Journal of Optimization in Industrial Engineering, Vol.11, Issue 2, Summer and Autumn 2018, 17-24
DOI: 10.22094/JOIE.2018.538229

Three Approaches to Time Series Forecasting of Petroleum Demand in OECD Countries

Majid Khedmati[a,*], Babak Ghalebsaz-Jeddi[b]

[a,b] *Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran*

Received 08 June 2014; Revised 16 February 2016; Accepted 17 February 2016

Abstract

Petroleum (crude oil) is one of the most important resources of energy and its demand and consumption is growing while it is a non-renewable energy resource. Hence forecasting of its demand is necessary to plan appropriate strategies for managing future requirements. In this paper, three types of time series methods including univariate Seasonal ARIMA, Winters forecasting and Transfer Function-noise (TF) models are used to forecast the petroleum demand in OECD countries. To do this, we use the demand data from January 2001 to September 2010 and hold out data from October 2009 to September 2010 to test the sufficiency of the forecasts. For the TF model, OECD petroleum demand is modeled as a function of their GDP. We compare the root mean square error(RMSE) of the fitted models and check what percentage of the testing data is covered by the confidence intervals (C.I.). Accordingly we conclude that Transfer Function model demonstrates a better forecasting performance.

***Keywords*:**Time series forecasting, OECD countries, Petroleum demand

1. Introduction

Forecasting energy demand is an important part of energy planning in which, more accuracy of it leads to better plans. Energy resources are divided into two main categories: renewable and non-renewable resources. Considering this categorization, petroleum (crude oil) is a non-renewable energy resource where it cannot be replaced after it has been used. With development of societies and industries around the world, the demand and consumption of this energy resource that is a basis for a wide variety of oil-based products and is utilized in many applications, increases against its limited supplies (and as a result, the supplies of it may decrease in the long run). Therefore, it is necessary to plan appropriate strategies for petroleum demand to cope with the problem of deficiency of this energy resource in future and ensure energy security. Nevertheless, in the planning of future strategies for energy demand, we need to have a prediction of demand and consumption in future.

Time series forecasting is one of the widely used forecasting methods, in which we use past observations of the same variable to develop a model representing the underlying process generating mechanism that has produced the given observations, and use this model to forecast the future observations and the future behavior of the process. There are many applications for time series forecasting in different contexts and energy is one of the most important of them.

In a comprehensive study, Gooijer and Hyndman (2006) reviewed past 25 years of research works in time series forecasting published in different journals and provided some directions and recommendations for future research in this field (in time series forecasting). Huntington (2011) provided a comparison of 10 year backcast projections of US petroleum consumption based on several functional forms for the period of 2000 to 2009. One structural approach is proposed to consider other factors, except past demand values, affecting the oil demand in the model and showed that autoregressive distributed lag model, a model that allowed oil demand to respond differently to price increases and decreases, is better than all other models including univariate models. Cheong (2009) used a flexible autoregressive conditional heteroscedasticity (ARCH) model to investigate the time-varying volatility of the West Texas Intermediate and Europe Brent crude oil markets and concluded that the volatility in the WTI is greater than in the Brent. Ye et al. (2005) proposed a short-term forecasting model of monthly WTI crude oil spot prices using readily available OECD (Organization for Economic Co-operation and Development) industrial petroleum inventory levels and by comparing it with other models, showed good in-sample and out-of-sample forecasts(in-sample forecasts are the fitted values to the available data whereas out-of-sample forecasts are the forecasted values of the future).

Pedregal et al. (2009) developed an econometric model to estimate the elasticities of five most important crude oil products demand in Spain. The proposed model took advantage of monthly information and included a multivariate unobserved components model. Unobserved components models decompose time series into a number of unobserved though economically meaningful components (mainly trend, seasonality and irregularity). According to their results, the main factor deriving demand was real income with prices having little impact

*Corresponding author Email address: majid.khedmati@yahoo.com

on energy consumption. Akkurt et al. (2010) estimated the gas consumption of Turkey using exponential smoothing, Winters forecasting and Box-Jenkins methods and concluded that for annual data, double exponential smoothing and for monthly data, Seasonal ARIMA model outperforms other models based on forecasting performance. Jiping and Ping (2008) presented a co-integration and vector error correction model to forecast the China's crude oil demand based on four factors of GDP, population, share of industrialsector in GDP, and the oil price, and used this model to forecast China's crude oil demand in 2008-2020 period.

However there is not any research work on forecasting the petroleum demand in OECD countries and this is the first work on it. In this paper we develop time series forecasting model to forecast the petroleum demand in OECD countries. The OECD originated in 1948 as the Organization for European Economic Co-operation (OEEC) to help administer the Marshall Plan for the reconstruction of Europe after World War II. Later, its membership was extended to non-European states. In 1961, it was reformed into the Organization for Economic Co-operation and Development by the Convention on the Organization for Economic Co-operation and Development. Now, OECD is an international economic organization of 34 countries founded in 1961 to stimulate economic progress and world trade. OECD represents itself as a forum of countries committed to democracy and the market economy, providing a platform to compare policy experiences, seeking answers to common problems, identifying good practices, and coordinating domestic and international policies of its members.

The rest of this paper is organized as follows. In section 2, different forecasting methodologies including univariate ARIMA, Winters forecasting and Transfer Function models are described. The data set of the petroleum demand is described and the comparison and evaluation of the forecasts of developed models are provided in Section 3. Finally, the paper is concluded in Section 4.

2. Time Series Forecasting Methodologies: a Briefing

2.1. Univariate ARIMA model

ARIMA is the most popular model for forecasting time series, especially for short-term forecasting, and is useful for longer terms when periodic patterns are present. An ARIMA model is an algebraic statement showing how a time-series variable Z_t is related to its own past values Z_{t-1}, Z_{t-2}, …, and this model is built based on the assumption that the generating process is stationary, at least weakly stationary, i.e. a process that its mean, variance and covariance matrix remain the same by time shifts. An ARIMA model of order (p,d,q) can be written as

$$\Phi(B)\,(1-B)^d Z_t = \theta(B)\, a_t$$

in whicha_t is a random error element at time t generally considered to be a white noise process with mean zero, B is a backshift operator defined as$Bz_t = z_{t-1}$, so (1-B) is differencing and d represents how many times the original data is differenced, called the differencing order. $\Phi(B)$ and $\theta(B)$ are autoregressive (AR) and moving average (MA) operators of order p and q, respectively defined by following equations:

$$\Phi(B) = 1- \varphi_1 B- \ldots- \varphi_p B^p$$

$$\theta(B) = 1- \theta_1 B- \ldots- \theta_q B^q$$

where$\varphi_1, \ldots, \varphi_p$are autoregressive coefficients and$\theta_1,\ldots,\theta_q$ are moving average coefficients, see Box et al. (1994), for example. However, when there is seasonal or periodic patterns in a process in which the observations in, says$s$, interval apart are similar, seasonal ARIMA models are useful. One general form of such models are in the form of multiplicative Seasonal-Nonseasonal ARIMA(p, d, q)(P, D, Q)$_s$, where P, D, Q and s are the seasonality parameters, for more details see Pankratz (1983) or Box et al. (1994), for example.

A multiplicative Seasonal-Nonseasonal model can be written as:

$$\Phi_p(B)\, \Phi_P(B^s)(1-B^s)^D\, (1-B)^d Z_t = \theta_Q(B^S)\theta_q(B)\, a_t$$

Where$\Phi_p(B)$ and $\Phi_P(B^s)$ denotes autoregressive operators Nonseasonal and Seasonal terms, and $\theta_q(B)$ and $\theta_Q(B^S)$ denotes moving average Nonseasonal and Seasonal terms.

2.2. Wintersforecasting model

Winters forecasting model smoothes the observations by Holt-Winters exponential smoothing and provides short to medium-range forecasting. This model is used when both trend and seasonality are present, with these two components being either additive or multiplicative. When the seasonal pattern in the data depends on the size of the data we use multiplicative model whereas, when the seasonal pattern in the data does not depend on the size of the data we use additive model. Winters model calculates dynamic estimates for three components: level, trend, and seasonal. At first we define following variables:

$$L_t = \alpha\, (Y_t/S_{t-p}) + (1-\alpha)\, [L_{t-1} + T_{t-1}]$$

$$T_t = \gamma\, [L_t - L_{t-1}] + (1-\gamma)\, T_{t-1}$$

$$S_t = \delta\, (Y_t /L_t) + (1-\delta)\, S_{t-p}$$

in which$L_t$is the level at time t,$\alpha$is the weight for the level,$T_t$is the trend at time t,$\gamma$is the weight for the trend,$S_t$is the seasonal component at time t,$\delta$is the weight for the seasonal component, p is the seasonal period and Y_t is the data value at time t.

Using the above variables and equations, the Winters multiplicative model is:

$$\hat{Y}_t = (L_{t-1} + T_{t-1})\, S_{t-p}$$

and Winters additive model is:

$$\hat{Y}_t = L_{t-1} + T_{t-1} + S_{t-p}$$

where$\hat{Y}_t$ is the fitted value, or one-period-ahead forecast, at time t.

Winters model uses the level, trend, and seasonal components to generate forecasts. The forecast for m periods ahead from a point at time t is$L_t + mT_t$.

2.3. Transfer function model

The Transfer Function model relates an input process X_t to an output processY_t, using an impulse response function that determines the Transfer Function for the system through a dynamic linear relationship between input X_t and outputY_t. When the conditions under whichthe data for the time series process is collected change over time, we can describe these changes by introducing certain inputs and consequently we can use Transfer Function models. Transfer Function models enable us to consider some more parameters, other than the past values of the same variable, affecting the underlying processes in the models and consequently we can describe and model these processes more precisely. The general form of the Transfer Function model with b time delay between the response and the input is:

$$y_t = \vartheta(B)x_t + N_t = \frac{\omega(B)}{\delta(B)}x_{t-b} + \frac{\theta(B)}{\varphi(B)}\varepsilon_t$$

where $\vartheta(B) = \sum_{i=0}^{\infty} \vartheta_i B^i$.

Furthermore N_t represents unobservable noise process and can be modeled as an ARIMA(p,d,q) model and, ε_t represents the independent random shocks. Also x_t and N_t is assumed to be independent, see Montgomery et al. (2008).

3. Model Fitting and Forecasting

3.1. Data set

Due to the importance of petroleum and its products in wide variety of industries and applications, in this paper a time series forecasting model will be proposed for forecasting the petroleum demand of OECD countries. To do this, we used theOECD countries petroleum demand data from 2001 to 2010,extractedfrom U.S. Energy Information Administration (EIA). This data set contains 117 observations (in thousands of barrels per day) for monthly petroleum demand of OECD countries from January 2001 to September 2010.

The data were divided into two groups, the data in the first group in order to build the forecasting models and testing data in order to verify the accuracy and performance of the proposed models. Hence, first 105 observations from January 2001 to September 2009 were selected as the first group and the rest of the observations from October 2009 to September 2010 were assigned to the second group as the testing data to test and verify the performance of the proposed forecasting models.

The observations of the training data are shown in Figure 1. According to Figure 1, the process is not stationary and the mean of the process changes during the time and also, we can find a seasonality in the process since for example, the values of observations of May is less than the values of other months within the same year and this pattern is repeated in other years.Hence, the period of seasonality is considered to be$s$ = 12. Whereas the ARIMA models are designated for stationary time series, it is necessary to transform this process into a stationary one. To make the mean of a series stationary, we difference successive observations of the random variable Z_t, and also we use seasonal differencing to eliminate seasonal non-stationarity as $w_t = (1-B)(1-B^s)Z_t$.

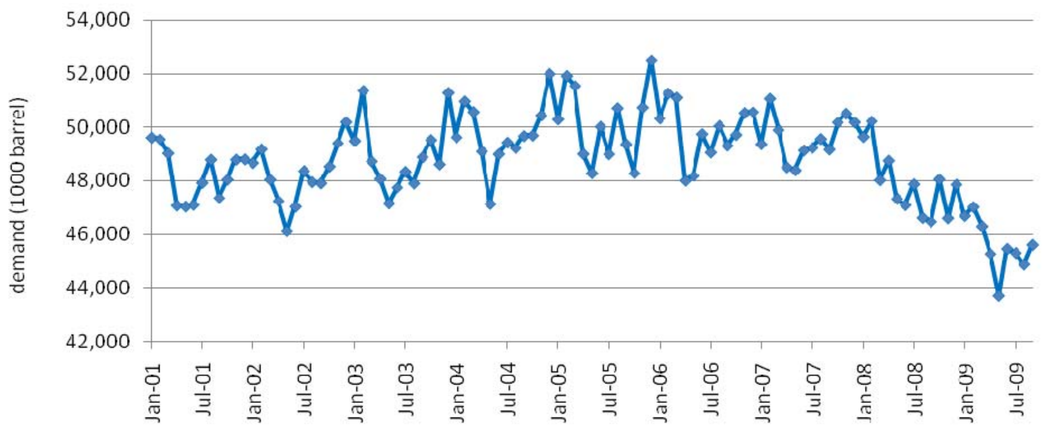


Fig. 1. OECD countries monthly petroleum demand from Jan-2001 to Sep-2009 (12 observations set aside as testing data set).

According to Figure 2, the new variable w_t is stationary and seasonality is removed from it. Although, it seems that seasonality with period of s = 6 might exist in the process, however, we skip it here and examine it more precisely in developing the models.

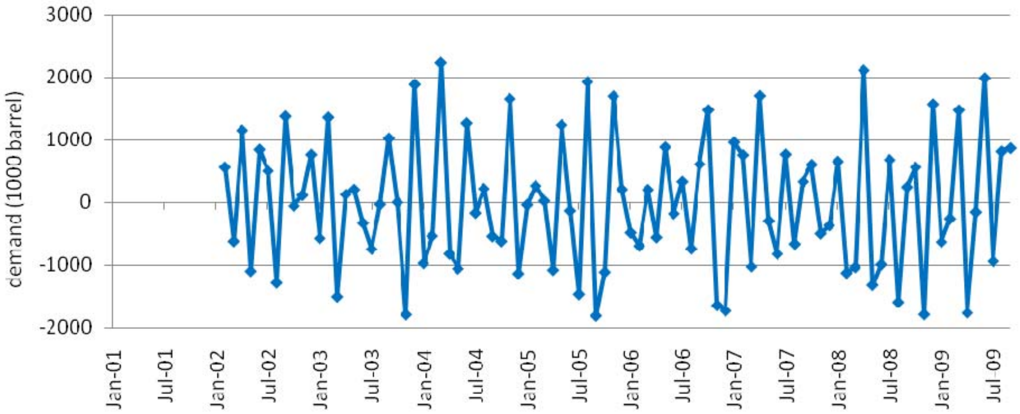


Fig. 2. OECD countries simple and seasonal differenced monthly petroleum demand.

Now, we use new variable w_t to develop time series forecasting models.

3.2. ARIMA modelformulation

Now, using the autocorrelation function (acf) and partial autocorrelation function (pacf) ofdifferenced observations, we identify twoARIMA models and by comparing them, determine the best onebased on one-step-ahead forecasting performance.

Analyzing the acf and pacf of the observations and residuals, we identified two different seasonal ARIMA models including:ARIMA$(2,1,0)(0,1,0)_{12}$ andARIMA$(2,1,0)(0,1,0)_6(0,1,0)_{12}$.

Then we estimated the parameters of these models and calculated the RMSE, AIC and BIC measures to compare and evaluate the performance of the proposed forecasting models. The proposed ARIMA models in the backshift notationare, respectively:

$$(1 + 0.741B + 0.584B^2)\,(1 - B)\,(1 - B^{12})\, z_t = a_t (1)$$

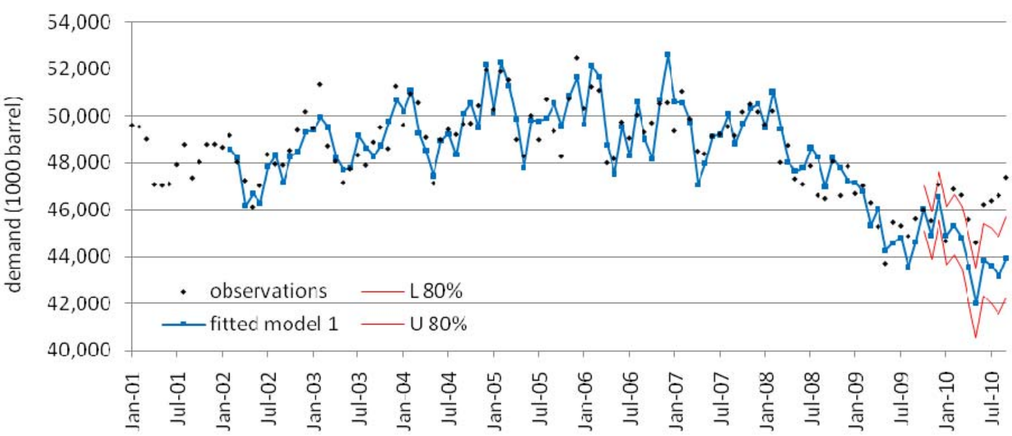
$$(1 + 0.704B + 0.475B^2)\,(1 - B)\,(1 - B^6)\,(1 - B^{12})\, z_t = a_t (2)$$

Here, we limit reporting the coefficients up to 3 decimals accuracy. Table 1 compares the performance measures of two seasonal ARIMA models. As another goodness criterion, we also consider the number of (last 12) observations correctly predicted by the 80% confidence interval in each model.

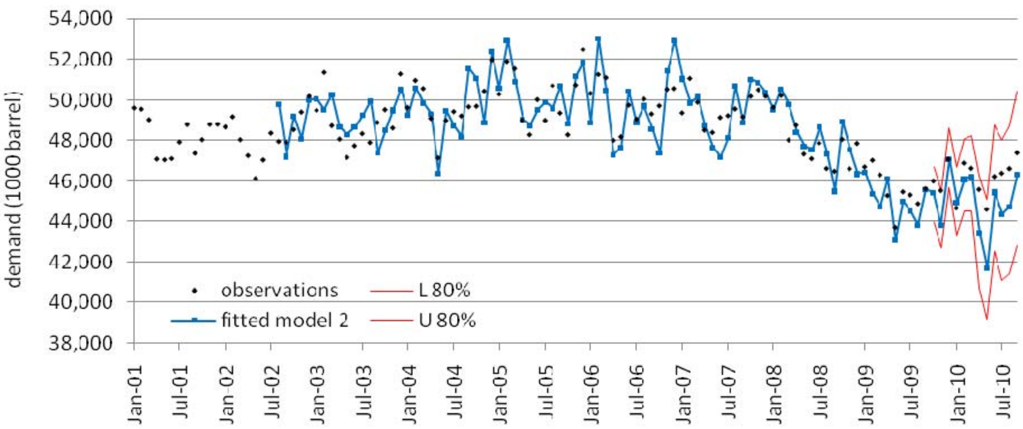
Table 1
Comparison of two seasonal ARIMA models based on RMSE, AIC and BIC

Number	ARIMA Models	RMSE	AIC	BIC	No. of data in 80% C.I.
(1)	ARIMA$(2,1,0)(0,1,0)_{12}$	774.87	1488.13	1495.69	4 out of 12
(2)	ARIMA$(2,1,0)(0,1,0)_6(0,1,0)_{12}$	1049.69	1443.48	1450.84	11 out of 12

Figure 3 shows the fitted values and forecasted values with 80% confidence intervals. So, we use both results presented in Table 1 and Figure 3 to determine the best model. The results in Table 1 show that the AIC and BIC measures for model (2) are less than model (1)and adversely, the RMSE for model (1) is less than the RMSE of model (2). Relying only to these results we can not specify which model is better, but considering the results of Figure 3, we can see that in model (1), 8 original observations is not included in the 80% confidence interval.



(a)



(b)

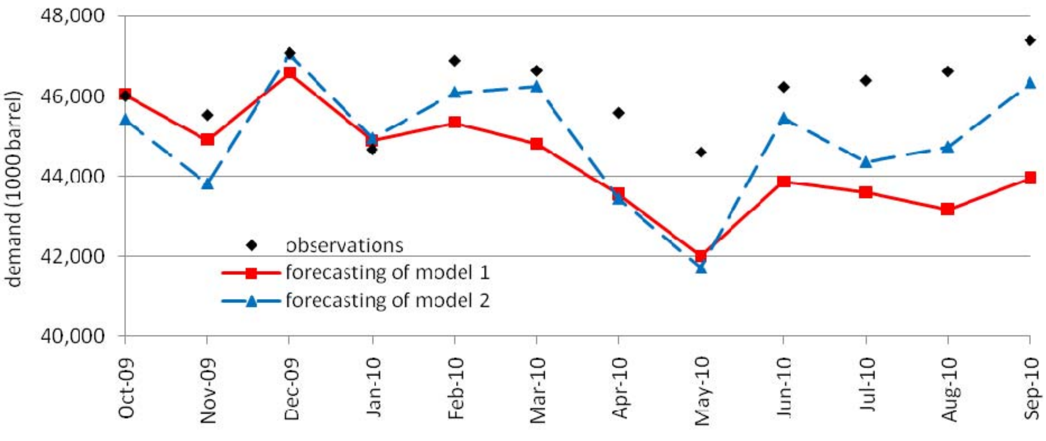


Fig. 3. Data and fitted forecasts with 80% C.I. for: (a) ARIMA(2,1,0)(0,1,0)$_{12}$ and (b) ARIMA(2,1,0)(0,1,0)$_6$(0,1,0)$_{12}$, (c) comparison of observations and forecasts of the models.

In other words, in model (1) the 80% confidence interval covers only 4 observations out of the last 12 testing data. In contrast, the 80% confidence interval of model (2) covers 11 observations out of 12. Therefore, forecasting performance of model (2) is far better considering the test data.

Finally, considering the performance measures of RMSE, AIC and BIC,in addition to the forecasting performance, we conclude that model (2) is superior to model (1). In fact, the best model fitted to this process is model (2).

3.3. Wintersforecasting modelformulation

In this section we fit a Winters forecasting model to the observations. The first step in this procedure is determining the appropriate Winters model, either multiplicative or additive, for fitting to the observations. Since the size of seasonal pattern for OECD petroleum demand data is relatively proportional to the observations, we choosemultiplicativeWinters model.Then, we should estimate the parameters value for the selected model. The fitted model to the observations and the forecasting values for 12 testing data are shown in Figure 4.

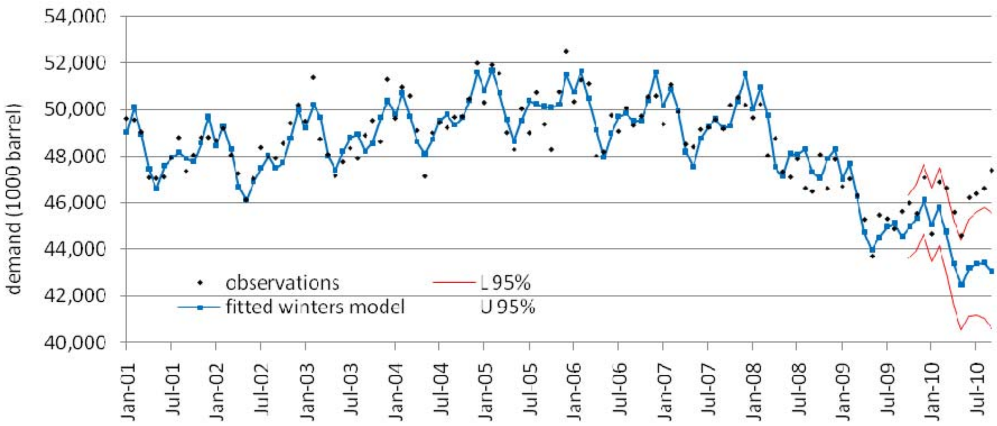


Fig. 4. Data and fitted forecasts with 95% C.I. for multiplicative Winters model.

The optimal values for the model parameters (α, γ, δ)are(0.274, 0.115,0.104). The RMSE, AIC and BIC for this model areshown in Table 2. The RMSE of this model isconsiderably small in comparison to the RMSE of the best ARIMA model. Also, Figure 5 compares the forecasted and original values with 95% confidence interval.

Table 2
Performance results of Winters model based on RMSE, AIC and BIC

Model	RMSE	AIC	BIC	No. of data in 95% C.I.
Winters model	680.90	1375.90	1383.90	5 out of 12

The forecasting results in Figure 5 show that the 95% confidence interval covers only 5 observations out of 12. Considering both RMSE and forecasting performance of this model, we conclude that this model has a good performance in terms of RMSE, while it has poor performance in forecasting and only 42% of the original observations are covered by a wider 95% confidence interval.

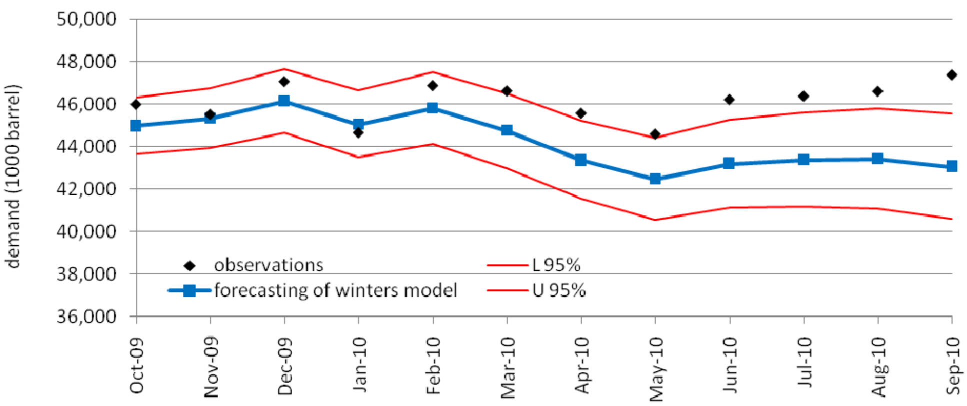


Fig. 5. Comparison of observations and forecasts of Winters model

3.4. Transfer function model formulation

Using Transfer Function models, we can introduce new influential variables to the model that affects the underlying process.Consequently, we can explain the behavior of the process in more details and develop a more accurate model with better forecasting performance.
A variable identifiedas an influential variable on the behavior of the OECD countries petroleum demand is GDP. GDP is a measure of the output of an economy resulting from the production of marketed goods and services within the national boundary. Goods and services are (for this purpose) valued at their market price and output is gross of capital depreciation. Hence, we collected the GDP percent changes from January 2001 to September 2010 and used Transfer Function method to add this variable to the model.
At the first step, we fitted an ARIMA(2,1,0) model to GDP data and then, we calculated the cross correlation values between petroleum demand and GDP.With analyzing the cross correlation values,the time delay considered to be 4and according to different impulse response functions, see Montgomery et al. (2008), the selected impulse response function is $\upsilon_t = \omega_0 B^4$.
Estimating ω_0, the initial model for petroleum demand will be$\hat{y}_t$ = -131.16x_{t-4} + N_t. Analyzing the residuals (i.e. noise N_t) of this relation and fitting an ARIMA model on them, we obtain the following TF formulation:

$$y_t = -131.16x_{t-4} + \frac{(1-0.40B)(1+0.76B^{12}+0.36B^{24})}{(1-0.87B)(1-0.29B^3+0.22B^{19})}\varepsilon_t \tag{3}$$

The results in Table 3 show the RMSE, AIC and BIC for this model. Fitted and forecasted values with 80% confidence interval are also shown in Figure 6.

Table 3
Performance results of Transfer Function model based on RMSE, AIC and BIC

Model	**RMSE**	**AIC**	**BIC**	**No. of data in 80% C.I.**
TF model (3)	798.97	1628.12	1648.96	12 out of 12

Based on the Figure 6, we see that the 80% confidence level for this model covers all 12 testing observations which is an acceptable accuracy performance. Moreover, the RMSE of the Transfer Function model is significantly less than the RMSE of the best ARIMA model.

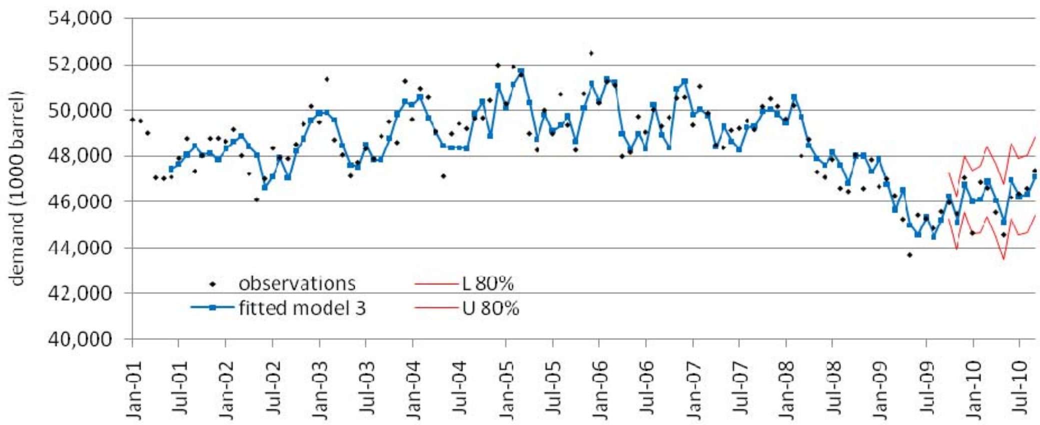


Fig. 6. Data and fitted forecasts with 80% C.I. for model (3).

Comparing the results of this model with the results of ARIMA and Winters forecasting models in the previous sections, we recognize that the RMSE of the Transfer Function model is considerablyless than model (2) -the best ARIMA model- however, a little greater than the RMSE of Winters forecasting model. Also all of the forecasted values of this model are within the 80% confidence interval and significantly close to the actual values. In spite of the good performance of the Winters model in terms of RMSE, this model has poor results in covering the testing observations for a given confidence level. In contrast, Transfer Function model provides satisfactory performances based on both RMSE and confidence level.Therefore, introducing and adding the new variable GDP to the modelhas led to obtain a model that describes the underlying process generating mechanismmore accurate than other models.As a consequence, we obtained a better forecasting

performance results than ARIMA and Winters forecasting model.
It is important to note, based on the results, that univariate models generally underestimates the future observations and almost all of the forecasted observations are less than the original ones. This poor performance of the univariate model may be due to a sudden decrease in petroleum demand starting January 2008 untilMay 2009. However, the Transfer Functionmodel adequately describes the model and the forecasts of this model are very close to the original observations.

4. Conclusions

Using time series methodologies, we proposed five forecasting models for petroleum demand of OECD countries including two seasonal ARIMA models, aWinters forecasting model and aTransfer Function model. We compared the two seasonal ARIMA models and selected the best of them. Then, we compared it with the proposed Winters forecasting and Transfer Function models and concluded that the best one among all of them is the proposed TF model. Although the RMSE of the Winters model was considerably small, however its forecasting performance was poor so that only 42% of the original observations (12 test data) were within the forecastedconfidence interval. In developing the TF model, we used the GDP percent changes as an influential variable that can describe the underlying process in more details. The TF model improved the fitting/forecasting performance in terms of both RMSE and the capability of the C.I. in covering the future (testing) data. To compare the models, we also used AIC and BIC as performance criteria.

References

Akkurt, M., Demirel, O.F.& Zaim, S. (2010). Forecasting Turkey's natural gas consumption by using time series methods. European Journal of Economic and Political Studies, 3, 1-21.

Box, G.E.P., Jenkins, G.M.& Reinsel, G.C. (1994). Time series analysis forecasting and control. 3rd edition, Prentice-Hall International, New Jersey.

Cheong, C.W. (2009). Modeling and forecasting crude oil markets using ARCH-type models. Energy Policy, 37, 2346-2355.

Gooijer, J.G.D.& Hyndman, R.J. (2006). 25 years of time series forecasting. International Journal of Forecasting, 22, 443-473.

Huntington, H.G. (2011). Backcasting U.S. oil demand over a turbulent decade. Energy Policy, 39, 5674-5680.

Jiping, X.& Ping, W. (2008). An analysis of forecasting model of crude oil demand based on cointegration and vector error correction model (VEC). International Seminar on Business and Information Management, 485-488.

Montgomery, D.C., Jennings, C.L.&Kulahci, M. (2008). Introduction to time series analysis and forecasting. John Wiley & Sons, New Jersey.

Pankratz, A. (1983). Forecasting with univariate Box-Jenkins models- concepts and cases. John Wiley & Sons, New York.

Pedregal, D.J., Dejuan, O., Gomez, N.& Tobarra, M.A. (2009). Modelling demand for crude oil products in Spain. Energy Policy, 37, 4417-4427.

U.S. Energy Information Administration (EIA), 2011.http://www.eia.gov/emeu/ipsr/demand.html.

Ye, M., Zyren, J.& Shore, J. (2005). A monthly crude oil spot price forecasting model using relative inventories. International Journal of Forecasting, 21, 491-501.

This article can be cited: Khedmati, M., & Ghalebsaz-Jeddi, B.(2018). Three Approaches to Time Series Forecasting of Petroleum Demand In OECD Countries. *journal of Optimization in Industrial Engineering*. 11(2),2018,17-24.

URL: http://qjie.ir/article_538229.html
DOI: 10.22094/joie.2018.538229