Journal of Industrial Strategic Management

Time Series Models to Predict the Monthly and Annual Consumption of Natural Gas in Iran

Arash Farrokhi[a] Reza Hassanzadeh[b*]

a Department of Industrial Engineering, University College of Ayandegan, Tonekabon, Iran
b Department of Industrial Engineering, University College of Rouzbahan, Sari,, Iran, Iran

CHRONICLE

Article history:
Received: 08/09/2017
Received in revised 03/11/2018
Accepted: 03/15/2018

Keywords:
** Forecast Gas*
** Consumption*
** ARIMA*
** SARIMA*

Abstract

Considering the fact that natural gas is a widely used energy source, the prediction of its consumption can be useful (Derek LAM, 2013). As Iran has one of the largest gas reserves in the world, its consumption in the country can affect the worldwide price of gas, Therefore, the current research is useful both from economic and environmental point of view.
The goal of the study is to select the best model for the prediction of gas consumption. To achieve the goal time series analysis are used. The findings indicate that ARIMA (0, 1, 0) is the best model for the prediction of annual gas consumption, while SARIMA (1, 0, 0) (1, 1, 0) for the prediction of monthly gas consumption.

Introduction

Energy plays an important role in maintaining social and economic development of a country. Natural gas, as source of energy, is one of the cleanest burning alternative fuels. About 17 percent of natural gas reserves in the world belongs to Iran, which makes it the second largest gas reserve in the world.

Normally, gas consumers can be divided into three groups: local suppliers, industrial and commercial consumers. The demand of these groups is different. Therefore, gas companies should make sure that the right amount of pressure is delivered to the customer whose consumption is predicted. There are many prediction methods that can be used to predict the gas consumption with high accuracy. Some of those methods are presented in Table 1 (Aras, H. and Aras, N. 2004).

Table 1: Methods of prediction of gas consumption

Prediction method	Researcher
Time series model to predict monthly consumption	**Aras and Aras (2004)**
Fuzzy neural network	**Viet and Mandziuk, (2003)**
SUPPORT VECTOR Method	**Liu et al., (2004)**
SARIMA, ARIMA models	**Ediger Volkan et al., (2007)**

[*]Corresponding author Tel: 09113955919 rhz_1974@yahoo.com

Kamari and Mohammad (2017) proposed reliable models for the prediction of cumulative gas production developed on the basis of powerful artificial intelligence techniques such as the artificial neural network (ANN), least square support vector machine (LSSVM), adaptive neuro-fuzzy inference system (ANFIS), and decision tree (DT) method. Lingxue Zhu (2017) proposed a novel end-to-end Bayesian deep model that provides time series prediction along with uncertainty estimation. He provided detailed experiments of the proposed solution on completed trips data, and successfully applied it to large-scale time series anomaly detection at Uber. Prema and UmaRao (2015) proposed time series models for short-term prediction of solar irradiance. The predictions are done for 1 day ahead using different time-series models. For each model, these predicted values are compared with the actual values for the next day and graphs are plotted. Basic time-series models such as moving average and exponential smoothing were tested.

The goal of this study is to develop a model for the prediction of the daily and monthly natural gas consumption in Iran using time series method.

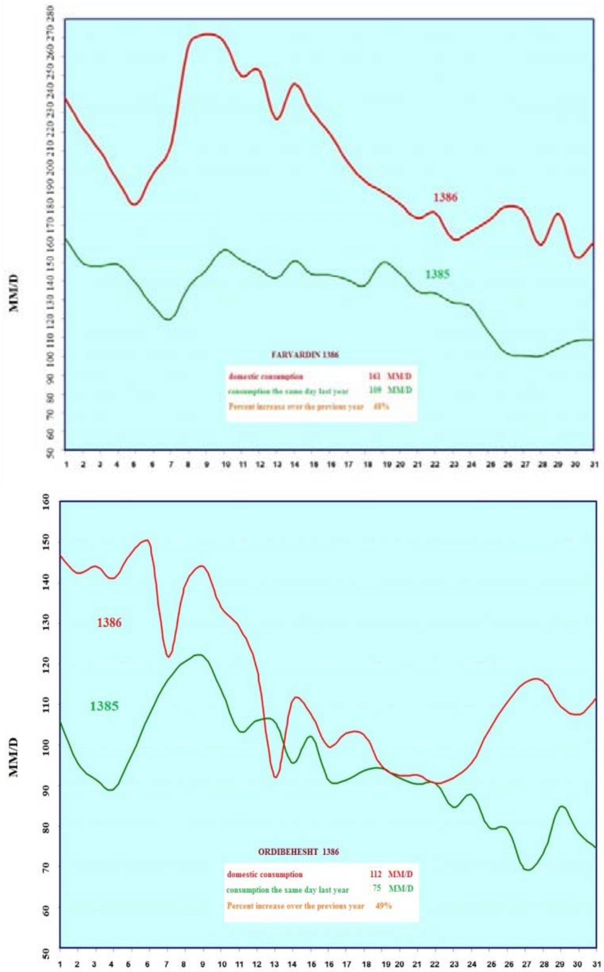

Figure 1: Monthly forecast for 2 months (Farvardin, Ordibehesht) in 1386

Introduction to Time Series

A time series is a set of data points listed in time order. A time series is a sequence taken at successive equally spaced points in time. Thus, it is a sequence of discrete-time data. Time series analysis comprises methods for analyzing time series data toextract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values. While other types of analysis are often employed to test if the current values of one or more independent time series affect the current value of another time series, this type of analysis of time series is focuses on comparing values of a single time series or multiple dependent time series at different points in time (Brown et al., 2005).

Box Jenkins Method

In time series analysis, the Box–Jenkins method, named after the statisticians George Box and Gwilym Jenkins, applies autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA) models to find the best fit of a time-series model to past values of a time series. The model is presented in Figure 2 (Hadiza Yakubo Bako, 2013).

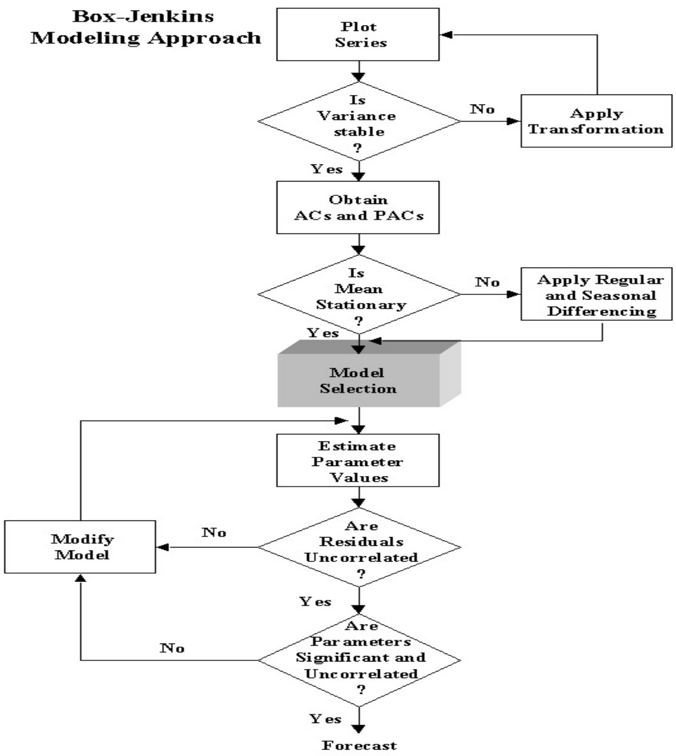

Figure 2: Box-Jenkins model flowchart

Non-seasonal ARIMA models are generally denoted ARIMA(p,d,q) where parameters p, d, and q are non-negative integers, p is the order (number of time lags) of the autoregressive model, d is the degree of differencing (the number of times the data have had past values subtracted), and q is the order of the moving-average model. Seasonal ARIMA models are usually denoted ARIMA(p,d,q)(P,D,Q)m, where m refers to the number of periods in each season, and the uppercase P,D,Q refer to the autoregressive, differencing, and moving average terms for the seasonal part of the ARIMA model. (Chatfield, 1996 and Sungki Kim et al., 2017).

The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged values. The formula of AR is expressed by the following formula:

$$Y_t = \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \varepsilon_t$$

The MA part indicates that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The formula is as follows:

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

The I (for "integrated") indicates that the data values have been replaced with the difference between their values and the previous values. The purpose of each of these features is to make the model fit the data as well as possible.

Autocorrelation function (ACF):

Autocorrelation, also known as serial correlation, is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them. The analysis of autocorrelation is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is often used in signal processing for analyzing functions or series of values, such as time domain signals.

Unit root processes, trend stationary processes, autoregressive processes, and moving average processes are specific forms of processes with autocorrelation.

Partial autocorrelation function (PACF):
In time series analysis, the partial autocorrelation function (PACF) gives the partial correlation of a time series with its own lagged values, controlling for the values of the time series at all shorter lags. It contrasts with the autocorrelation function, which does not control for other lags.
This function plays an important role in data analysis aimed at identifying the extent of the lag in an autoregressive model. The use of this function was introduced as part of the Box–Jenkins approach to time series modeling, where by plotting the partial autocorrelative functions one could determine the appropriate lags p in an AR (p) model or in an extended ARIMA (p,d,q) model. Partial autocorrelation (total number of remaining data becomes zero) specifies the order of MA (moving averages) and shows q stage. If the ACF had a progressive pattern, it was defined as a non-static. In this case it needs change which was made in Figure 3.

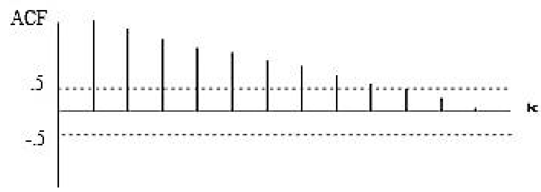

Figure 3: Non-static

If the ACF and PACF go down to zero at the beginning of the function, it shows autoregressive model of order p, AR (p). P depends on the number of significant breaks in Figure 4.

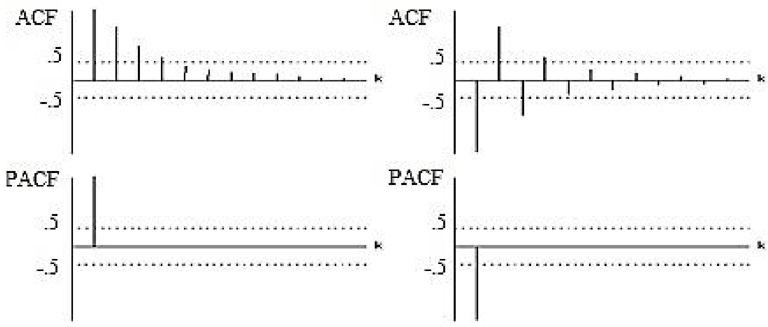

Figure 4: AR

If there is a trend toward zero and PACF and ACF have significant interruptions at the beginning of the function, the pattern has an order q. MA (q) in Figure 5.

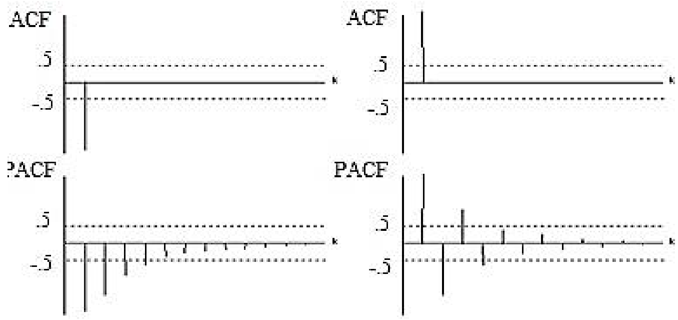

Figure 5: MA

If both ACF and PACF decline to zero, a complex pattern of ARMA (p, q) is formed like in Figure 6.

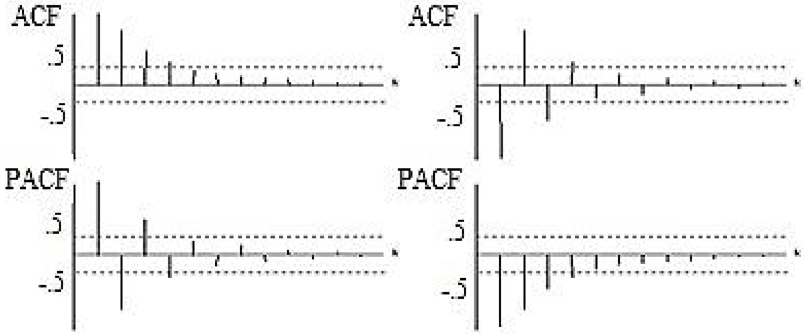

Figure 6: ARMA (p, q)

In non-static model, ARMA (p, q) changes to the ARIMA (p, d, q).

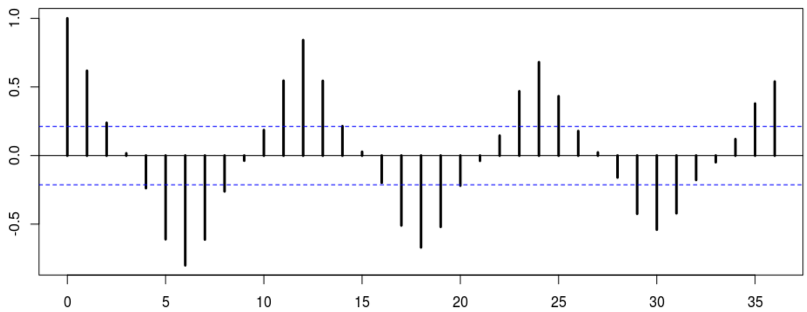

Figure 7: ARIMA (p, d, and q)

Predicting Monthly Gas Consumption in Iran

The following figure presents the data of 9-year monthly consumption of natural gas in Iran. The data is provided by the natural Iranian gas company and these data will be used in our analysis.

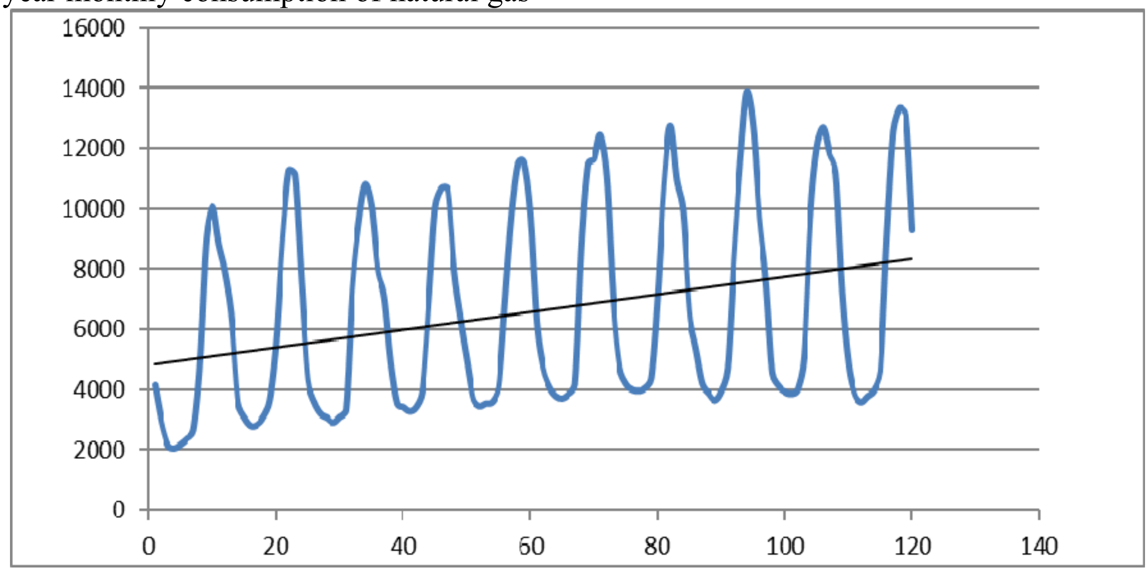

Figure 8: Monthly consumption of natural gas in Iran from 2006 to 2015

We employed software to predict the gas consumption for 2015 based on the data 2006-2014. The methods we used for the prediction and corresponding results are illustrated in Figures 9-12.

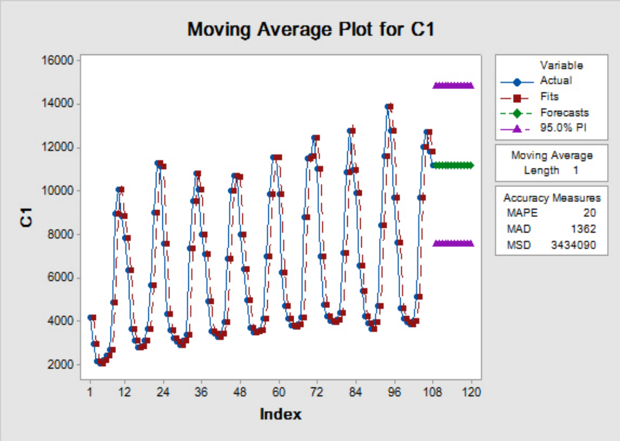

Figure 9: The results predicted by MA

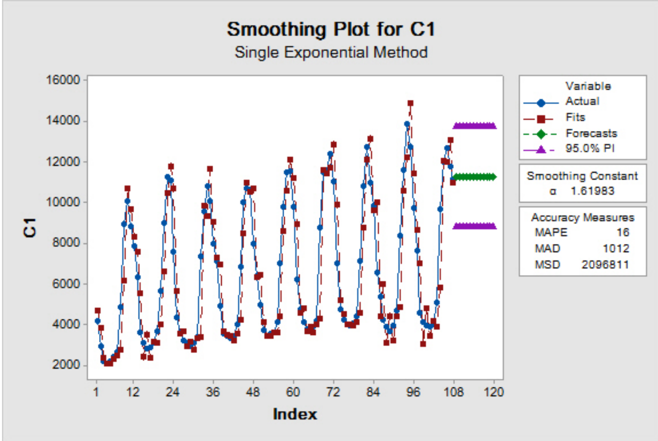

Figure 10: The results predicted by SE

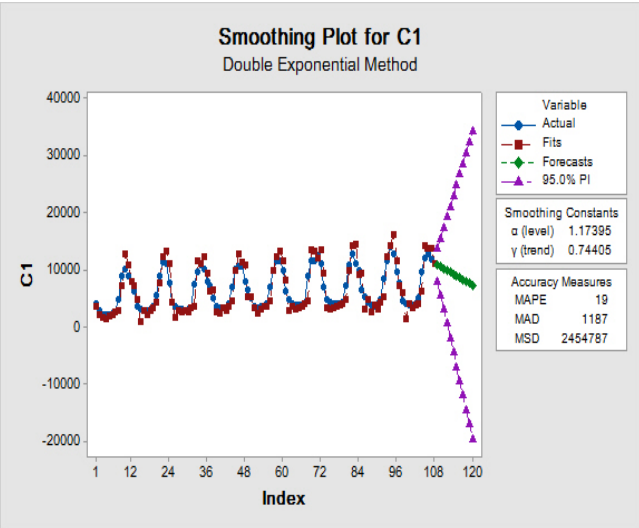

Figure 11: The results predicted by DES

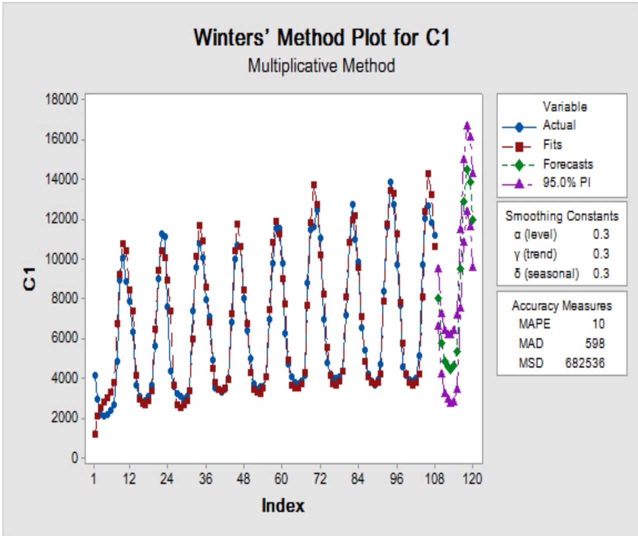

Figure 12: The results predicted by WINTERS METHOD

WINTERS hybrid model with weight 0.3, has fewer errors and therefore it has been considered as the best among the mentioned methods. The results are listed in Table 2.

Further, we employed Box-Jenkins SARIMA for predicting monthly gas consumption in Iran in 2015. According to SPSS, SARIMA time series models (1, 0, 0), (1, 1, 0) are estimated to static models where the parameter estimate of 0.57 AR (1) to 0. 4373 and AR SEAEANAL was -0.594. Predictions by SARIMA model are in the Table 3. Model SARIMA, is less than MAD (mean absolute deviation). Hence, SARIMA model is selected as the model for predicting the monthly consumption of gas.

Table 2: Monthly gas consumption predicted by WINTERS method

winter multicative (0.3-0.3-0.3)					
			2015		
	real	Prediction	Upper bound	Lower bound	error
March	7263	8035	9499	6571	-772
April	4844	5728	7244	4211	-884
May	3825	4826	6403	3249	-1002
June	3579	4553	6199	2908	-975
July	3776	4440	6159	2720	-664
August	3979	4614	6414	2815	-636
September	4625	5312	7197	3428	-688
October	8997	9495	11469	7522	-499
November	12540	12906	14972	10839	-365
December	13329	14529	16692	12366	-1200
January	13099	13877	16138	11615	-778
February	9288	11934	14298	9571	-2646
					-926

Table 3: Monthly gas consumption predicted by the model SARIMA

SARIMA (1,0,0)(1,1,0)					
			2015		
	real	prediction	Upper bound	Lower bound	error
March	7263	7794.76	9426.55	6383.65	-532
April	4844	5579.36	6862.43	4484.43	-736
May	3825	4603.41	5679.44	3687.38	-779
June	3579	4299.84	5308	3441.98	-721
July	3776	4108.31	5072.12	3288.25	-332
August	3979	4361.76	5385.14	3491.03	-383
September	4625	5339.96	6592.88	4273.93	-715
October	8997	9788.91	12085.7	7834.72	-792
November	12540	12948.19	15986.26	10363.31	-408
December	13329	14700	18200	11800	-1371
January	13099	13607.07	16799.73	10890.65	-508
February	9288	11310.01	13963.71	9052.16	-2022
					-775

Predicting Annual **Gas Consumption**

As it can be seen in the figure below the demand for gas has increased from 1999 to 2015.

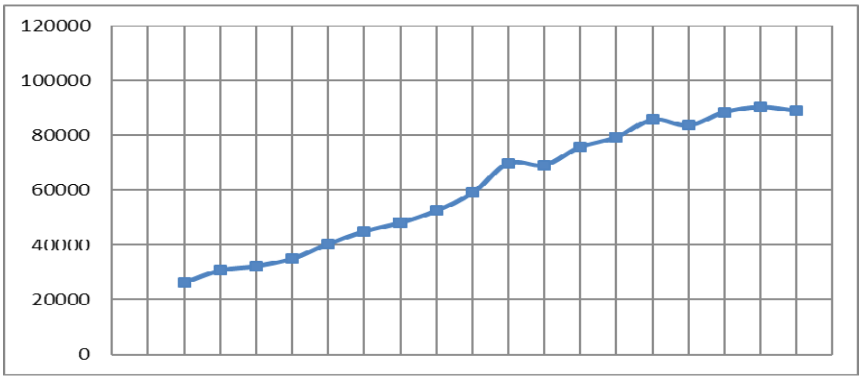

Figure 13: Iran's annual gas consumption from 1999 to 2015

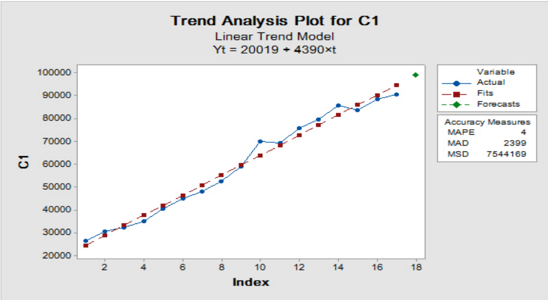

Figure 15: Consumption with the Grade 2

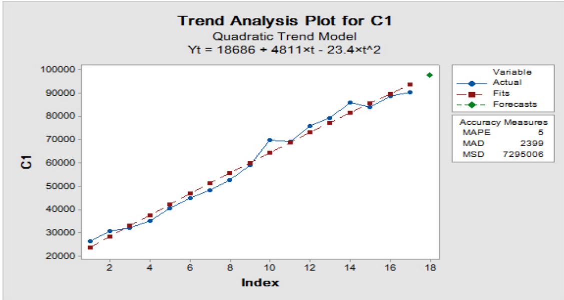

Figure 14: Linear model

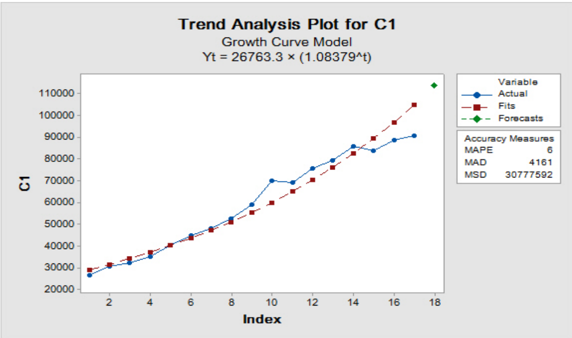

Figure 17: SES model

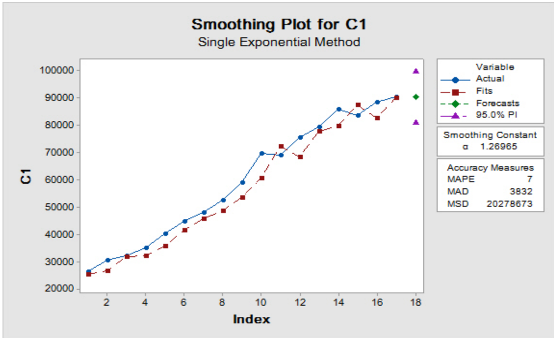

Figure 16: Exponential model

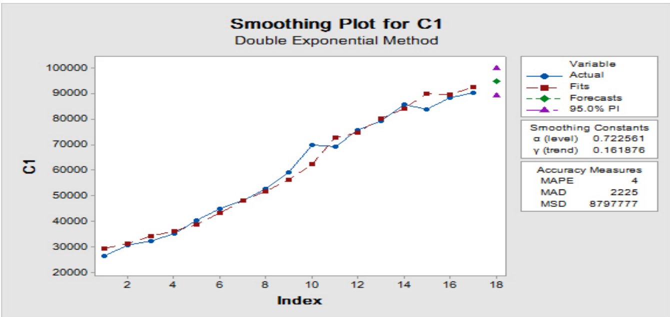

Figure 18: DES model

Analysis can be carried out using different methods. Most common method is linear and quadratic. In this paper, the development methods are used to pave double annual forecast using MINITAB software. The methods and the corresponding results are presented in Figures 14-18 and the wrong data are listed in Table 4.

The mean absolute percentage error (MAPE) and mean absolute deviation (MAD) are less than other data. To get a better result for the constant level (constant values for the smooth development double between 0 and 1 respectively) values of 0.2:0.4 ratio and MAD will be considered.

As Table 4 lists, the results show that the parameter values (0.05: 0.65) have a lower MAD and the forecasted amount is equal to 95,765.3 in 2015.

Table 4: Fixed parameters for the development process and double smooth development

Alfa	landau	MAD
0.65	0.05	2093
0.7	0.1	2117
0.7	0.2	2159
0.6	0.2	2166
0.6	0.4	2230
0.6	0.6	2350

Forecasting Annual Gas Consumption with ARIMA

According to the Figure 13, gas consumption in Iran has increased yearly. This data is dynamically and this type of series should be fixed in accordance with the BOX-JENKINS model. Thus, the new series ($Y_1,\ldots,Y_{t-1}$) are obtained from the previous ($X_1,\ldots,X_{t-1}$)

by the following formula:

$$Y_t = X_{t+1} - X_t$$

In some cases, the second-order difference for the removal process is also used in accordance with the following formula:

$$Y_t^2 = X_{t+2} - 2X_{t+1} + X_t$$

First we kill time in terms of Y_t graph form (19). As you can see, the chart does not have a trend line. ACF and PACF are used to draw the graphs in Figures 19-21 (Chatfield, 1996).

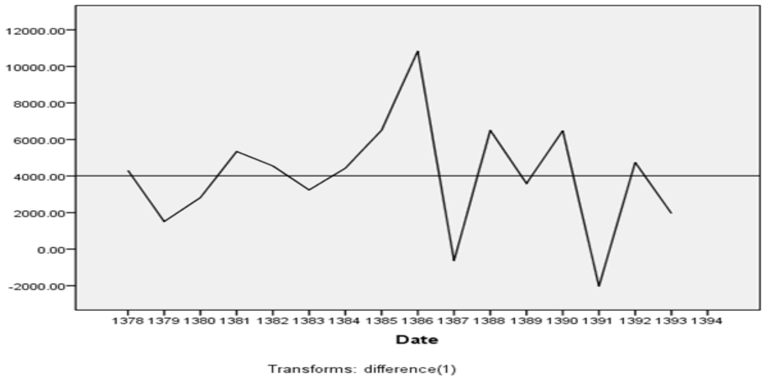

Figure 19: The series after the first difference of annual gas consumption

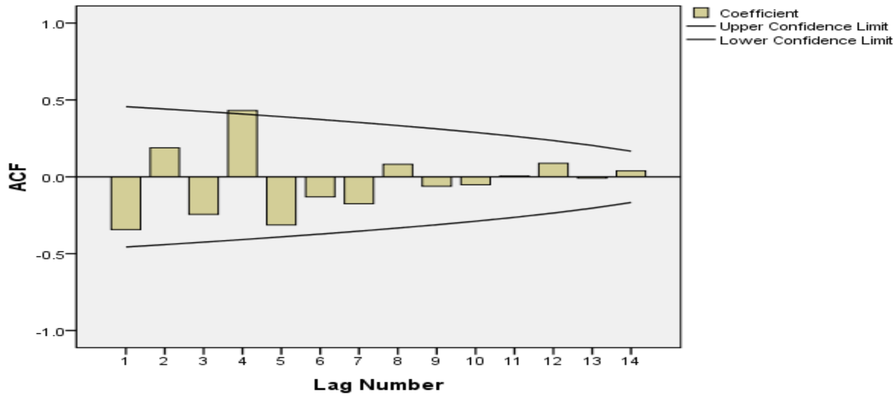

Figure 20: ACF for the first difference for the annual consumption of gas

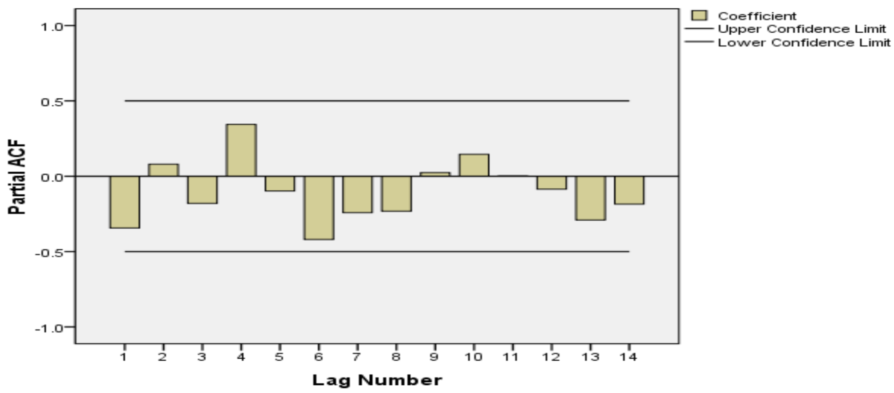

Figure 21: PACF first annual consumption of gas

Further, in this model ACF coefficients are detected (Suhartono, 2010). Heights in the functions (PACF, ACF) command to determine the model ... but we see a graph is not the highest point. The ARIMA (0, 1, and 0) model predicted annual gas consumption in Iran. The ARIMA (0, 1, and 0) is double compared with flat growth (Table 5). In sum, it will have fewer errors. The final model is selected to predict Iran's annual gas consumption.

Table 5:Ccompares the ARIMA (1, 2, 1) with a double smooth development

2015	real	prediction	Upper bound	Lower bound	MAPE
DES(٠.6:.05)	89141	95765.3	102300	8٩940	4
ARIMA(0,1,0)	89141	94500	101000	88000	3.55

Conclusion

Considering the importance of natural gas in everyday life, the prediction of gas consumption can be useful (Derek LAM, 2013). Particularaly, that regional and world gas prices can be influenced by the consumption of gas in Iran. According to the findings, we can predict gas consumption in Iran using the models offered by this study. The results indicate that ARIMA (0, 1, 0) is the best model for the prediction of annual gas consumption, while SARIMA (1, 0, 0) (1, 1, 0) for the prediction of monthly gas consumption.

Reference

1. Akkurt, M., Demirel, O.F. and Zaim, S. (2010). Forecasting Turkey's natural gas consumption by using time series methods. *European Journal of Economic and Political Studies*, 3(2), 1–21.
2. Al-Fattah, S.M. (2006). Time series modeling for U.S. natural gas forecasting. *E Journal of Petroleum Management and Economics Petroleum Journals Online*, 1(1), 1-17.
3. Aras, H. and Aras, N. (2004). Forecasting Residential Natural Gas Demand, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 26(5), 463-472.
4. Baets, S. and Harvey, N. (2018). Forecasting from time series subject to sporadic perturbations: Effectiveness of different types of forecasting support. International Journal of Forecasting, 34(2), 163-180.
5. Brown, R. H., Marx, B. M., and Corliss, G. F. (2005). Mathematical Models for Gas Forecasting. Gas Day Report, No.129.
6. Chatfield, C. (1996). The Analyses of Time Series: An Introduction. Fifth Edition. Chapman & Hall.
7. Derek LAM, C., (2013). Time Series Modeling of Monthly WTI Crude Oil Returns, Lincoln College University of Oxford.
8. Ediger Volkan S., and Sertac, A. (2007). ARIMA Forecasting of Primary Energy Demand by Fuel in Turkey. Energy Policy, 35(3), 1701-1708.
9. Hadiza, Y. B. (2013). Predictive modeling of pelagic fish catch in Malaysia using seasonal ARIMA models, Agriculture. *Forestry and Fisheries*, 2(3), 136-140.
10. Kamari, A., & Mohammad, H. (2017). Decline curve based models for predicting natural gas well performance. Petroleum, 3(2), 242-248.
11. Lingxue, Z. (2017). Deep and Confident Prediction for Time Series at Uber, IEEE International Conference on Data Mining Workshops (ICDMW).
12. Liu, H., Liu , D., Liang, Y. M. and Zheng, G. (2004). Research on natural gas load forecasting based on least squares support vector machine. Proceedings of 2004 International Conference on Machine Learning and Cybernetics. Shanghai, China.
13. Peter, J. Brockwell, & Richard A. D. (2016). Introduction to Time Series and Forecasting. Second Edition, Springer International Publishing Switzerland.
14. Primož, P. (2007). Forecasting risks of natural gas consumption in Slovenia, Faculty of Mechanical Engineering, Laboratory of Synergetics, Aškerčeva 6, SI–1000 Ljubljana, Slovenia.
15. Prema, V. K., and Uma, R. (2015). Development of statistical time series models for solar power prediction. Renewable Energy, 83, 100-109.
16. Suhartono, C. (2010). Time Series Forecasting by using Seasonal Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive Model. Journal of Mathematics and Statistics, 7 (1), 20-27.
17. Sungki, K., Wonil, K., and Hyoon, N. (2017). Statistical model for forecasting uranium prices to estimate the nuclear fuel cycle cost. *Nuclear Engineering and Technology*, 49(5), 1063-107.
18. Viet, N. and Mańdziuk, J. (2003). Prediction of natural gas consumption with feed-forward and fuzzy neural networks. Artificial Neural Nets and Genetic Algorithms, 1, 107-114.
19. Wong-Parodi, G., Lekov, A., and Dale, L. (2005). Natural Gas Prices Forecast Comparison--AEO vs. Natural Gas Markets