Periodicals of Engineering and Natural Sciences ISSN 2303-4521
Vol. 8, No. 2, June 2020, pp.656-667

Predictive study on time series modeling and comparison with application

Haifa Taha Abd[1], Nabaa Naeem Mahdi[2], Asia Hmoud Hussein[3]
[1,2,3]University of Mustansiriyah / collage of Management and Economics

ABSTRACT

Efficient time series modeling and forecasting are essential in different practice areas. Consequently, much active research work on this topic has been ongoing for several years. Given the importance of different prediction methods, this research aims to provide a brief description of some common time series prediction models used with their salient features. Therefore, Box-Jenkins and exponential booting models were compared, along with the strengths and weaknesses of the prediction. Our discussion on various time series models is supported by giving the experimental prediction results, which were made to the actual monthly sales of some fuel products for the period 2014-2017. While installing the Data Set template, special care is taken to select the most creative. To evaluate prediction accuracy in addition to comparing it, we used several criteria, mean square error (MSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE), and mean square error (RMSE). To obtain originality and clarity in our discussion on modeling and forecasting a time series, we were able to obtain assistance from various published research work from famous magazines and some standard books and it was concluded that the 3ARMA terms best model among the Box-Jenkins models built based on the dependence of gas oil sales in Iraq, as well as Simple exponential smoothing is the best exponential smoothing model to forecast in the coming years for sales of improved gas and gas oil in Iraq.

Keywords: Forecasting Sales, Methodology of Box-Jenkins, Exponential Smoothing

Corresponding Author:

Haifa Taha Abd
collage of Management and Economics
University of Mustansiriyah
E-mail: haefaa_adm@uomustansiriyah.edu.iq

1. Introduction

As a result of the development of statistical analysis techniques that the world has witnessed in recent years has increased interest in the study forecasting models, which led to the emergence of many developed to predict the statistical methods and the most important topics used in the analysis and interpretation of the behavior of different phenomena by studying the evolution of the historical and constrained events and linking factors and different variables across periods, but it is the subject of time series and that the most important use of the results of time series is to predict what will happen to the phenomena studied in the future less error as possible and to predict the future depends mainly on the past and previous experiences and this helps proper planning and policy-making and the development of plans and adjusted in addition to it has become time-series analysis an important role in the process of decision-making in various applied fields.

2. Research problem

Iraq is one of the countries that rely mainly on oil as energy occupies its oil role strategically important in stimulating and financing of other sectors, hence the need for a predictive study based on sound scientific principles contribute to improved policy-making and the development of strategic plans and future expectations to increase the level of production in addition to the optimal use of oil derivatives, for example, interest in sales of oil (such as derivatives Premium gasoline, Gas oil) which represent the foundation in the subject of this research study and become a problem here is how to reach the most appropriate and best modern statistical models for use in oil sales forecasting in terms of quality of data representation of products for which the estimated parameters and in terms of significant statistics so that the mean square error is going down to the lowest possible level, leading to the accuracy of the estimates so consistent with the theoretical and practical framework

3. Purpose of research

The study aims at the definition of statistical forecasting models that included the Box-– Jenkins and models of exponential smoothing and to clarify the necessary steps to do and describe the special features of the process under which each formed a series of time series studied and how to choose the appropriate means to determine the forces acting in the series and models able to describe the behavior of the evolution of this series over time and used in the forecasting process and Compared the models statistical with each other and choose the best for predicting sales of oil products (Premium gasoline, gas oil) in Iraq.

4. Materials and methods

Mixed prediction models, as well as exponential smoothing models, the self-regression model, and the moving average will be reviewed, as several models were compared based on some statistical measures, and the following of the theoretical basis for them.

4.1. Types of time series

There are two types of time series:

4.1.1. Time series stable

It is said that the time series is stable if all its properties does not change with time and stability be:

a) Full stability: The series is called it fully stable if each allowable points $(t_1, t_2, ..., t_n)$, and (k) the joint probability distribution That is, check the following condition

$$Fx(t_1), x(t_2), ..., x(t_n) = Fx(t_1+k), x(t_2+k), ..., x(t_n+k)$$

b) Stability rank M: can be said that the time series is stable for the rank of M if the points allowed $(k, t_1, t_2, ..., t_n)$ each common moments rank $M\{x(t_1), x(t_2), ..., x(t_n)\}$ and have a presence and identical with moments common as it is explained below:[1,2]

$$E\{(x_{t_{1+k}})^{m_1}, (x_{t_2+k})^{m_2}, ...(x_{t_{n+k}})^{m_n}\}$$

$$\mu_1 + \mu_2 + ...+ \mu_n \leq \mu$$

4.1.2. Unstable time series

If the values do not vibrate around a fixed mean or constant variance, this indicates that the series is unstable. Through the drawing of the series shows the descriptive features of the data, such as the general direction, seasonal changes, abnormal data [3].

4.2. Methods of checking the stability of the time series

4.2.1. The chart

The first step in analyzing any time series is to draw the series data and through it we have a good idea of the series containing seasonality, general trend, abnormal values or instability, and by drawing the series shows that it needs the appropriate transformation to stabilize its average and variance before any analysis.

4.2.2. Testing the roots of the unit

A procedure examines the time series and determines whether they are stable or not and be by taking the first difference of the series and there are many tests of the root of the unit such as:

- *Dickey-Fuller Test (ADF) Augmenter*

1- A simple Dickey-Fuller test neglected hypothesis of the possibility of correlation errors because it is assumed that the random error is a random shock while the test of Dickey-Fuller Augmenter came on the ruins of a simple test where the listed hypothesis probability of correlation errors and this test is based on the choice of significant statistical estimator of the models calculated by the least-squares method and according to the models listed below:

❖ Auto-Regressive AR(p)

That is according to the following formula:

$$\nabla_{y_t} = \lambda y_{t-1} - \sum_{j=2}^{p} \phi \nabla_{y_{t-j+1}} + \varepsilon_t \qquad (1)$$

❖ Auto-Regressive model AR (p) with the constant presence of which is in accordance with the following formula:

$$\nabla_{y_t} = \lambda y_{t-1} - \sum_{j=2}^{p} \phi \nabla_{y_{t-j+1}} + c + \varepsilon_t \qquad (2)$$

❖ Auto- Regressive model AR (p) with fixed and vehicle presence general trend which is in accordance with the following formula:

$$\nabla_{y_t} = \lambda y_{t-1} - \sum_{j=2}^{p} \phi \nabla_{y_{t-j+1}} + b_t + c + \varepsilon_t \qquad (3)$$

Accordingly, the decision in accordance with the following:

If $T_t < T_c$ the decision to reject the null hypothesis and accept the alternative hypothesis that any time series is stable but if $T_t > T_c$ the decision to accept the null hypothesis and rejected any alternative hypothesis that the series is unstable.

Here we use a statistical table differs from the table in the first test and the degree of delay determines depending on Akaike or calculable Schwarz or Hannan Quinn is determined by the degree of delay of the series before applying Dickey Fuller test for the purpose of determining the type of the test used in the detection of unilateral root and the general trend in the series.

- *Phillips- Peron (P-P)*

In this test we will go beyond the problems of auto-correlation of the residuals and the instability of the random variation of the error that was experienced by Dickey Fuller simple test, this test procedure and follow the following:

▪ estimate the model by ordinary least squares.

▪ estimate the variance which is called (short-term) according to the following formula:

$$\sigma^2 = \frac{1}{2} \sum_{t=1}^{n} e_t^2 \qquad (4)$$

Where (e_t) is the estimated residual error.

• estimate the corrected coefficient (δ_t^2), which is called Long-term variances, which is extracted from the structure of common variations of the residuals previous models whereas:

$$\delta_t^2 = \frac{1}{n} \sum_{t=1}^{n} e_t^2 + 2 \sum_{i=1}^{n} (1 - \frac{i}{\tau + 1}) \frac{1}{n} \sum_{t=i+1}^{n} e_t + e_{t-1} \qquad (5)$$

In order to estimate this variance must be defined in terms of the number of delays estimated observation (n).

We calculate statistica by Phillips and Peron the following formula [4,5]:

$$t_{\phi_1} = \sqrt{k} \times \frac{(\phi_1 - 1)}{\sigma_{\phi_1}} + \frac{n(k-1)\sigma_{\phi_1}}{\sqrt{k}} \quad (6)$$

4.3. Box – Jenkins methodology

The methodology of this method refers to the general ARIMA model, and this model is usually applied in time series analysis and forecasting, and the term ARIMA denotes an abbreviation for a combination of Autoregressive/Integration/Moving Average models.

4.3.1. Autoregressive model (AR (p))

Predictable value (Y_{t+k}) based on the conditional expectation $(E_t(Y_{t+k}))$ and who writes the following formula:

$$E_t(Y_{t+k}) = E(Y_{t+k} / Y_t, Y_{t-1}, Y_{t-2}, ...) \quad (7)$$

. If the time series follow the model of auto- regression from one degree and removal of the k of future periods, the written form model according to the following:

$$Y_t = \phi Y_{t+k-1} + a_{t+k} \quad (8)$$

When taking the expectation becomes the following formula:

$$E_t(Y_{t+k}) = \phi^k Y_t$$

For conditional variance to (Y_{t+k}) previous information based on the formula would be as follows:

$$\mathrm{var}_t(Y_{t+k}) = \mathrm{var}(Y_{t+k} / Y_t, Y_{t-1}, Y_{t-2}, ...) \quad (9)$$

$$\mathrm{var}(Y_{t+k}) = \frac{\sigma^2(1-\phi^{2k})}{1-\phi^2} \quad (10)$$

It is known that (Y_t) is a random process stable distribution gauss distributed naturally, the confidence interval 95% probability written as follows:

$$E_t(Y_{t+k}) \mp 1.96\sqrt{\mathrm{var}(Y_{t+k})} \quad (11)$$

And that confidence interval 95% probability values prediction and model AR (1) be according to the following formula:

$$\phi^k Y_t \mp \frac{\sigma^2}{1-\phi^2} \quad (12)$$

And increase the amount of displacement k, the variation of the predictive values f the model AR (1) becomes:

$$\mathrm{var}_t(Y_{t+k}) = \frac{\sigma^2}{1-\phi^2} \quad (13)$$

4.3.2. Moving Average Model (MA (q))

The Moving Average (MA) model relates the current time series values to the random errors that occurred in the previous period instead of the actual series values themselves. The moving average model can be written as follows [4,6]:

$$Y_t = \mu + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \cdots \theta_q\epsilon_{t-q} \quad (14)$$

Where,

y_t is the response or variable adopted at period t.

μ is the mean about which the series fluctuates.

θ is the moving average parameters to be estimated.

ϵ_{t-q} is the error terms (q=1,2,…,q) assume to be independently distributed over time.

4.3.3. Mixed models ARMA(p, q)

The Mixed Autoregressive Moving Average (ARMA) Model was the combination of AR model and MA model and was assume stationary. In other words, the series yt is assumed stationary (no need differencing) and the ARMA model is written as:

$$\emptyset_0 + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q y_{t-q} \quad (15)$$

The Mixed Autoregressive Integrated Moving Average (ARIMA) Model was formulated when the assumption of stationary assumption was not met. The differencing is required to achieve stationary. The general term as ARIMA (p, d, q), where p represents the AR model, d denotes the number of times the variable y_t needs to be differenced to achieve stationary and q represented MA model [7,8,9,10].

4.3.4. Steps to build the Box Jenkins model

4.3.4.1. Identification stage

At this stage is plotted time series under study and note whether they contain a general trend or the impact of seasonal then auto-correlation calculation and partial auto-correlation and then are selected simple types of models (ARIMA) to estimate the appropriate value to the degree of the model (q, d, p) the objective of this phase subsides in the choice of values for (q, d, p) in the model when containing the string on a general trend, it is reconciling the model of the regression of the general trend and put forward the direction of the series and deal with error as stable and free general trend or series through the differences and series often be stable after taking the second difference.

4.3.4.2. Estimation stage

The parameter estimation to model auto-regression and parameter of moving average model chosen according to the model degree to which have been relied upon in the first step and is appreciated by using the method (Maximum Likelihood Techniques) or (Exact Maximum Likelihood Techniques) or method (Back casting).

4.3.4.3. Diagnostic checking stage

To choose the best model from among the estimated models, we rely on several criteria, and from these criteria [11]:

- *Akaka information criterion*

This test is used to determine the degree of auto- regression (P) model, and according to the following formula:

$$AIC(k) = lin(\hat{\sigma}_k^2) + \frac{2k}{T} \quad (16)$$

The higher the value of this small scale was better and this is the best form prescribed.

- *Schwarz criterion*

It is the standard used for determining the degree of auto- regression model, according to the following formula:

$$SC = lin(\hat{\sigma}_k^2) + \frac{lin(T)k}{T} \quad (17)$$

A scale model test class (k) to lower the value of this scale and that the degree of this scale approaching the estimated degree of the model whenever the larger sample size [12].

4.3.4.4. Forecasting stage

It is the final stage where the use is estimated to describe the model form prescribed parameters and then calculates the predictive values of the time series to predict whether a point or period with efficiency values f predictive standards [13].

4.4. Exponential smoothing model

Exponential Smoothing models easy to use and accurate results in addition to the efficiency of their estimates and this way to do not be bound any condition of statistics pertaining to errors form prescribed by the statistical distribution and has a characteristic loss of memory because according to these models are given observation facilities have a weight and grow up this weight with observation time period near the to predict the latest observation, making the importance of the current observation better than its predecessor and in this research we will address some non-seasonal models that have been applied to the practical side, namely:

4.4.1. Simple exponential smoothing

As the data for the time series is prepared according to the following relationship:

$$\mathrm{F_t} = \alpha \mathrm{x_t} + (1-\alpha)\mathrm{F_{t-1}} \qquad (18)$$

Where 0<α<1 The constant preamble, that the prediction is made according to the following relationship:

$$F_t = f_t \qquad (19)$$

4.4.2. Single exponential smoothing with linear trend

This model is called a model Holt's Exponential Smoothing It is rewarded in terms of forecasting to the ARIMA model (0,2,2) and by parameters $\theta_2 = a-1, \theta_1 = 2-a-ab$ since $a \& b$ represent the constants of the exponential smoothing and there are many treatments for the initial values one of them $T_0 = 0, F_0 = x_1$ or that the linear general trend model is reconciled to half of the series data where $\hat{a} = F_0$ and the regression parameter $\hat{b} = T_0$ and the equations of the model are described in the following way:

$$F_t = \alpha x_t + (1-\alpha)(F_{t-1} + T_{t-1}) \qquad (20)$$

$$T_t = b(F_t - F_{t-1}) + (1-b)(T_{t-1}) \qquad (21)$$

For the purpose of predicting a later period of time is by applying the following equation [4, 14]:

$$f_{(t+h)} = F_t + hT_t \qquad (22)$$

4.4.3. Brown linear exponential smoothing

We can combine the regression model and exponential smoothing if the data contains a linear trend, this will lead us to get a good prediction model and the prediction value is calculated according to the following formula:

$$f_{(t+h)} = 2F_t - F_t^* + h\left(\frac{a}{1-a}\right)\left(F_t - F_{t-1}\right) \qquad (23)$$

The initial value is equal to $F_0 = F_0^* = x_1$ the constant of exponential smoothing is (a) be confined between (0,1) [15,16].

5. The practical side

The time series of both sales of gas oil and perfect petrol was used for the period from 1/1/2008 to 1/11/2017 which is equivalent to N = 119, and the application side includes a presentation of the results of the techniques that use to forecast, and it is worth noting the use of the Eviews program to obtain the results.

5.1. Time series analysis

The first stage in any time-series analysis should be to draw time-series observations. This is often a very important part of any data analysis, as qualitative characteristics such as trend, seasonality and external values will usually be visible if present in the data. Figure 1 shows monthly sales of gas oil with time from 2008 to November 2017.

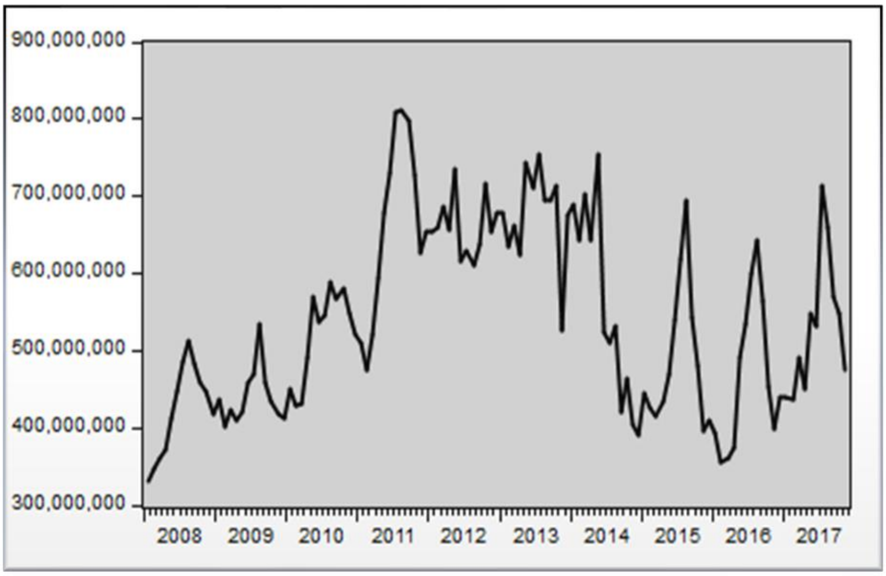


Figure 1. Graph for gas oil over time

Box Jenkins' methodology for building a time series model begins with examining the time series to ensure whether or not the series is stable. To achieve this purpose, the Ducky Fuller and Philip Peron tests were used, Table 1 shows that the chain is unstable by contrast due to the general trend according to the Dickey Fuller test (ADF) and Philip Peron test (PP) where the value of significance level was 0.6931 and 0.1769 respectively, which is greater than 0.05, and therefore the null hypothesis cannot be rejected and from it, we conclude that the chain is unstable. It requires taking the first differences to achieve stationary.

Table 1. Trend test result

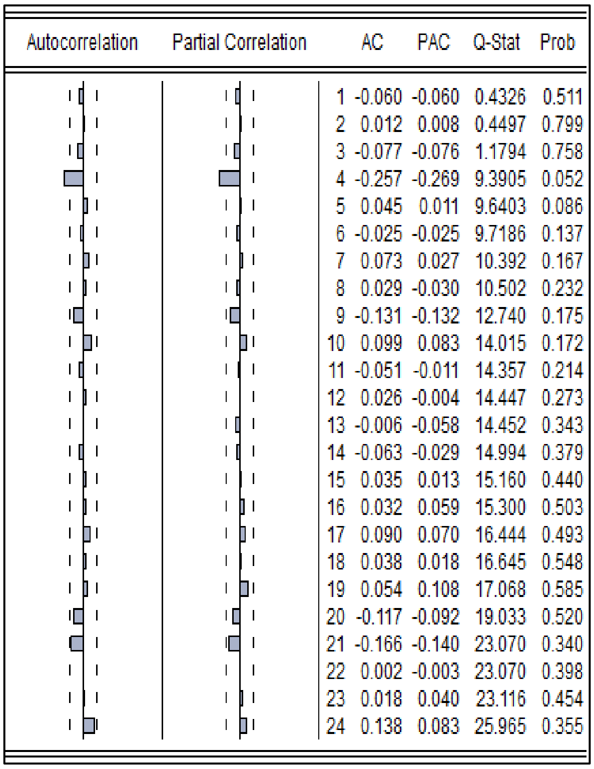
	Statistics	P-Value
ADF Level	-1.81086	0.6931
First differences	-7.27556	0.0000
PP Level	-2.86754	0.1769
First differences	-14.2755	0.0000

Correlogram analysis also confirmed the stationary of the series, not after taking the first differences.
Depending on the theoretical aspect, several models were built. A section of these models is included according to Table 2 to choose the best model according to R^2 and the lowest standards of AIC and SBIC, and we must mention that there can be no accurate or perfect model from ARIMA because it is "an art of science", and we must mention that the construction of any model depends on the characteristics of both ACF, PACF.

Table 2. ARMA Model for Difference gas oil

	ARIMA (1,1,4)	ARIMA (2,1,2)	ARMA (3,1,3)	ARIMA (6,1,6)
Significant coefficient	1	2	2	2
$Sigma^2$ (Volatility)	3.27E+15	3.70E+15	3.61E+15	2.77E+15
Adj R^2	0.015	0.021	0.045	0.268
AIC	38.7617	38.7559	38.7316	38.4877
SBIS	38.8556	38.4898	38.8256	38.5816

From the results of Table 2, we conclude that the best model is ARIMA (6,1,6) because it has the lowest AIC AND SBIC and also the highest R2. Therefore, the parameters of this model will be estimated and plot AIC, PAIC for error and forecast the time series of gas oil based on this model Figure 2.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.060	-0.060	0.4326	0.511
		2	0.012	0.008	0.4497	0.799
		3	-0.077	-0.076	1.1794	0.758
		4	-0.257	-0.269	9.3905	0.052
		5	0.045	0.011	9.6403	0.086
		6	-0.025	-0.025	9.7186	0.137
		7	0.073	0.027	10.392	0.167
		8	0.029	-0.030	10.502	0.232
		9	-0.131	-0.132	12.740	0.175
		10	0.099	0.083	14.015	0.172
		11	-0.051	-0.011	14.357	0.214
		12	0.026	-0.004	14.447	0.273
		13	-0.006	-0.058	14.452	0.343
		14	-0.063	-0.029	14.994	0.379
		15	0.035	0.013	15.160	0.440
		16	0.032	0.059	15.300	0.503
		17	0.090	0.070	16.444	0.493
		18	0.038	0.018	16.645	0.548
		19	0.054	0.108	17.068	0.585
		20	-0.117	-0.092	19.033	0.520
		21	-0.166	-0.140	23.070	0.340
		22	0.002	-0.003	23.070	0.398
		23	0.018	0.040	23.116	0.454
		24	0.138	0.083	25.965	0.355

Figure 2. ACF, PACF for the residual ARIMA (6,1,6)

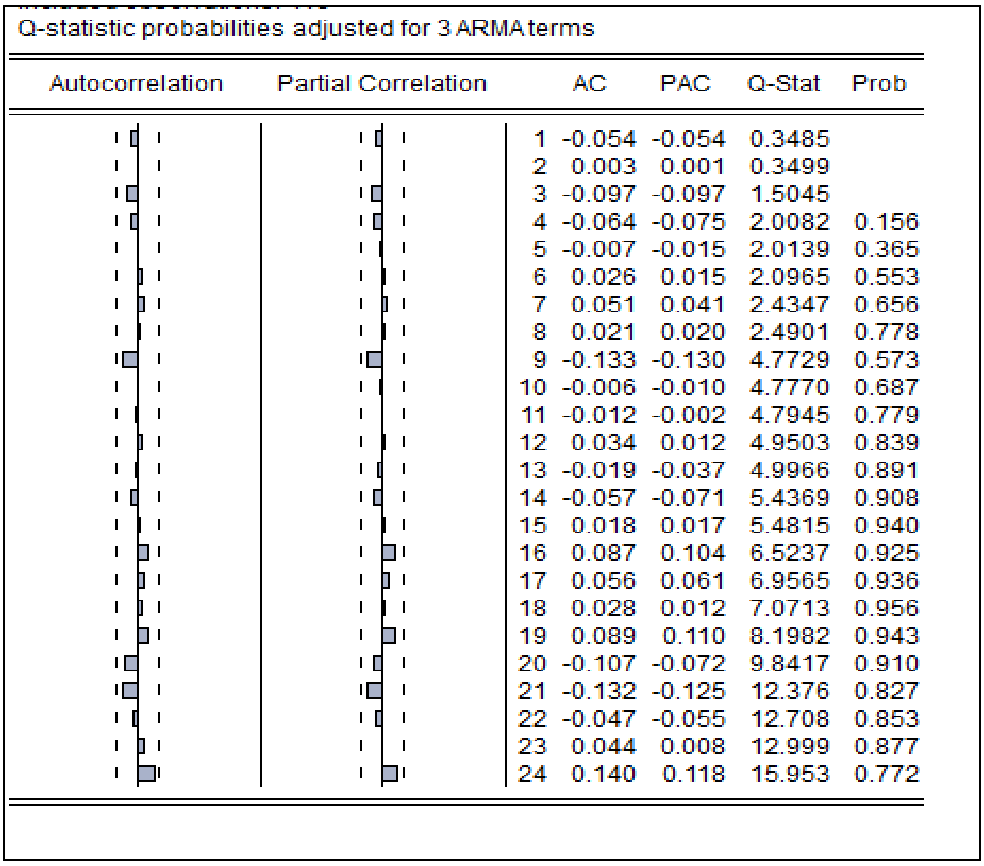
By drawing a functions AIC, PAIC, we note that some values deviate from the confidence limits at a significant level 0.05, exactly when in lag 4 and accordingly, models will be reconstructed based on the same previous criteria for selecting the best model and as shown in the Table 3.

Table 3. Some ARIMA models

	ARIMA(6,1,6)	AR(6)AR(4)MA(6)	AR (6) MA (6) MA(4)	AR(1)AR(4) A(1)
Significant coefficient	2	2	3	3
Sigma^2 (Volatility)	2.77E+15	2.90E+14	2.96E+14	3.93E+14
Adj R^2	0.268	0.28	0.31	0.16
AIC	38.4877	38.4788	38.4406	38.6166
SBIS	38.5816	38.5264	38.5580	38.7340

From the results of Table (3), it was found that the best model is AR (6) MA (6) MA (4). To choose the best, the time series of gas oil will be predicted based on the two models ARMA (6,1,6) adjusted [4] for 3ARIMA model {(AR(6), MA(6) MA(4)}.

From Figures 4 and 5, we find that the best model is 3 ARIMA(ar(6) ma(6) ma(4)) model modified based on the behavior of the function ACF, PACF, as we found from the analysis of the original and expected series graph that the difference between months from the first to the fourth of 2017 for the modified 3 ARIMA model is better ARIMA (6,1,6) and the difference is very little between the original and expected series, and from the fourth to the sixth month we find that the prediction using the ARIMA model (6,1,6) is better than the modified 3ARIMA model while noting the good prediction of the two series with the original series, While we find that the period from the sixth to the eighth month is a modified ARIMA 3 model better than the ARIMA (6,1,6) model, but the two models have moved away from A good instance of the series, again we have the best representation of the adjusted 3ARIMA model compared to ARIMA (6,1, 6) where we note the very small difference between the original and predicted series until it matches in some points. We conclude that the prediction, in general, is very good and can depend on the modified model for prediction.



Q-statistic probabilities adjusted for 3 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.054	-0.054	0.3485	
		2	0.003	0.001	0.3499	
		3	-0.097	-0.097	1.5045	
		4	-0.064	-0.075	2.0082	0.156
		5	-0.007	-0.015	2.0139	0.365
		6	0.026	0.015	2.0965	0.553
		7	0.051	0.041	2.4347	0.656
		8	0.021	0.020	2.4901	0.778
		9	-0.133	-0.130	4.7729	0.573
		10	-0.006	-0.010	4.7770	0.687
		11	-0.012	-0.002	4.7945	0.779
		12	0.034	0.012	4.9503	0.839
		13	-0.019	-0.037	4.9966	0.891
		14	-0.057	-0.071	5.4369	0.908
		15	0.018	0.017	5.4815	0.940
		16	0.087	0.104	6.5237	0.925
		17	0.056	0.061	6.9565	0.936
		18	0.028	0.012	7.0713	0.956
		19	0.089	0.110	8.1982	0.943
		20	-0.107	-0.072	9.8417	0.910
		21	-0.132	-0.125	12.376	0.827
		22	-0.047	-0.055	12.708	0.853
		23	0.044	0.008	12.999	0.877
		24	0.140	0.118	15.953	0.772

Figure 4. ACF, PACF for residual 3 ARIMA term

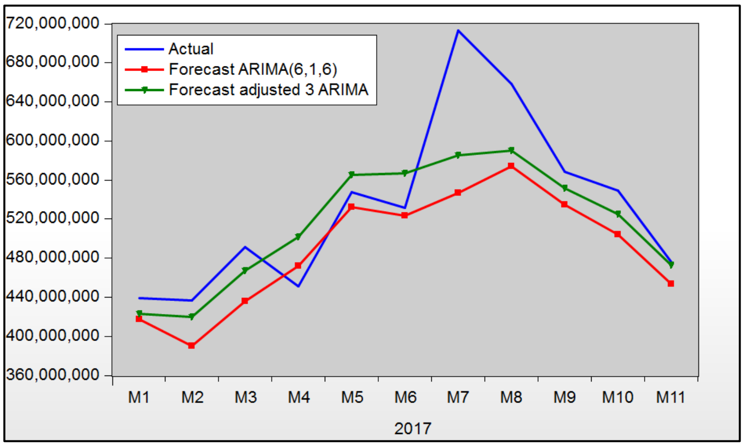


Figure 5. Graph of real and predictive values for models ARIMA (6,1,6) and adjusted 3 ARIMA for gas oil

5.2. Exponential smoothing

Exponential smoothing models were applied to the time series data for gas oil, where many models were built and we chose the best model based on some comparison criteria as shown in Table 4.

Table 4. The exponential smoothing patterns applied to the gas oil series

Model	RMSE	MAE	MAPE	BIC
Simple exponential smoothing	61893194	46221013	8.477	35.922
Browns liner exp. Smoothing	62149357	46187245	8.473	35.970
Holt linear Exp. smoothing	68679429	5154291244	9.466	36.130

Depending on the selection of the best model that has the lowest scale of error, we find that the single exponential smoothing model is the best as predictive values have been calculated for all of the coming months of the year 2018 with confidence limits for it as shown in table (5) as well as drawn as in Figure (6).

Table 5. The predictive values for the simple exponential smoothing of the gas oil series

Period	Forecast	Lower Limit 95%	Upper Limit 95%
Jan-18	483674029	319021846	648326212
Feb-18	483674029	285688142	681659916
Mar-18	483674029	257208840	710139218
Apr-18	483674029	231931002	735417056
May-18	483674029	208969420	758378638
Jun-18	483674029	187784367	779563691
Jul-18	483674029	168017947	799330111
Aug-18	483674029	149418390	817929668
Sep-18	483674029	131800610	835547448
Oct-18	483674029	115023825	852324233
Nov-18	483674029	98977989	868370069
Dec-18	483674029	83575150	883772908

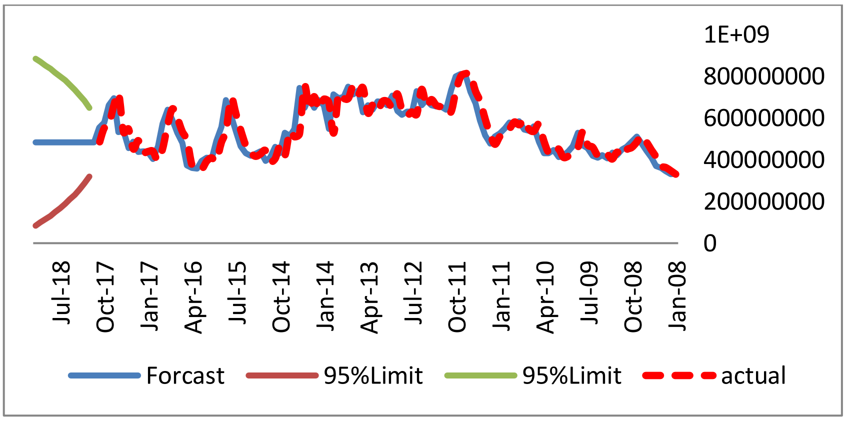


Figure 6. Predictive values for the best model to represent gas oil data

In addition to the above, multiple models were built from the random step model based on efficiency measures for these models. Models were described in the Table 6. By observing AIC standard, we find that the best model among the models ARIMA is ARIMA(6,1,6) that it has reached a value AIC for it 37.35674, and it must From mentioning that ACF, PACF function of ARIMA(6,1,6) model was all within the confines of trust, and this confirms the preference of the model. As for exponential smoothing models, the single exponential smoothing model is the best model based on the AIC standard.

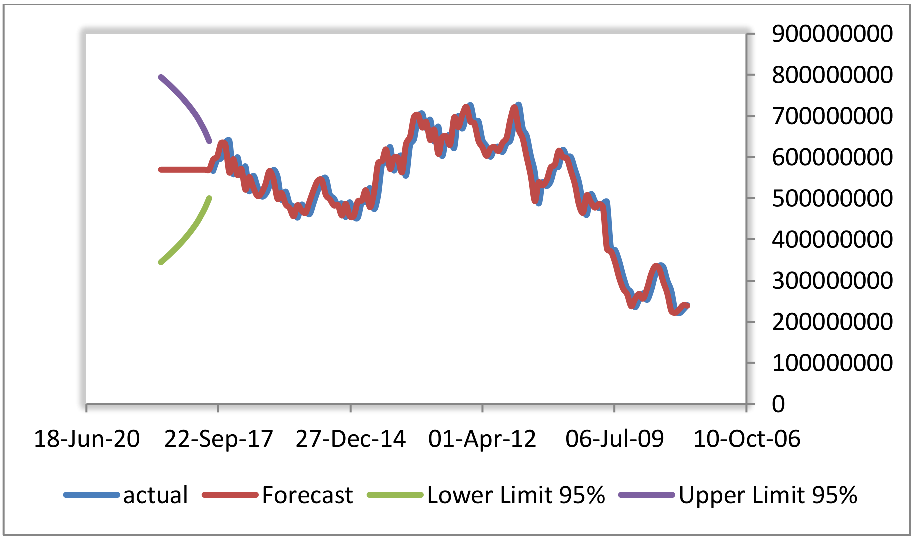
Table 6. The models applied to a series of excellent petrol

Model	RMSE	MAE	MAPE	AIC
ARIMA(2,1,3)	2.06E+08	1.70E08	29.63024	37.62
ARIMA(1,1,2)	1.84E+08	1.42E+08	24.48044	37.60257
ARIMA(6,1,2)	1.38E+08	1.13E+08	20.13429	37.52405
ARIMA(6,1,1)	1.41E+08	1.15E+08	20.27801	3.369412
ARIMA(6,1,6)	1.36E+08	1.13E+08	20.06519	37.35674
Simple exponential smoothing	3508388.24	28682308.41	23.143	34.787
Browns liner exp. smoothing	37873395.27	31466275.03	19.856	34.940
Holt linear Exp. smoothing	35115780.03	28635299.99	22.742	34.829

Depending on the best model, which has the least error and is the single exponential smoothing model, the predictive value has been calculated for the next 12 months of the year 2018 with confidence limits for it as shown in the Table 7 and Figure 7.

Table 7. The predictive values for the single exponential smoothing of the excellent petrol series

Period	Forecast	Lower Limit 95%	Upper Limit 95%
Jan-18	662454492	476551414	569502953
Feb-18	681096462	457909444	569502953
Mar-18	697042100	441963806	569502953
Apr-18	711204600	427801306	569502953
May-18	724074876	414931030	569502953
Jun-18	735952952	403052954	569502953
Jul-18	747038090	391967816	569502953
Aug-18	757470630	381535276	569502953
Sep-18	767353832	371652074	569502953
Oct-18	776766295	362239611	569502953
Nov-18	785769491	353236415	569502953
Dec-18	794412575	344593331	569502953



Figurer 7. Predictive values for the best model (simple exponential smoothing) to represent excellent petrol

6. Conclusions

1) It turned out that both the gas oil and the excellent gasoline chains were unstable as a general trend was observed in the two chains and they were converted to a stable series after taking the first differences.
2) After addressing the instability, ARIMA(6,1,6) the model was chosen to represent gas oil data, because this model has the lowest value of the Akaike standards and the BIS information for Schwartz, compared to the Box Jenkins models that are appropriate, and therefore it is the most appropriate model for the series to use for prediction.
3) The model was reconstructed based on the error behaviour of ACF and PACF for ARIMA(6,1,6) model, the adjusted 3 ARIMA term was chosen because it had the lowest value for the Akaike and Schwartz parameters, and was used for purpose of prediction which proved to be preferable forecast from ARIMA(6,1,6) model.

4) Through the application of exponential smoothing methods in research, it was noted that the single exponential smoothing method was superior to other methods adopted in the research, where results were obtained with the lowest values for the statistical standards used to calculate prediction errors, so this method is the best way to predict future values.
5) Through the models applied to the improved gasoline chain, it was noted that the single exponential smoothing method was superior to other methods used in the research (Box Jenkins models, Holt linear Exp. smoothing, Browns liner exp. smoothing) where the lowest values were obtained for the statistical standards used, so this method is the best way to predict future values of a series Excellent petrol.

References

[1] S. Makridakis, S.C. Wheelwright and V. E. McGee, Forecasting: Method and Applications, 2nd ed, New York: Wiley, 1983.

[2] M. Wooldridge, Introductory Econometrics: A Modern Approach, South-Western, 5th ed, 2013.

[3] J. Heizer and B. Render, Principles of operations management, 6th ed, Prentice-Hall, USA, 2001.

[4] N. Gujarati, Basic Econometrics, McGraw- Hill Higher Educations, 4th ed, 2003.

[5] G. Box, and G.M. Jenkins, Time series Analysis Forecasting and Control, Revised Edition, Holden-Day, 1976.

[6] S. Shuhaili and A. Saba," Comparing the Univariate Modeling Techniques, Box-Jenkins and Artificial Neural Network (ANN) for Measuring of Climate Index," *Applied Mathematical Sciences*, vol 8, no 32, pp 1557-1568, 2014.

[7] G. Box, G. M. Jenkins and G.C. Reinsel, Time Series Analysis and Control, 3rd ed,Holden-Day, San Francisco, 2008.

[8] S. Makridakis, S.C Wheelwight and R.J. Hyadman, Forcasting Methods and Applications, 3rd ed, John Wiley and Sons, Inc, USA,1998.

[9] M. K. Newas, "Comparing the Performance of Time Series Models for Forcasting Exchange Rate," *BRAC, University journal*, vol v, no 2, pp 55-65, 2008.

[10] L.M. Liu, Time Series Analysis and Forcasting, 2nd ed, Scientific Computing Associates Grop,USA, 2006.

[11] W. S. Wei, Time Series Analysis: Univarate and multivariate Methods, Addition Wesley, 1990.

[12] E. Stawan, N. Herawati and K. Nisa, "Modeling Stock Return Data Using Asymmetric Volatility Models: A performance Comparison based on the Akaike Information Criterion and Schwarz Criterion," *Journal of Engineering and Scientific Research (JESR)*, vol 1, Issue 1, pp 37-41, June 2019.

[13] R. Nochai and T. Nochai, ARIMA model for forecasting oil palm price. In *Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and applications*, pp. 13-15, 2006.

[14] K. Ryu and A. Sanches," The Evaluation of Forecasting Methods at an Institutional Foodservice Dining Facility," *Journal of Hospitality Financial Management*, vol. 11, no. 1, 2003.

[15] G. Box, G. M. Jenkins and G. C. Reinsel, Time Series Analysis and Control, 3rd ed, Holden-Day, San Francisco ,1999.

[16] M. Alysha, J. Rob and D. Ralph, "Forcasting Time Series with Complex Seasonal Patterns Using Exponential Smoothing," Monash university, 2010.