. Rice is the only grain which is sown specifically for human consumption. In Iran, it is the second highly-consumed agricultural product after wheat. In the years after Islamic Revolution, Iranian government has always paid a lot of attention to its production. Rice is a strategic crop in developing and Asian countries. Almost 90% of the production and consumption of this crop is done in Asia and 96% of this value belongs to developing countries (Nouri, 2005, p2).

Internal rice production does not cover its demand and Iranian government imports rice in order to balance supply and demand. Iran is one of the main rice importers and it is necessary to adopt appropriate policies in order to reduce rice import and increase its internal production. This is not possible without having enough information about future production status (HoseiniKordKalayee&Kalashemi, 2010, p 93).

Therefore, the main question in the present research is: what is the appropriate model for prediction of high-yield rice production in Iran (out of ARIMA, VAR and ANN models) in terms of exactness and error criteria and prediction power?

Theoretical literature

Prediction concept:

In Aryanpour (English-to-Persian) dictionary, forecasting means guessing future events out of present evidence and prediction means telling future events before their occurrence. In spite of this, both words are approximately the same and mean "pishbini and pishgouyee" (foretelling) (KhaleqiTbar&Mashayekhi, 2011, 7).

Prediction means description of qualitative and quantitative changes in a phenomenon or event which takes place in future.

Prediction is an essential element in management decision-making because final efficiency of any decision depends on a series of events. Prediction enables us to guess uncontrollable aspects of events before making decisions so that better selections can be made (Niroumand, 2004, 73).

*Corresponding Author: Mahboobeh Mohammadi, Department of Business Management, Faculty of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran. Email: gcmohamadi@yahoo.com, Tel: (+98)21-66921499

Decision-making concept:

Decision-making is an inseparable element of management and it refers to selection of a solution out of many solutions and managers must make decisions at different settings (ibid, p 75).

Artificial neural network (ANN):

A neural network is in fact a data-processing system which has some executive characteristics similar to that of human neural networks. An artificial neural network is in fact a simplified model of central neural system of human and conveys data rule to network structure via empirical data process just like human brain (Shayegan, Mohammadi&Mousavi, 2005, p 89). Artificial neural networks structures are like biologic networks. ANNs are systems which are able to do operations like natural neural networks. In other words, they can copy human brain features. They have many applications in solving different problems and algorithms like signal processing, time series prediction, and image processing and so on. They are used also as instruments for data mapping from entrance space to exit space. Neural network is made up of artificial neurons (Menhaj, 2008, p 25).

Neuron or node:

A neuron or node is the smallest data processing unit which is the base of neural networks performance. Each neuron receives inputs and produces an output after processing the inputs (ibid, p 35).

Biological neural network: human brain is a data processing system and is made up of main structural elements called neuron. In fact, neurons are the simplest structural units of neural systems. The number of neurons in human brain is estimated to be about 100 trillion and there are more than hundreds of types of neurons. These neurons are connected to each other via 10 communication channels. Neurons are classified into groups called network or texture and each network includes several thousands of neurons. Neural textures are responsible for conveying information and message from one part of body to another. Therefore, human brain can be considered as a set of neural networks (SHayegan, Mohammadi&Mousavi, 2005 p 89).

Research models:**VAR model:**

VAR model is demonstrated as follows in matrix form:

$$Y_t = X_t \cdot \beta + \sum A_s \cdot Y_{t-s} + U_t$$

$$U_t = IN(0, \sigma^2), E[u_t \cdot u_t] = \psi$$

VAR methodology is largely similar to simultaneous equations. The only difference is that in simultaneous equations, some variables are endogenous and some variables are exogenous or predetermined but this is not true about VAR model. One of the characteristics of auto-regression models is their non-theoretical bases and there is no need to theoretical fundamentals for building model. The first stage in estimation of this model is investigation of the fact of being stationary of the variables of time series. If variables were static, there is no problem but if they were non-static, their correlation degree must be specified and this is determined via augmentedDickey-Fuller test (ADF).

If Dickey-Fuller statistic is less than the calculated value, the variables are stationary or correlated with zero degree L(0). If variables become zero with a subtraction, they are of degree one I(1). In the latter case, the next step will be autocorrelation test using Touhanson test. Further, it must be noted that long-term relationship between variables is not intended in the present research but we look for variables prediction (Isfahaniyan& Amin Naseri, 2008, 366).

1. ARIMA model (Autoregressive integrated moving average):

Box and Jenkins (1970) is a prediction instrument which is technically known as ARIMA methodology. This method is based upon principle of conservation. In other words, a model with fewer parameters is preferred to models with more parameters. This preference has some advantages: 1. Selection of a simple parameter increases exactness of parameters estimation. In other words, decrease in the number of estimated parameters increases degree of freedom. 2. Selection of a simple model helps avoid large number of models. In general, this method has 4 steps. The first step is called experimental identification. In this step, autocorrelation of sample is conducted using autocorrelation function of sample and partial function. In the second step, parameters are estimated (estimation stage).

Third stage is called fit exactness test. In this stage, sufficiency of experimental identification and estimation is evaluated. If the model was proved to be unfit, the model must be adjusted.

Identification methods help researchers with making decisions about way of model adjustment and improvement. After reaching final model, it can be used for prediction of future values of time series (prediction stage). Box-Jenkins prediction method is an iterative method. In other words, if an identified experimental

model is proved to be unfit, we must turn back to experimental identification stage again and select a better model and after estimation of model parameters, its fitness must be investigated. This cycle continues over and over until an appropriate final model is reached.

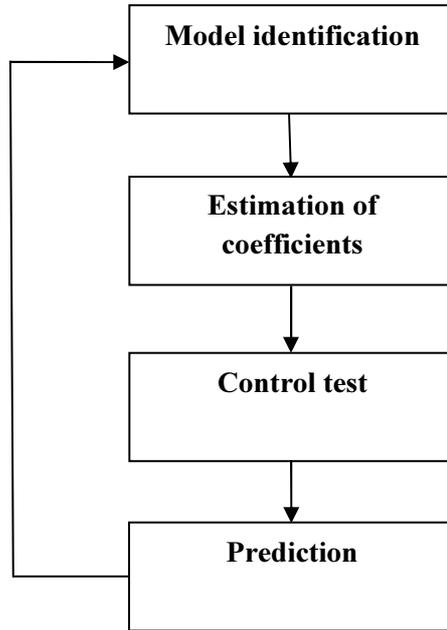


Figure 1.Steps of implementation of Box and Jenkins methodology

Box-Jenkins method is applicable to continuous sectional and discontinuous data. Therefore, data must be measured in equal time periods.

In general, a process is called ARIMA(p, q) when it contains p orders of autoregressive terms and q orders of moving average terms. In other words, it contains p orders of term with lag from dependent variable and q orders of term with lags from disorder terms. A general model of ARIMA is as follows:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Artificial neural network

Neural networks are computational models which are able to determine relationship between inputs and outputs of a physical system by a network of nodes which are all connected to each other. Further, activity level of each connection is regulated by historical information (learning process) and finally, the model will be able to discover rules concerning inputs and outputs, although these rules are non-linear and complex. A neural network includes constituents of layers and weights. Network behavior depends on relationship among members.

In order to conduct a prediction and obtain output from network, first we must consider some lag for data before writing program. The number of lags or the very inputs is calculated via trial and error and different modeling methods.

In order to specify the optimum number of lags, we can consider model sensitivity analysis. To this end, we keep other parameters of the program constant and therefore we can eliminate ineffective and extra lags from our model. By ineffective we mean a lag from which the network learns the minimum learning.

After this stage, we make data random. The final result of this stage is a set of inputs and outputs without any particular system. After data randomization, the amount of statistics and information which must be taught is specified in the program. Therefore, some data are considered for education. A part of the other data is used for data calibration and some other data are considered as test data. data Calibration is used for re-measurement of information. Another important point which must be considered when writing a program is stimulation function of learning coefficient and momentum. For sigmoid function, learning coefficient was assumed to be equal to 1, which has been obtained by means of trial and error and experience. Momentum also links current change in weight to previous change in weight and also current error. After network structure was specified and inputs collection was given to network, the network is prepared for instruction. The network provides output using the weights which are selected in the first step at random. Then, it compares its output with real values using performance function of sum of errors squares and regulates weights using the lowest expected error. It iterates these steps until it reaches minimum expected error.

Investigation of models predicting powers:

After prediction steps end and values are obtained, the following criteria are used for comparing prediction power and selection of the best prediction method:

$$RMSE = \left(\sum_{i=T}^{T+h} (\hat{y}_t - y_t)^2 / h \right)^{1/2}$$

RMSE: root mean square error

$$MAPE = \left(\sum_{t=T+1}^{T+h} |\hat{y}_t - y_t / y_t| / h \right)$$

MAPE: mean absolute percentage error

$$MAE = \left(\sum_{t=T+1}^{T+h} \hat{y}_t - y_t \right) / h$$

MAE: mean absolute error

$$MSE = \frac{1}{T} \sum (\hat{y}_t - y_t)^2$$

MSE: mean square error

The model which has the lowest prediction RMSE is selected as the best prediction model.

Research questions:

Which model has the best prediction power in terms of exactness and error criteria?

Subsidiary questions are as follows:

1. is ANN model better than ARIMA model in terms of RMSE, MAPE, MAE and determination coefficient in prediction of high-yieldrice production prediction?
2. is ANN model better than VAR model in terms of RMSE, MAPE, MAE and determination coefficient in prediction of high-yieldrice production and import prediction?

Statistical population

Statistical population of the research was produced rice in provinces Guilan, Mazandaran, Khuzestan, Fars and Golestan from 2002 to 2012. Considering the fact that necessary data were collected from Agriculture Ministry databank and Commerce Chamber of Ministry of Industry databank, statistical population and sample were the same.

Data collection method and instrument(s)

Field data were collected from website of Ministry of Commerce, Industry and Mining Chamber of Commerce and also from databank of National Iranian Commerce Holding and Ministry of Agriculture databank.

First, 10-year data on independent variables were collected from Ministry of Agriculture for prediction of national rice production. Then, three prediction models (VAR, ARIMA, and ANN) were used for prediction of national rice production. Afterwards, the models were compared in terms of prediction power evaluation indices like MAD, MSE, and so on. Finally, the fittest model for prediction of rice production was identified. MATLAB, EXCEL and other software were used for data analysis.

Data analysis

In this section, data were collected and then analyzed in order to use prediction models (VAR, ARIMA and ANN) in order to predict rice production.

Modeling by ANN

Characteristics of neural network models which were built for the present research are summarized in table 1.

Table 1: characteristics of neural network models

| R ² | MSE | SSE | Maximum occurrence | Middle layer nodes number | Momentum | Learning coefficient | input | model |
|----------------|--------|----------|--------------------|---------------------------|----------|----------------------|-------|----------|
| 0.9763 | 1.4834 | 8231.53 | 7000 | 16 | 0.7 | 1 | 8 | first |
| 0.9784 | 1.3572 | 7534.26 | 10000 | 12 | 0.7 | 1 | 6 | second |
| 0.9814 | 1.1738 | 6517.38 | 20000 | 10 | 0.7 | 1 | 5 | third |
| 0.9774 | 1.4001 | 7752.86 | 10000 | 40 | 0.7 | 1 | 20 | fourth |
| 0.9827 | 1.0657 | 5901.05 | 20000 | 40 | 0.7 | 1 | 20 | fifth |
| 0.9793 | 1.3092 | 7268.69 | 10000 | 10 | 0.7 | 1 | 5 | sixth |
| 0.9791 | 1.3296 | 7273.21 | 10000 | 10 | 0.7 | 0.5 | 5 | seventh |
| 0.9742 | 1.7055 | 9441.47 | 20000 | 16 | 0.7 | 1 | 8 | eighth |
| 0.9744 | 1.4892 | 8196.46 | 20000 | 40 | 0.7 | 1 | 20 | ninth |
| 0.9738 | 1.6678 | 9218.12 | 10000 | 40 | 0.7 | 1 | 20 | tenth |
| 0.9706 | 1.9922 | 11036.93 | 10000 | 14 | 0.7 | 1 | 7 | eleventh |
| 0.9742 | 1.6819 | 9318.24 | 10000 | 14 | 0.85 | 1 | 7 | twelfth |

As it can be seen in table 1, a comparison of the sixth and seventh models reveals that decreasing learning coefficient- while keeping other parameters fixed- can lead to MSE increase. Therefore, as learning coefficient of the model increases and approaches to 1, better results are achieved.

On the other hand, if we look at the eleventh and twelfth models, we can conclude that if we increase momentum number in the model (while keeping other parameters constant), better results are achieved and MSE and determination coefficient (R²) decrease and increase, respectively.

Ratios of Sensitivity coefficients of outputs of each model to inputs (or the very وقفهها) have been summarized in table 2. It must be mentioned that sensitivity coefficient indicates the level of interaction or learning of network output from inputs.

Table 2. Sensitivity coefficients of outputs of each model to inputs

| twelfth | eleventh | tenth | ninth | eighth | seventh | sixth | fifth | fourth | third | second | first | input |
|---------|----------|-------|-------|--------|---------|-------|-------|--------|-------|--------|-------|-------|
| 0.215 | 0.205 | 0.139 | 0.209 | 0.205 | 0.105 | 0.140 | 0.163 | 0.097 | 0.177 | 0.100 | 0.056 | L1 |
| 0.094 | 0.101 | 0.109 | 0.123 | 0.142 | 0.126 | 0.134 | 0.145 | 0.090 | 0.141 | 0.135 | 0.109 | L2 |
| 0.130 | 0.099 | 0.102 | 0.095 | 0.122 | 0.156 | 0.108 | 0.098 | 0.058 | 0.104 | 0.082 | 0.093 | L3 |
| 0.081 | 0.085 | 0.057 | 0.092 | 0.084 | 0.105 | 0.092 | 0.063 | 0.094 | 0.081 | 0.106 | 0.093 | L4 |
| 0.069 | 0.056 | 0.061 | 0.092 | 0.90 | 0.109 | 0.128 | 0.052 | 0.063 | 0.100 | 0.101 | 0.068 | L5 |
| 0.016 | 0.023 | 0.069 | 0.009 | - | - | - | 0.032 | 0.030 | - | 0.076 | 0.055 | L6 |
| 0.039 | 0.073 | 0.047 | 0.021 | - | - | - | 0.037 | 0.025 | - | - | 0.078 | L7 |
| - | - | 0.043 | 0.007 | - | - | - | 0.022 | 0.031 | - | - | 0.056 | L8 |
| - | - | 0.018 | 0.028 | - | - | - | 0.025 | 0.062 | - | - | - | L9 |
| - | - | 0.034 | 0.045 | - | - | - | 0.018 | 0.023 | - | - | - | L10 |
| - | - | 0.025 | 0.016 | - | - | - | 0.01 | 0.004 | - | - | - | L11 |
| - | - | 0.008 | 0.006 | - | - | - | 0.027 | 0.031 | - | - | - | L12 |
| - | - | 0.018 | 0.004 | - | - | - | 0.015 | 0.003 | - | - | - | L13 |
| - | - | 0.054 | 0.007 | - | - | - | 0.012 | 0.013 | - | - | - | L14 |
| - | - | 0.024 | 0.016 | - | - | - | 0.007 | 0.017 | - | - | - | L15 |
| - | - | 0.017 | 0.002 | - | - | - | 0.011 | 0.001 | - | - | - | L16 |
| - | - | 0.015 | 0.016 | - | - | - | 0.010 | 0.024 | - | - | - | L17 |
| - | - | 0.003 | 0.002 | - | - | - | 0.028 | 0.003 | - | - | - | L18 |
| - | - | 0.001 | 0.003 | - | - | - | 0.005 | 0.014 | - | - | - | L19 |
| - | - | 0.005 | 0.004 | - | - | - | 0.002 | 0.021 | - | - | - | L20 |

In addition to MSE and determination coefficient parameters, it can be said that network performance will be optimum when it is affected mostly by the first input because the first input has only one day lag. A network which receives the highest impact from the second and third inputs instead of the first input has a kind of disorder in its performance because it has more learning from further lags.

If we look at table 2, we will see that the eighth, ninth, eleventh and twelfth models have higher sensitivity coefficients with respect to the fifth model but have higher MSEs and lower determination coefficients.

Modeling by ARIMA:

In order to model prediction using ARIMA, after conducting stationary test on variables, autocorrelation graphs were investigated and primary experimental model for prediction is identified considering peaks and troughs of the diagrams. Then, model parameters are estimated and model sufficiency remaining elements are investigated using stationary test.

We used Dickey-Fuller test in order to investigate stationary of variables. The results are summarized in table 4.

Table 4. Results of testing variables stationary

| ADF Test Statistic | | -3.526.525 | | 1% Critical value | -3.4347 |
|--------------------|-------------|-----------------------|-------------|--------------------|---------|
| | | | | 5% Critical value | -2.8626 |
| | | | | 10% Critical value | -2.5674 |
| Variable | Coefficient | Std-Error | t-statistic | Prob | |
| DOP(-1) | -1.1010 | 0.0312 | -35.2652 | 0.0000 | |
| D(DOP(-1)) | 0.0952 | 0.0277 | 3.4277 | 0.0006 | |
| D(DOP(-2)) | 0.0574 | 0.0236 | 2.4236 | 0.0154 | |
| D(DOP(-3)) | 0.0181 | 0.0190 | 0.9523 | 0.3410 | |
| D(DOP(-4)) | 0.0423 | 0.0134 | 3.1518 | 0.0016 | |
| C | 0.0047 | 0.0078 | 0.6092 | 0.5424 | |
| R-squared | 0.5046 | Mean dependent Var | | 0.0004 | |
| Adjuster R-squared | 0.5042 | S.D dependent Var | | 0.8276 | |
| S.E of regression | 0.5827 | Akaike info criterion | | 1.7590 | |
| Sum squared | 1883.187 | Schwarz criterion | | 1.7661 | |
| Log likelihood | -4876.186 | F- statistic | | 1129.818 | |
| Durbin-Watson | 1.9990 | Prob(F- statistic) | | 0.0000 | |

As it can be seen in table 4, data are stationary in (1%), (5%) and (10%) levels. In the next step, we investigate the number of optimum lag for MA and AR.

Table 5. Results of ARIMA model implementation

| Variable | Coefficient | Std-Error | t-statistic | Prob |
|--------------------|-------------|-----------------------|-------------|--------|
| C | -0.0155 | 0.0147 | -1.0562 | 0.2909 |
| @Trend | 7.18E | 4.59E | 1.5640 | 0.1179 |
| AR(1) | 0.228 | 0.037 | 6.1785 | 0.0000 |
| AR(2) | 0.1409 | 0.0420 | 3.3507 | 0.0008 |
| AR(3) | 0.0661 | 0.0333 | 1.9849 | 0.0472 |
| AR(4) | 0.0011 | 0.0318 | 0.0375 | 0.79 |
| AR(5) | 0.1859 | 0.2995 | 6.209 | 0.0000 |
| AR(6) | -0.1191 | 0.0360 | -3.3035 | 0.0010 |
| AR(7) | 0.5827 | 0.0291 | 20.008 | 0.0000 |
| AR(8) | -0.3889 | 0.0358 | -10.863 | 0.0000 |
| AR(9) | 0.4902 | 0.0377 | -12.985 | 0.0000 |
| MA(1) | -0.237 | 0.0354 | -6.715 | 0.0000 |
| MA(2) | -0.1701 | 0.0415 | -4.098 | 0.0000 |
| MA(3) | -0.092 | 0.0356 | -2.596 | 0.0095 |
| MA(4) | 0.0274 | 0.0338 | 0.8116 | 0.4170 |
| MA(5) | -0.224 | 0.0314 | -7.131 | 0.0000 |
| MA(6) | 0.1063 | 0.3038 | 2.761 | 0.0058 |
| MA(7) | -0.569 | 0.0325 | -17.512 | 0.0000 |
| MA(8) | 0.3701 | 0.033 | 11.035 | 0.0000 |
| MA(9) | 0.5094 | 0.0376 | 13.538 | 0.0000 |
| MA(10) | 0.0326 | 0.0153 | 2.132 | 0.0330 |
| R-squared | 0.0189 | Mean dependent Var | | 0.0043 |
| Adjuster R-squared | 0.0153 | S.D dependent Var | | 0.5841 |
| S.E of regression | 0.5796 | Akaike info criterion | | 1.751 |
| Sum squared | 1856.807 | Schwarz criterion | | 1.7761 |
| Log likelihood | -4835.536 | F- statistic | | 5.333 |
| Durbin-Watson | 1.9974 | Prob(F- statistic) | | 0.0000 |

As it can be seen in table 5, Durbin-Watson index is also in good condition.

Modeling by VAR

Investigation of stationary of time series variables is the first step in estimation of this model. if the variables were stationary, there will be no problem. Otherwise, their correlation order must be specified. This is specified by means of augmentedDickey-Fullertest. If Dickey-Fuller statistic is less than the calculated value, the variables are stationary or correlated with zero degree $L(O)$. If variables become zero with a subtraction, they are of degree one $I(1)$. In the latter case, the next step will be autocorrelation test using Johansson test.

Johanssonintroduced two tests called Trace test and maximum Eigen values test in order to determine cointegratedvectors (r) within the framework of vector error correction model. These two tests are conducted based upon the highest lag length(k) obtained from AIC and SIC criteria. If the calculated statistic of trace test and maximum Eigen values test is smaller than critical values of the tests in tables of distribution of and λ_{max} statistics which were calculated by Hall et al (1989), H_0 (presence of r co integrated vectors).

In the present research, in order to identify the best hypothesis of Johansson’s cointegration test, this test was conducted considering different assumptions of the trend in data in terms of presence or absence of intercept and linear and non-linear trend. Finally, presence of intercept and absence of a specified trend was verified.

Table 6. determination of the number of cointegrated vectors test (stationary test)

| Critical values | | Test statistic | hypothesis | |
|-----------------|-------|----------------|------------|-------|
| %5 | %1 | | H1 | H0 |
| 28.58 | 33.73 | 55.55 | r = 1 | r = 0 |
| 22.30 | 27.06 | 20.746 | r = 2 | r ≤ 1 |
| 15.89 | 20.16 | 12.468 | r = 3 | r ≤ 2 |
| 9.16 | 12.76 | 10.653 | r = 4 | r ≤ 3 |

Conclusion and recommendations

Now we compare prediction values of the three models with real values of production within the framework of time domain of the research.

| ANN | ARIMA | VAR | Real value | year |
|----------|----------|----------|------------|------|
| 10397.35 | 11541.21 | 10653.05 | 11534.22 | 2002 |
| 12374.39 | 11597.56 | 10981.89 | 12387.73 | 2003 |
| 11414.99 | 11658.73 | 11128.22 | 13326.75 | 2004 |
| 10795.75 | 11733.35 | 10832.89 | 12872.59 | 2005 |
| 11813.69 | 11826.42 | 11327.72 | 13867.72 | 2006 |
| 12071.06 | 11939.08 | 11924.72 | 14515.95 | 2007 |
| 13356.39 | 12069.22 | 12276.72 | 18160.2 | 2008 |
| 13673.36 | 12212.43 | 12328.05 | 17840.14 | 2009 |
| 13571.9 | 12363.18 | 12737.55 | 16167.07 | 2010 |
| 14487.73 | 12515.89 | 13075.72 | 17261.44 | 2011 |
| 13655.19 | 12810.02 | 12134.22 | 16676.58 | 2012 |

Selection of the best model

In the present research, four indices (MSE, MAPE, MAE, and RMSE) were used.

The following table indicates different indices of measurement of estimation error for the three prediction methods. Considering this table and the above discussions, ANN model is the best method and has the smallest error in comparison with VAR and ARIMA methods.

| Appropriate model | VAR | Neural network | ARIMA | Error measurement index |
|-------------------|----------|----------------|----------|-------------------------|
| ANN | 4127.073 | 3301.54 | 3920.557 | MAE |
| ANN | 0.2436 | 0.1920 | 0.2217 | MAPE |
| ANN | 2192.1 | 1471.8 | 2336.8 | MSE |
| ANN | 4681.9 | 3836.5 | 4729.5 | RMSE |

Therefore, ANN model is the best model for prediction of Iranian high-yield rice production.

Applied recommendations

1. considering the results of the research, it can be said that ANN has a better performance in comparison with the other two models. Officials and decision-makers are advised to use his method for prediction of high-yield rice production. Further, ANN model can be used for prediction of other economic variables like products price. Therefore, we advise managers and decision-makers of Ministry of Industry, Mining and Commerce to use this method for short-term prediction of basic products prices so that they can plan for price variations in market.
2. We advise to those researchers and experts who intend to predict rice production level to use other instruments besides ANN method because having experience in production and consumption market and then use experiences for better predictions.

REFERENCES

1. Esfahaniyan, M & Amin Naseri, M. (2008). Presenting a neural network model for short-term prediction of crude oil price. International journal of engineering sciences, Iranian Science and Technology University, number 1, pp: 27-35.

2. Hafiznia, M. (2011). An introduction to research methodology in humanities. Tehran: SAMT publications.
3. Hoseini, S, KordKalayee, M & Kalashemi, M (2010). Investigation of competitiveness ability of Iranian rice: long-size grain form in Guilan Province, journal of agricultural economy studies, vol 2, number, 3, pp: 91-104.
4. KhaleghiTabar, S & Mashayekhi. (2011), comparative investigation of bankruptcy models of Stock Market Companies using neural network optimization algorithms. Master degree thesis, Tehran University, 7.
5. Shayegan, M, Mojmamadi, H & Mousavi, S. N. (2005). Prediction of rice and maize imports using artificial neural networks method, quarterly of economic policies and studies, number 89, 44.
6. Menhaj, M. B. (2008), fundamentals of neural networks. Tehran: Amirkabir University of Technology press.
7. Nouri, investigation of rice supportive policies in Iran. Agriculture and development economy. 13th year, 1 & 2.
8. Niroumand, H. (2004). Time series. Tehran: Payam-e-Nour University.