Cumhuriyet Üniversitesi Fen Fakültesi
Fen Bilimleri Dergisi (CFD), Cilt:36, No: 3 Ozel Sayı (2015)
ISSN: 1300-1949

Cumhuriyet University Faculty of Science
Science Journal (CSJ), Vol. 36, No: 3 Special Issue (2015)
ISSN: 1300-1949

Determining an optimal pattern for the Prediction of potato and onion consumption in Iran

Ahmad FATTAHI ARDAKNI

Assistance Professor of Agricultural economics of Ardakan University, Iran

Received: 01.02.2015; Accepted: 06.06.2015

Abstract. Given the importance of agriculture in the economy and the increasing role of potato and onion in the diet, this paper predicts the consumption of these products using a statistical model. Therefore, based on single and multi-variable models, data from 1970-2013 was applied to predict potato consumption using double exponential adjustment model and ARIMA (2, 1, 3) model was applied to predict onion consumption in 2015-2018. The results not only show a suitable prediction based on error criteria for the above mentioned models, but also reveals the fact that the application of an annual series using appropriate and valid models can provide acceptable results. The comparison of these Predictions with actual values is a sign for the high prediction power of these models and confirm this claim. According to our study, it is anticipated that per capita consumption of potato and onion will reach to 31.4 and 18.9 kg at the end of the year 2018.

Keywords: Prediction, Dual power adjustment, ARIMA, VAR، ARDL, Potato and onion

1. INTRODUCTION

Investigating the pattern of food consumption and anticipated consumption for various products is of great importance since it contains valuable information to the country's food policy. Comparison of food consumption at different times in a community makes it possible to examine the changes in the use of these products. In addition, the impact of different policies on the food consumption can be seen including changes in production, imports and subsidies. In this context, predicting the use of various products, especially household goods of important base can help planners and policy makers in various aspects to better understand the situation and achieve the desired status. Therefore, it is essential that comprehensive research take place in the field of consumer behavior and how to predict the effects of different factors. The importance of addressing the expected consumption of basic goods increases when it is possible to reach more accurate predictions through advanced theories of the consumption of these products. This discussion has led to various methods and models to predict the consumption in recent decades. Among these methods, time series patterns emerged as a powerful tool in predicting in recent years and evolved and are widely used to predict food products consumption and factors affecting.

Prediction is described as the art and science of predicting future events and news, including getting historical information and extending them into the future with the help of a variety of mathematical models. Demand and price are more important for consumers and producers of agricultural products among economic variables, and play the central role in optimizing production, marketing and investment. Prices guide in agricultural production and consumption decisions. So, the producers consider current prices of inputs and the price of future products to produce the desired products and decide about production resource allocation altogether. On the other hand, consumers of agricultural products decide about their consumption, according to current prices and future prices. Farmers predict and control their production and supply

* Corresponding author. *Email address: Ahmad Fattahi Ardakni*

Special Issue: The Second National Conference on Applied Research in Science and Technology

according to price fluctuations in the past few years, and consumption of the products in the future. These fluctuations decrease the income and welfare and increase product waste in the years of increased production. On the other hand, supply fluctuations affect the general population (consumers). Therefore, predicting consumption and cost for products with more volatility seems essential. Among Iranian agricultural products, potatoes and onions are of particular importance regarding the general consumption, and the annual fluctuations in supply and demand caused discontent among farmers and consumers. In this article, regarding the importance of the role of potato and onion in the diet of consumers and their place in the household food costs, the consumption of potatoes and onions were reviewed from 1970 to 2007 (annual series) and prediction has been made for 2011 to 2014.

Monroe (1990) claims that price aims at maximizing the well-being of an economy and the allocation of resources to maximize income. In other words, in a free market economy, production and consumption of the product can be determined by prices. Deaton and Laroque (1992) have used econometric analysis and seasonal variations to examine price fluctuations. Adhikari et al (2003), in a study to predict required water for the birds using data poultry industry Georgia in the years 2001-1971 with model (0, 1, 2) ARIMA forecast water consumption by 2010. They concluded that the lack of economic parameters results in overestimation of the required water. Reviewing various studies in Iran and other countries, no study has been done in this field.

2. MATERIALS AND METHODS

In general, there are different models to predict a time series variable. First, time series data should be analyzed in order to obtain a model to provide effective predictions. To predicting time series and prediction model, forecasting techniques are used, including simple moving average, Simple exponential smoothing, Double exponential smoothing, Triple exponential smoothing (Holt Winters), AR single integrated moving average (ARIMA) and seasonal (SARIMA). In general, given the nature of the series, these methods are used. If series are static, simple and exponential averages and ARIMA are applied and if they follow trends, a double exponential moving average is adjusted. For seasonal fluctuations, triple exponential modulation method is used.

Based on observed behavior of the time series and analyzes its components, the data from the analysis of this data, a repeatable pattern for the future acquires. In general, to fit a time series model based on observations, there are three basic steps which should be taken:

1) Identification of the model: The model must be selected for data.

2) Estimating coefficient of the model: different methods can be used for estimating the coefficients of the model such as Yule- Wakker, M.L .Method, Mean Square Method

3) Verify the model: In this method, after the identification of coefficients, the right fit, will be examined. One the most important ones is examination of the remains.

Due to the use of dual power adjustment, ARIMA, (univariate) and VAR and ARDL (multiple), a brief explanation of each one is provided.

Univariate models

1. Adjusted dual view: This method is similar to adjusted unique view in which trend has been added.

Selecting and with minimum MSE and prediction could be made.

2. ARIMA (p, d, q) Process:

ARIMA for X variable could be written as follows:

Usually and. If ARIMA becomes ARMA. Usually to estimate the patterns, Box and Jenkins (Bj) is used which consists of four stages:

1. Diagnosis: In this stage, actual values of q, d and p are selected from by ACF and PACF. According to these models and ACF and PACF numbers which are the peak and decline, the process of time series can be identified.

2. The estimate stage is after the diagnosis.

3. Diagnostic control: The model's fit is controlled by disturbing element static test.

4. The final model is used to predict.

5. There are different statistics in relation to the validity and predictive power of the model, such as MAD, RMSE which are calculated by the following index:

MAPE

The actual value

The predicted value

The lesser the prediction error, the more ability the model has to predict. Zero for each statistic represents a perfect fit. If the seasonal time series data is observed, which means that the variable is a function of the same month in the previous year, it is necessary that the variables levels come in the current season and in previous months or years before the model is taken. In other words, this model is shown as SARIMA (Q, D, P) and (q, d, p) in which D is Seasonal differences, P is AR degree, seasonal (SAR) and Q is MA seasonal degree.

Multivariate models

1. VAR

VAR, in fact is a linear relationship between the dependent variable and all the variables in the equation system in which the number of interrupts are determined experimentally. The general form of an equation system with n dependent variable (n equation) is as follows:

In This Equation

L**:** Operator interrupt

C: Intercept matrix equations

: Random disturbing elements

They are assumed to have a normal distribution with mean zero random and constant variance of. Also, the elements of the matrix an are defined as:

Where, i represents the number of the equation, j represents the number of variables in the equation, K represents the number of interruptions to the system. In a system of (VAR), if the variable demand is a function of the traps, all variables in the system, the system is called non-structural vector regression. According to Sims (1980), the main issue in this case is to determine the appropriate lag length and determine the parameters of the system. Sometimes limited degree of freedom determines the number of interrupts, but when the number of observations is large, it is necessary to determine the optimum interval.

Clear vector regression model is very simple, with minimal reliance on set theory. In this model, it is sufficient to specific variables in the system (based on the analysis of economic relations) and specified number of interrupts will be determined. Then, a linear relationship between the variables in the system is simply established. In Vector regression models, since the linear correlation between the right variables is likely and lead to severe Multi co linearity, therefore, the criteria of the test function (t) cannot be used for individual coefficients to minimize the model.

In VAR model, the degree of stability and continuous optimization variables should be determined in the mode first. For the static test variables involved in the investigation, the Dickey-Fuller test generalized (ADF) was used to optimize the number of intervals based on (AIC) and (SBC).

In the estimation of VAR, it is necessary to determine the stability degree and continuous optimization variables used in the model. For the static test variables involved in the investigation, the Dickey-Fuller test generalized (ADF) was used to optimize the number of intervals based on Akaike (AIC) and Schwartz (SBC).

The fact that the variables in the VAR model together from the first class (I (1)) refer to the valid results of estimation, forecasting and statistical tests, when the relationship between the convergence model variables exist (Milo and Nell, 2001). So, in the next step, vector pattern convergence test between variables was done based on Johansson test. Determining the number of vectors converging in this way is done using the effect test ( ) and specific maximum test ( ).

2. ARDL

In this type of time-series models, the value of that variable depends on the past and the present values of other independent variables in the related model.

The general form of is as follows:

Dependent variable

Vector of explanatory variables

K: The number of explanatory variables

The optimal number of intervals for each of the explanatory variables

P: The optimal interval of the dependent variable models

Vector categorical variables such as intercept, seasonal variations and trends

The model is estimated using OLS method. At first, maximum intervals are specified and then, the model is estimated for all values p=0**, 1, m and q=0, m**. Then, using a measure of Akaike, Schwartz or Hanan-Quinn (HQC) the optimal orders are selected. This approach is in contrast to VAR and has no obligation to be static of all variables of degree one. Also, based on the estimated coefficients for the variables of interest ARDL model, long-term coefficients can be also achieved. The data used in this study consisted of per capita consumption of these products (the Order of households / families) between 1970-2013 (annual) that is extracted from the Central Bank. It should be noted that for 1980-83 years due to lack of implementation of the project, Data was obtained by interpolation of other years.

3. RESULTS AND DISCUSSION

In this part, anticipation of onions and potatoes consumption **is** discussed separately.

Onion

In order to provide accurate results in time series variables, static variables should be ensured. First, annual fluctuations in consumption during this period was familiar with time series curves (Run Sequence). ACF also shows annual changes and annual differences were calculated for static. According to rapid reduction of ACF, static was determined. Then, P and q primary values of the confidence interval was determined by displacement interrupts. With the Over Fitting, valid models were specified and the appropriate model was determined by AIC and SBC statistics. In order to identify the right type of the model, Box- Pierce test and normal White noise components interfere with ACF, histogram curve, probability of normal and so on were used and the assumption of independence and stochastic series was established in model (3, 1 and 2) ARIMA. Data survey showed annual fluctuations; therefore, dual power adjustment method was used in this method.VAR model was also applied. First, Granger causality test is applied to assess causality performing expenditure price index and per capita income. The results of this test indicate the impact of these two factors on per capita consumption. Dickey-Fuller test was done and the results showed that consumption logarithm (C), Per capita (Kg), income (I) , per capita (Rials) and price index(P) were I(1). Then, the optimal interval (pause) was estimated using Akaike criterion (AIC) and Schwartz (SBC) under the following model.

LC=7/2 **LC (**-1) + 0/1 LI(-1) - 0/3LP(-1) +8/5

(t) 6/3 2/2 0/5 3/2

R=0/61 F=18/5

Johansen test for convergence was studied. The results indicate that in the pattern of onions consumption, the quantity of statistics in both tests are less than the critical values provided by Scarlett Johansson in 95%. Therefore, in both cases, the existence of a balanced relationship between variables will be accepted. Assessment with ARDL is another multivariate model. The best model with variables ARDL (1,0,1) were estimated as follows. ARDL pattern using different explanatory variables and determining the optimal lag length each, according to Schwartz criterion (SBC) was fitted.

Table 1. The results ARDL (1,0,1) with continuous dependent variable per capita consumption of onions and other variables.

LP (-1)	LP	LY	LC (-1)	Const	
-0/28	-0/31	-0/001	0/77	0/82	LC
2/72/5	0/1	7/04	1/2	t-stat	
R-Squared=0/74					

The cusum, cusumq tests were done to stability parameters which show the stability coefficients.

To compare the models, prediction for the years 2009-2013 of each model was done. Then, FMSE was compared to select the most suitable model. According to Table 2, ARIMA has less forecast errors.

Table 2. Comparison of Model errors.

VAR	ARDL	Dual power adjustment	ARIMA	Model
1.31/02	0.59	0.42	FMSE	

Given the choice of the final model, consumption prediction is shown in Table 3.

Table 3. Actual and forecast values predicted consumption of 2009-2013 years and 2015-2018 (Unit: kg).

2018	2017	2016	2015	2013	2012	2011	2010	2009	Year
Value									
- -	-	-	15.17	9.16	3.17	9.16	6.16	actual	
99.18	34.19	63.17	33.17	18.17	9.18	2.17	2.16	5.16	Forcast

According to the table, per capita consumption in 2015 is 17.33 kg and will be 18.99 kg by the end of the year 2018.

Potato

To predict consumption, using existing data (annual) time series curves are drawn and the data trend is specified, and ACF showed non-static data. Then, with one simple difference of static data and using the steps listed in predicting in onion, the best Double exponential adjustment model was selected. Since the process for predicting time series, ARIMA method is used, the method of (VAR) with variable consumption per capita (kg) per capita income (thousand dollars) and retail price (with optimal interrupt **1**) in log is calculated as follows.

LC=0/83 LC (-1) +0/03 LI (-1) - 0/03 LP (-1) +0/48

(T) 8/4 2/7 0/7 1/4

R=0/74 F=27

Johansen test for convergence was studied. The results indicate that in the pattern of onions consumption, the quantity of statistics in both tests is less than the critical values provided by Scarlett Johansson in 95%. Therefore, in both cases, the existence of a balanced relationship between variables will be accepted. ARDL pattern using different explanatory variables and determining the optimal lag length each, according to Schwartz criterion (SBC) was fit. Therefore, the pattern of ARDL (1,0,0) for potatoes consumption (both in logarithmic) as a Model was selected.

Table 4. The results ARDL (0,0,1) with continuous dependent variable per capita consumption of potato and other variables.

LP	LY	LC (-1)	Const	
06/0-	06/0-	8/0	8/0	LC
7/08/2	7/7	8/1	t-stat	
R-Squared=0/75				

The cusum, cusumq test was done to stability parameters which showed the stability of coefficients.

Table 5. The comparison of FMSE models.

VAR	ARDL	Adjusted dual view	ARIMA (2,1,2)	Model
043/1	1/04	03/1	54/2	FMSE

The table shows that FMSE is lower in the Dual power adjustment than other methods. Therefore, selecting this method predicting the consumption for the years 2015-2018 will be as following.

Table 6. The actual and forecast values for the years 2009-2013 and 2015-2018.

2018	2017	2016	2015	2013	2012	2011	2010	2009	Year
Value									
- -	-	-	6.29	9.28	3.27	2.26	8.26	actual	
4.31	31	5.30	1.30	28	6.27	2.27	7.26	4.26	forcast

According to the table, per capita demand in 2015 is 30.1 kg and will be 31.4 by the year 2018.

REFERENCES

[1] Adhikari.M. at all (2003). Water Demand Forecasting for Poultry Production Selected Paper Prepared for AAEA

[2] Campo, I.S. and J.C. Beghin.(2005). Dairy Food Consumption, Production, and Policy in Japan. Working Paper 05-WP 401. Center for Agricultural and Rural Development

[3] Chorng.s.(2005). Model identification of arima family using genetic algorithms. Applied mathematics and computation pp 164

[4] -Deaton, A. S. & G. Laroque. 1992. On the behavior of commodity prices, Review of Economic Studies, 59:1-23

[5] Halicioglu, F. (2004). An ARDL Model of International Tourist Flows to Turkey. Global Business and Economics Review.

[6] Houston. & at all.(2003). Forecasting Broiler Water Demand Econometric & Times Series Analysis Selected Paper Prepared For Western Ag. Econ

[7] Joy.H. and Barry. r (2001). Principles of operations management. Prentice Hall Inc. new jersey

[8] Levin.R. (1989). Quantitative approaches to management. Mac grow hill

[9] Monroe . k.(1990) : Pricing' making profitable decision Mac grow hill international ' Editions

[10] Mello, M.D. and K. S. Nell.(2001). The Forecasting Ability of a Co integrated VAR Demand System With Endogenous vs. Exogenous Expenditure Variable: An Application to the UK Imports of Tourism From Neighboring Countries"

[11] Sabur.r.(1993). Analysis of rice price of in immensity town market. Pattern and forecasting Bangladesh. G . of. Ag. Eco (16)

[12] Soares, L and M.C. Medeiros. (1999). Modeling and forecasting short-term electric load demand: a two step methodology.

[13] Roger.k.(1984). Forecasting future price trends in the u.s. fresh and processed potato market (g.of food distribution research)

[14] Volkan.s.(2006). Forecasting production of fossil fuel sources in turkey using a comparative regression and arima model. Energy policy 3 4

[15] winter.a.(1996). Expectations' supply response and marketing boards: an example from Kenya: Ag. Eco (14).