UIC UNIVERSITY OF ILLINOIS AT CHICAGO

Time Series Analysis & Forecasting of US Petroleum Import from Kuwait

Siddhesh Tiwari

30th November 2015

ABSTRACT

Petroleum, a commodity, on which majority of the world is dependent is on the verge of extinction. This calls for proper management and utilization of petroleum in its usage. Countries, for example, United States of America, import oil and petroleum from the Middle East. If the amount of import can be forecasted before-hand, it would make it easier for countries to develop strategies and policies for smooth functioning of its relationship. Moreover, this would also affect the exchange rates in-between the countries involved. Our analysis involves ARMA modelling to be performed on the known import data, to forecast the future amount. Our approach involves removal of deterministic trend from the data and using the stationary residuals for ARMA (2n, 2n-1) modelling procedure. On successful execution of the modelling procedure, the future values were forecasted. This forecasted values were compared with different other approaches to determine its significance.

`

PROJECT OVERVIEW

Dataset: Our project deals with analysis of petroleum import of U.S. from Kuwait from 15th January 2000 to 15th April 2014, collected by U.S. Census Bureau, Foreign Trade Division.

Steps of Analysis:

- To summarize data for interpretation
- To find the trend in the data, followed by de-trending
- To identify the adequate ARMA model for forecasting
- Comparison with Exponential, Holt's Approach and Holt-Winter's Approach.

Software Used: R and Excel (XLMiner)

We followed the following steps in our project:

1. Identifying the trend followed by the data, checking residuals and ACF and PACF of the stochastic data.
2. Identifying the adequate order of the ARMA (n, m) model, using the standard (2n, 2n-1) approach.
3. Joint optimization of integrated model and check for stochastic trend & seasonality.
4. Forecasting
5. Comparing with other forecasting approaches:
    a) Exponential Forecasting
    b) Double Exponential Forecasting
    c) Holts-Winter Forecasting

Trend line fitting, ARMA modeling and joint optimization was done in R. Forecasting and comparison with other forecasting techniques was done using MS-Excel.

All the supporting graphs are attached in further parts of the report.

Introduction

Oil is a critical resource for U.S. economy. It meets nearly 40% of total U.S. energy requirements, including 94% of energy used in transportation and 41% of energy used by the industrial sector. Unlike other forms of energy such as coal and natural gas, which are largely supplied from domestic sources, half of U.S. oil consumption is currently supplied from foreign sources.

The United States consumed 18.6 million barrels per day (MMbd) of petroleum products during 2012, making it the world's largest petroleum consumer. The United States was third in crude oil production at 6.5 MMbd. Crude oil alone, however, does not constitute all U.S. petroleum supplies. Significant gains occur because crude oil expands in the refining process, liquid fuel is captured in the processing of natural gas, and we have other sources of liquid fuel, including Bio fuels. These additional supplies totaled 4.8 MMbd in 2012. The United States imported 11.0 MMbd of crude oil and refined petroleum products in 2012. It also exported 3.2 MMbd of crude oil and petroleum products, so it’s net imports (imports minus exports) equaled 7.4 MMbd.

In 2012, the United States imported 2.1 MMbd of petroleum products such as gasoline, diesel fuel, heating oil, jet fuel, and other products while exporting 3.1 MMbd of products, making the United States a net exporter of petroleum products. The U.S. Census Bureau's Foreign Trade Division (FTD) formulates, develops and implements plans and programs for the collection, processing and dissemination of statistical data relating to the United States’ merchandise trade with foreign countries and U.S. possessions.

By obtaining the data from U.S. Census Bureau's Foreign Trade Division (FTD), the U.S. monthly petroleum imports in the time period 2000-2014 has been analyzed.

At the end, the forecasted results are compared with different approaches namely Exponential Smoothing Forecast, Holt’s Forecast and Holt-Winter’s Forecast.

1. Forecasting Techniques

1.1 Autoregressive Models

We say that the process $\{x_t\}$ is autoregressive of order p (AR(p)) if there exist constants $a_1, \ldots,$ a_p such that

$$x_t = \sum_{k=1}^{p} a_k x_{t-k} + \varepsilon_t$$

where $\{\varepsilon_t\}$ is zero-mean white noise, and ε_t is uncorrelated with $x_{t-1}, x_{t-2} \cdots$. The AR (p) process exists and is weakly stationary if and only if all the roots of the polynomial $P(z)=1-a_1 z \cdots -a_p z^p$ lie outside the unit circle in the complex plane. An example of an autoregressive process which is not stationary is the random walk, $x_t = x_{t-1} + \varepsilon_t$, which is an AR (1) with $a_1=1$. We can verify that the random walk is not stationary by noting that the root of $P(z)=1-z$ is $z=1$, which is on the unit circle.

1.2 Moving Average Models

We say that $\{x_t\}$ is a moving average of order q ($MA(q)$) if there exist constants $b_1, \ldots, b_q$ such that

$$x_t = \sum_{k=0}^{q} b_k \varepsilon_{t-k}$$

where $b_0=1$. The $MA(q)$ process has a finite memory, in the sense that observations spaced more than q time units apart are uncorrelated. In other words, the autocovariance function of the $MA(q)$ cuts off beyond lag q. This contrasts with the autocovariance function of the AR (p), which decays exponentially, but does not cut off.

1.3 Autoregressive Moving Average Models

We say that $\{x_t\}$ is an autoregressive moving average process of order p ,q ($ARMA$ (p ,q)) if there exist constants $a_1, \ldots, a_p$, $b_1, \ldots, b_q$ such that

$$x_t = \sum_{j=1}^{p} a_j x_{t-j} + \varepsilon_t + \sum_{j=1}^{q} b_k \varepsilon_{t-k}$$

where $\{\varepsilon_t\}$ is zero-mean white noise, and ε_t is uncorrelated with $x_{t-1}, x_{t-2} \cdots$. The *ARMA* (*p* ,*q*) process exists and is weakly stationary if and only if all the roots of the polynomial $P(z)$ defined earlier are outside the unit circle. The process is said to be invertible if all the roots of the polynomial $Q(z)=1+b_1z+\cdots+b_qz^q$ lie outside the unit circle. A time series is invertible if and only if it has an infinite-order autoregressive (*AR* (∞)) representation of the form

$$x_t = \sum_{j=1}^{\infty} \pi_j x_{t-j} + \varepsilon_t$$

where π_j are constants with $\Sigma\pi_j^2<\infty$.

1.4 Exponential Smoothing

The simplest of the exponentially smoothing methods is naturally called simple exponential smoothing (SES). This method is suitable for forecasting data with no trend or seasonal pattern.

$$S_t = \alpha X_{\tilde{t}} + (1-\alpha)S_{t-1}$$
$$F_t = S_t$$

Since single exponential smoothing method is a one period ahead forecaster, it is modified and the actual demand of the most recently known period is used for future periods' forecast.

1.5 Holt's Method

Holt extended simple exponential smoothing to allow forecasting of data with a trend. This method involves a forecast equation and two smoothing equations i.e. one for the level and one for the trend.

$$F_t = a_{\tilde{t}} + kb_{\tilde{t}}$$
$$a_t = X_t + (1-\alpha)^2 e_t$$
$$b_t = b_{t-1} - \alpha^2 e_t$$
$$e_t = F_{t-1}(1) - X_t$$

Similar to single exponential smoothing method, the double exponential smoothing method (Holt's Method) is a one period ahead forecast, the last known period's a and b values for future periods' forecast are used.

1.6 Holt-Winter's Method

Holt and Winters extended Holt's method to capture seasonality. The Holt-Winters seasonal method comprises the forecast equation and three smoothing equations — one for the level, one for trend and one for the seasonal component with smoothing parameters.

$$F_t(k) = a_t + b_t k + 0.5 c_t k^2$$
$$a_t = X_t + (1-\alpha)^3 e_t$$
$$b_t = b_{t-1} + c_{t-1} - 1.5\alpha^2 (2-\alpha) e_t$$
$$c_t = c_{t-1} - \alpha^3 e_t$$

- $\mathbf{X_t}$ is the actual demand for period t.
- $\mathbf{F_t}$ is the forecast demand for period t.
- $\boldsymbol{\alpha}$ is the smoothing coefficient.
- **k** is the number of periods ahead forecast.

There are two variations to this method that differ in the nature of the seasonal component. The additive method is preferred when the seasonal variations are roughly constant through the series, while the multiplicative method is preferred when the seasonal variations are changing proportional to the level of the series. With the additive method, the seasonal component is expressed in absolute terms in the scale of the observed series, and in the level equation the series is seasonally adjusted by subtracting the seasonal component. Within each year the seasonal component will add up to approximately zero. With the multiplicative method, the seasonal component is expressed in relative terms and the series is seasonally adjusted by dividing through by the seasonal component.

Basic Programs for Production and Operations Management," by Pantumsinchai, P., Hassan, M. Z., and Gupta, I. D., Prentice-Hall, 1983, pp. 65-66.

http://people.stern.nyu.edu/churvich/TimeSeries/Handouts/Intro.pdf.

2. Time Series Analysis

2.1 Evaluation of Raw Data

The raw data consisted of 172 values i.e. 172 monthly data beginning from 15th January 2000 to 15th April 2014. For our analysis and modelling procedure, we chose 169 values for training data and 3 values as validation data. The training data was plotted and after analyzing and looking at the data, it seemed to have a polynomial trend. The ACF plot of the raw data also indicated presence of a deterministic trend in data

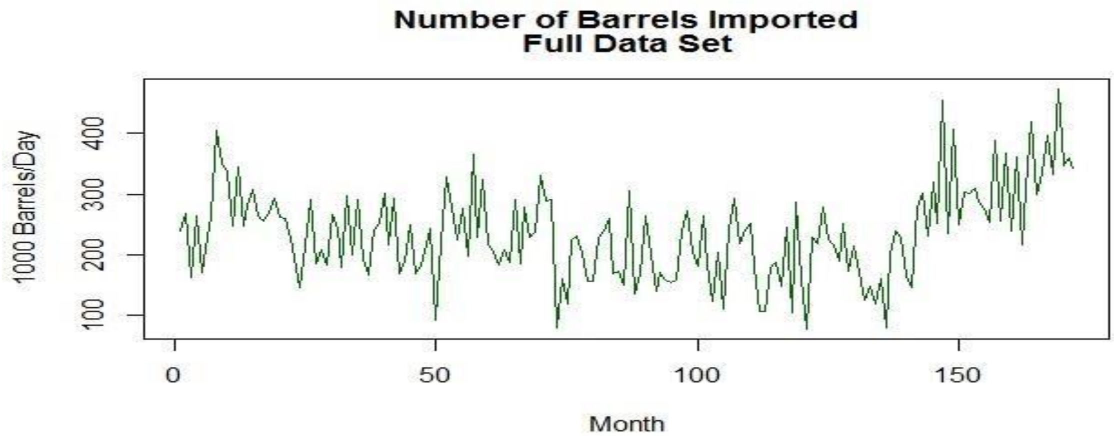


Figure 2.1 Full Data

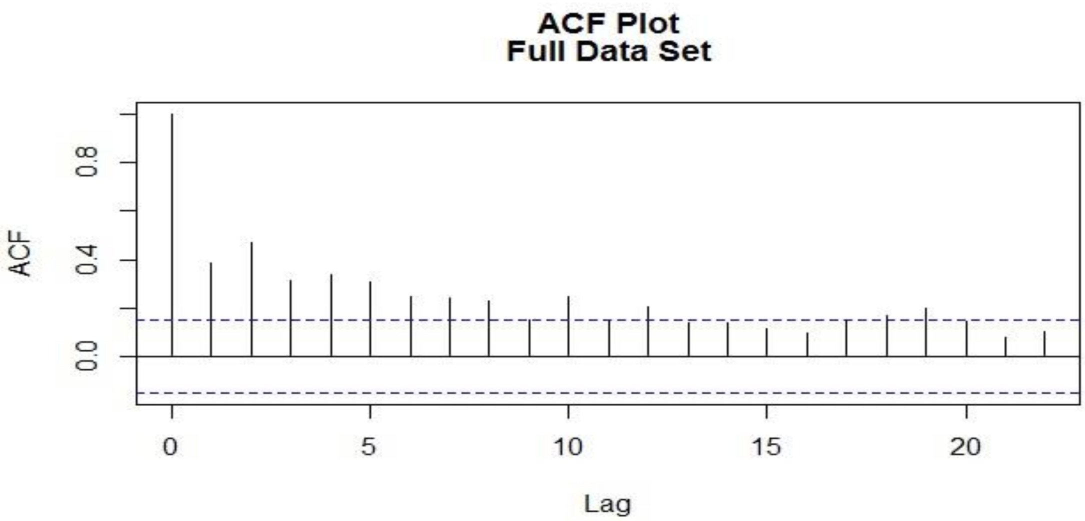


Figure 2.2 ACF of Full Data

2.2 Evaluation of Trend

The training data showed a deterministic trend of a polynomial order. Therefore, using R, various polynomial order were fitted into the training data to determine the best fitting polynomial. This was done using the RSS values and performing the F-test between various polynomial orders.

Table 2.1 Polynomial Order Equations

Order	Equation	RSS
Linear	$y = 0.1271x + 222.1513$	911744.6
Polynomial 2	$y = 0.01669x^2 - 2.71056x + 303.0240$	698390.4
Polynomial 3	$y= 0.000279x^3 – 0.041711x^2 + 1.2725x + 245.7658$	624663.7
Polynomial 4	$y= 1.965E\text{-}6x^4 – 4.39E\text{-}4x^3 + 0.03144x^2 - 1.51x + 270.35723$	614830.5
Polynomial 5	$y= -9.948E\text{-}9x^5 + 6.193E\text{-}6x^4 – 1.079E\text{-}4x^3 + 7.241E\text{-}2x^2 – 2.517x + 276.24615$	614376.3

Table 2.2 F-Statistics of Polynomial Orders

Polynomial Order	F-Statistic	F-Critical
1	-	2.3719
1-2	48.268	2.3719
2-3	18.5301	2.3719
3-4	2.4941	2.3719
4-5	0.1145	2.3719

The various equations for the polynomial orders were obtained along with their respective RSS values. These values were used to perform the F-test to check whether there was any improvement in RSS when moving to higher order polynomial. Firstly, F-test was done on increasing order from polynomial 1-2 followed 2-3,3-4 until polynomial 4 and polynomial 5, where F-test showed insignificant improvement in RSS, where we decided to stop. Polynomial order of 4 was considered as best fit for the deterministic trend of data. All the F-test for project were performed at 0.05 level of significance.

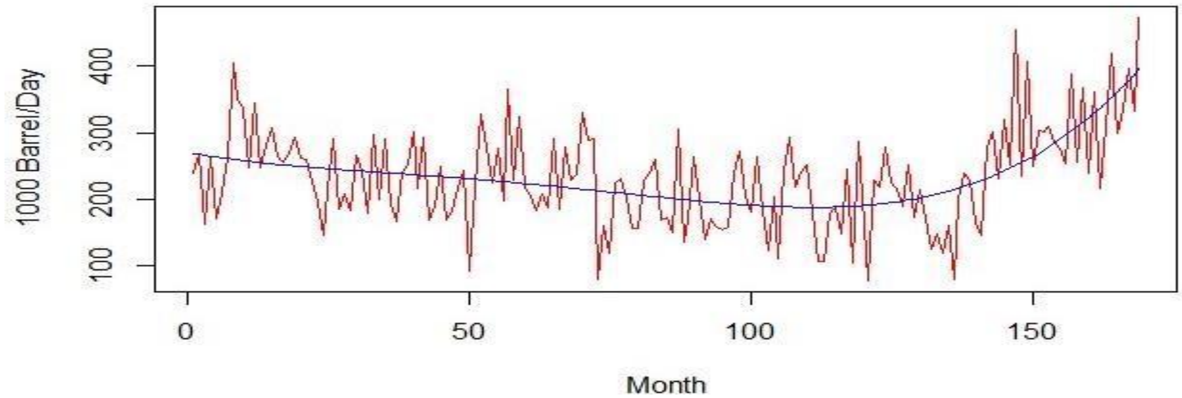


Figure 2.3 Trend Line Fitting

$$Y_t = f(t) + X_t \quad f(t) = 1.965E\text{-}6t^4 - 4.39E\text{-}4t^3 + 0.03144t^2 - 1.51t + 270.35723$$

After obtaining the equation of the trend, residuals were obtained for ARMA modelling. The following graphs show the stochastic nature of residual making them fit for ARMA modeling procedure.

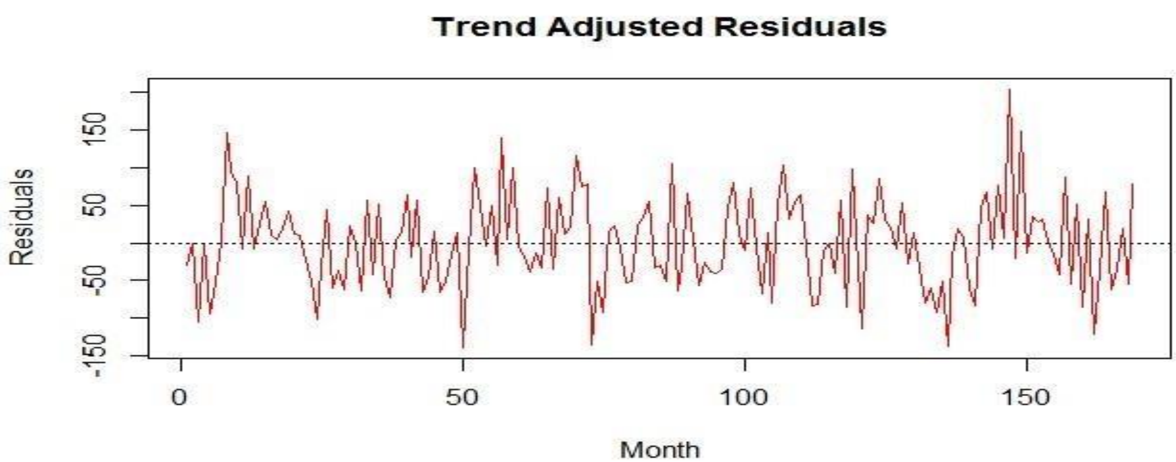


Figure 2.4 Residuals Plot

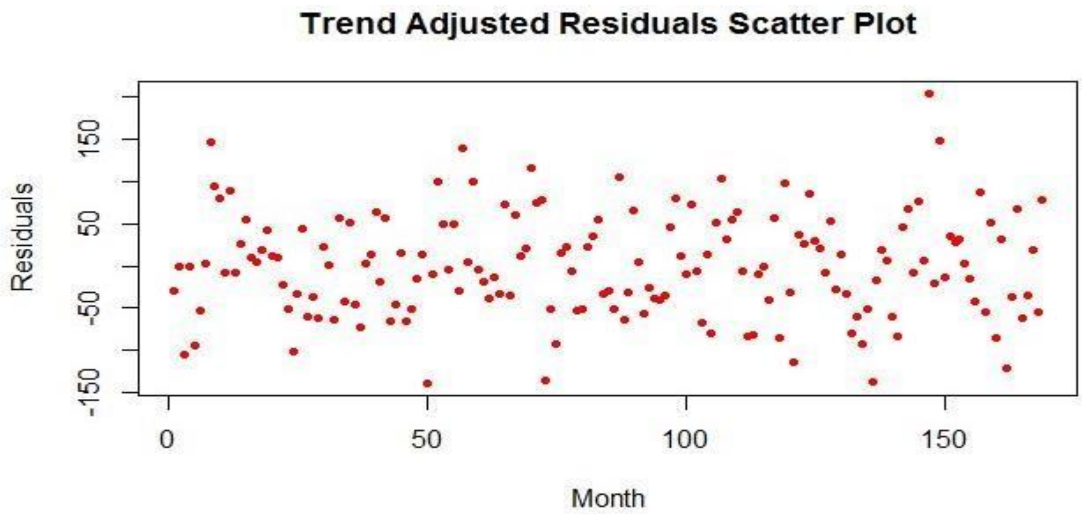


Figure 2.5 Residuals Scatter Plot

To further strengthen our believe regarding the stochastic nature of residual we plotted the autocorrelation and partial autocorrelation graph of residuals and compared it with the ACF of full data set.

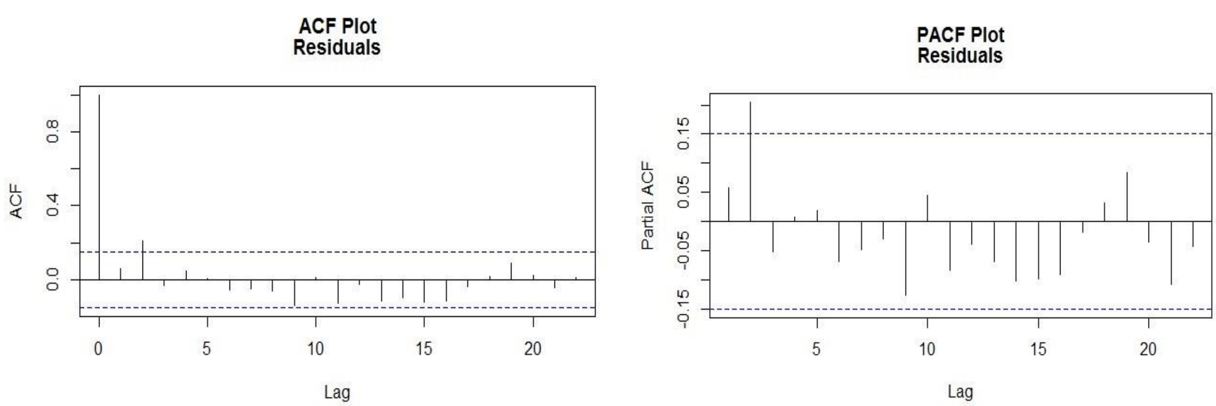


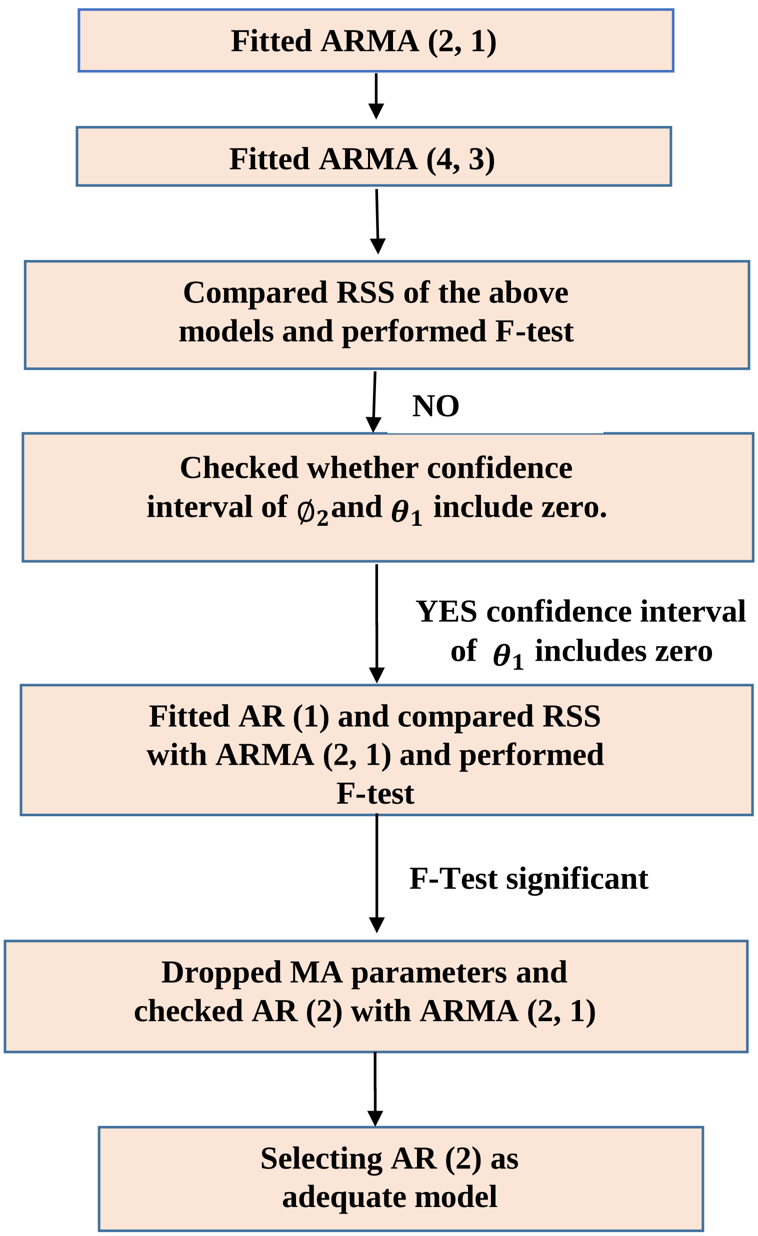
Figure 2.6 ACF & PACF Plot of Residuals

On analyzing the above ACF and PACF figure, it shows that there is no significant autocorrelation after the 2^{nd} lag. Also there is no significant partial autocorrelation after 2^{nd} lag. This points towards a lower ARMA model. Moreover, the ACF & PACF plots lie within the confidence intervals.

2.3 ARMA(2n,2n-1) Modelling

Following is the procedure used for ARMA modelling:

Start with n=1, which gives ARMA(2n,2n-1) = ARMA(2,1)



The residuals obtained after fitting of 4th order polynomial trend are then subjected to ARMA(2n,2n-1) modelling procedure. We start by inputting the value of n=1 to obtain ARMA(2,1). This ARMA(2,1) was fitted to the residual values and its parameters were obtained.

This was followed by inputting n=2 to obtain ARMA(4,3). Similarly, this ARMA(4,3) was fitted to the residual values and its parameters were also obtained.

Table 2.3 Parameters for ARMA Models

Parameters	ARMA(2,1)	ARMA(4,3)	AR(2)	AR(1)
ϕ_1	-0.1786 ± 0.5292	-1.1058 ± 0.7155	0.0446 ± 0.1523	0.0581 ± 0.1509
ϕ_2	0.2196 ± 0.1519	-0.7323 ± 0.7826	0.2057 ± 0.1524	
ϕ_3		0.1910 ± 0.6321		
ϕ_4		0.1827 ± 0.1599		
θ_1	0.2334 ± 0.5332	1.1909 ± 0.7231		
θ_2		1..0524 ± 0.8359		
θ_3		0.0457 ± .7277		
RSS	5845385.5	555111.2	585894.6	612754.1
N	160	160	160	160
AIC	1866.96	1869.49	1865.43	1870.7
Sigma^2	3461	3285	3471	3626
Loglikelihood	-928.48	-925.74	-928.72	-932.35

Table 2.4 F-Statistics for ARMA Models

	ARMA(2,1)-(4,3)	AR(2,1)-(1)	ARMA(2,1)-(2)
F-Statistic	2.0450	3.7026	0.449
F-Critical	2.3719	2.3719	2.3719

After obtaining parameters for ARMA(2,1) and ARMA(4,3) and their respective RSS values, F-Test was conducted between the two models which showed insignificant improvement in RSS. Therefore, ARMA(2,1) model was shown adequate at this particular point. Upon further observation of ARMA(2,1) model, we found that their parameters included zero in their confidence interval. This led to reduction of model to ARMA(2n-1,2n-2) and then performing the F-Test. This meant performing F-Test between ARMA(2,1) and AR(1). F-Test between ARMA (2,1) and AR (1) showed significance increase in RSS. On further proceeding with the ARMA(2n,2n-1) modelling procedure, it led to an adequate AR(2) model, since small MA parameters had to be drop from ARMA(2n,2n-1) model due to significant increase in RSS shown in F-Test. Therefore AR(2) model was found as the adequate model.

$$\mathbf{X_t = 0.0446X_{t-1} + 0.2057X_{t-2} + a_t}$$

2.4 Joint Optimization of the Combined Model

Presently we have two models, one pertaining to the deterministic trend in the data and the other pertaining to the stochastic nature of the residuals i.e. polynomial trend and AR(2) model. Now joint optimization is performed on the complete integrated model. The nls(non-linear least square) function of R was used for joint optimization which minimizes the sum of squared residuals to obtain the optimized parameters.

Table 3.1 Joint Optimization

Parameter	Initial Est.	Joint Estimate	Std. Error
ϕ_1	0.0446	0.04488	0.07782
ϕ_2	0.2057	0.2079	0.07791
β_0	270	276.2	27.98
β_1	1.51	1.956	2.181
β_2	0.0314	0.041	0.05052
β_3	0.000439	0.0005172	0.0004383
β_4	1.97E-06	0.000002181	0.000001265
RSS	585894.584	585644	

On analyzing the complete model parameters it can be observed that there is not much difference between the initial estimates and the joint optimization estimated values. Also, there is no huge difference between the RSS values.

We then had to calculate the characteristic roots of our jointly optimized parameters to check for presence of any stochastic trend or seasonality. We obtained following values of our roots

$$\boldsymbol{\lambda}_{\mathbf{1}} = 0.4791 \quad \boldsymbol{\lambda}_{\mathbf{2}} = -0.4343$$

The roots clearly indicate that there is not presence of stochastic seasonality or trend as they are neither close to unit nor they are a complex conjugate.

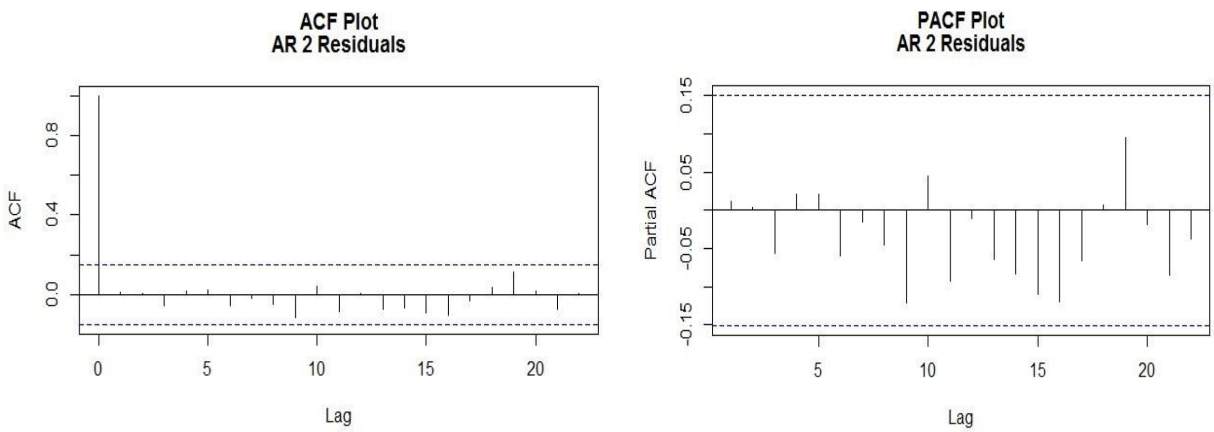


Figure 2.7 ACF & PACF Plots for AR(2) Model

The above mentioned figure of ACF & PACF plots for the adequate AR(2) Model were found to be within confidence intervals. This indicates that AR(2) proves to be a good fit for the residual data.

Following is the graph obtained for our fitted AR(2) value on the residual data. The blue line show the fitted AR (2) values and the red line shows the residual obtained from fitted polynomial trend. The fitted values does not demonstrate an excellent fit but is able to capture the fluctuation around the mean.

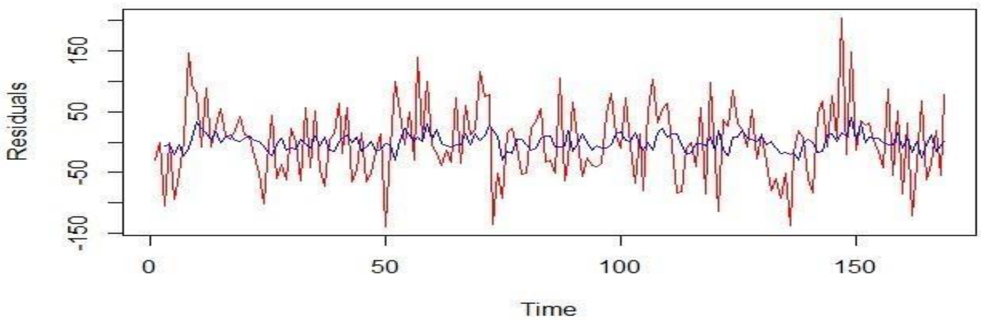


Figure 2.8 X_t vs AR(2) - a_t

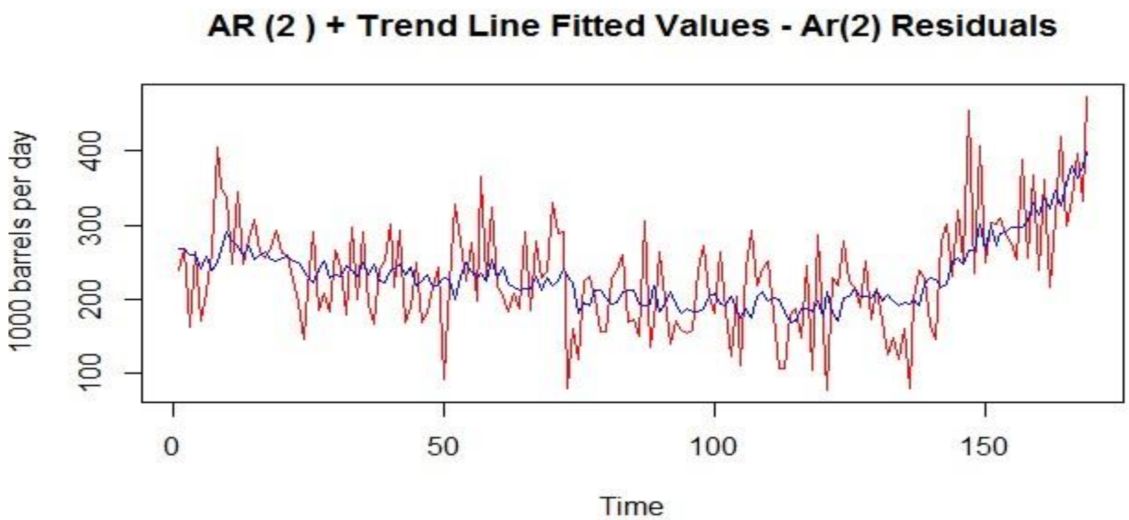


Fig 2.9 Complete model vs Raw Data

The above graph shows the AR(2) model superimposed on the fitted trend line which makes our complete model and is shown in blue color which makes our complete model. The raw data is plotted in red.

3. Forecasting

Forecasting is done on the whole model, i.e. combining the trend line equation along with the obtained ARMA model, which is, AR(2). It was done on excel using the traditional formulas given in the course book.

The whole model on which forecasting is to be performed is

$$\mathbf{Y(t) = f(t) + X(t)}$$

$$\mathbf{f(t)} = 1.965E\text{-}6\ t^4 – 4.39E\text{-}4\ t^3 + 0.03144\ t^2 - 1.51\ t + 270.35723$$

$$\mathbf{X_t} = 0.04488\ X_{t-1} + 0.2079\ X_{t-2} + a_t$$

$$\mathbf{Y(t)} = 2.181E\text{-}6\ t^4 – 5.172\text{-}4\ t^3 + 0.041\ t^2\text{-}1.956\ t + 276.2 + 0.04488\ X_{t-1} + 0.2079\ X_{t-2} + a_t$$

Forecasting of validation data was performed, i.e. modelling was performed on 169 data points, whereas forecasting was performed on the final 3 validating data points.

Time	X_t Forecast	Trend Line Forecast	Actual Forecast	Raw Values	Upper Limit	Lower Limit
170	-8.01997	409.16941	401.1494	378	516.6303	285.6685
171	15.73778	419.329563	435.0673	390	550.6641	319.4706
172	-0.9478	429.80637	428.8586	362	546.1494	311.5678

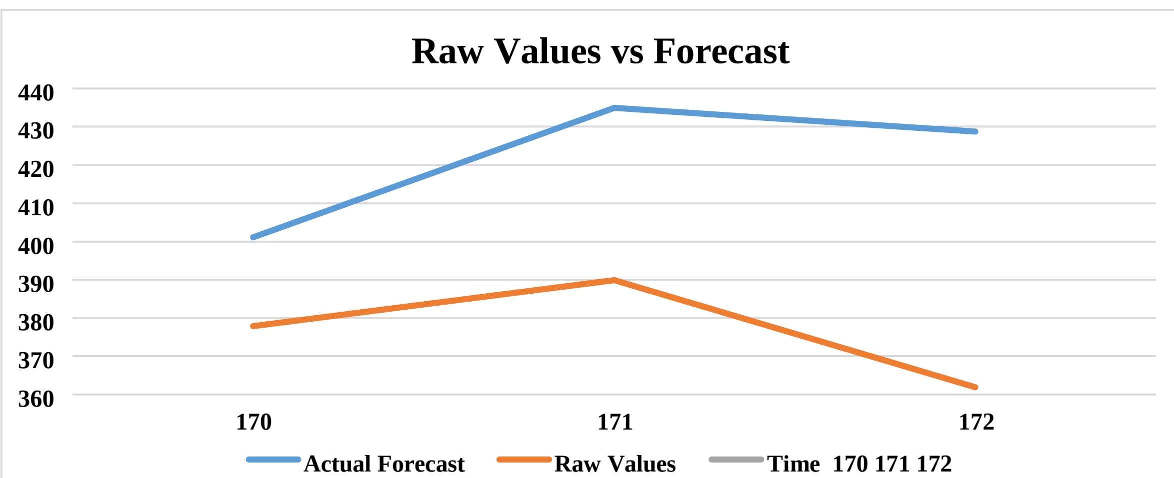


Figure 3.1 Forecast vs Raw Values

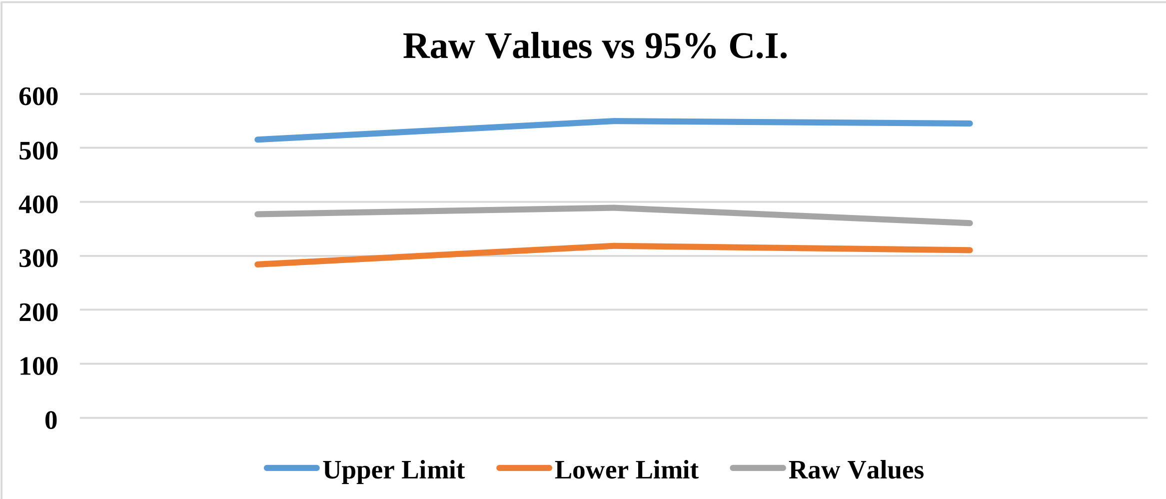


Figure 3.2 Raw Values and 95% Confidence Interval

The above two graphs plotted show the forecasted values and also the confidence interval plot. The forecasted values follow the same path as the raw values but depicts a difference in its magnitude. In the next graph plotted, it is seen that even though there is difference between the forecasted and raw values, they fall within the confidence interval limits which shows that the forecasting can be accepted.

Further, we will compare this forecasting result with some other smoothing techniques used for forecast with the help of XLMiner Tool.

3.1 Single Exponential Forecasting

The description of this concept was explained before in this report. On forecasting using this technique, the results were found are shown below.

Time	Actual	Forecast	Error	LCI	UCI
170	378	362.5887	15.4113	237.6777	487.4997
171	390	362.5887	27.4113	235.1509	490.0265
172	362	362.5887	-0.5887	232.6733	492.5041

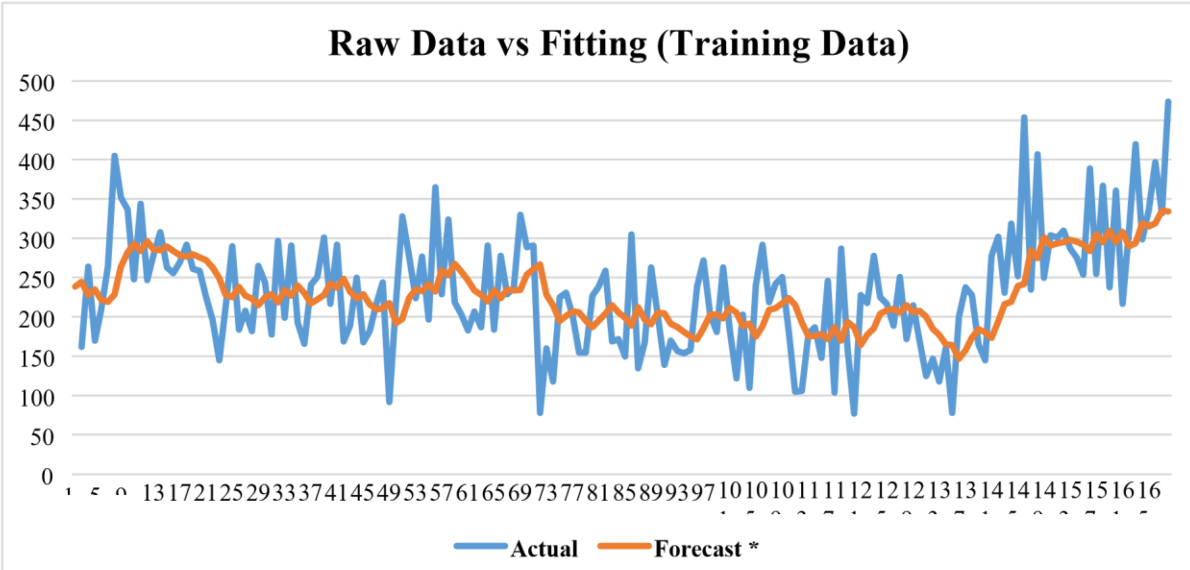


Figure 3.3 Single Exponential Training Data

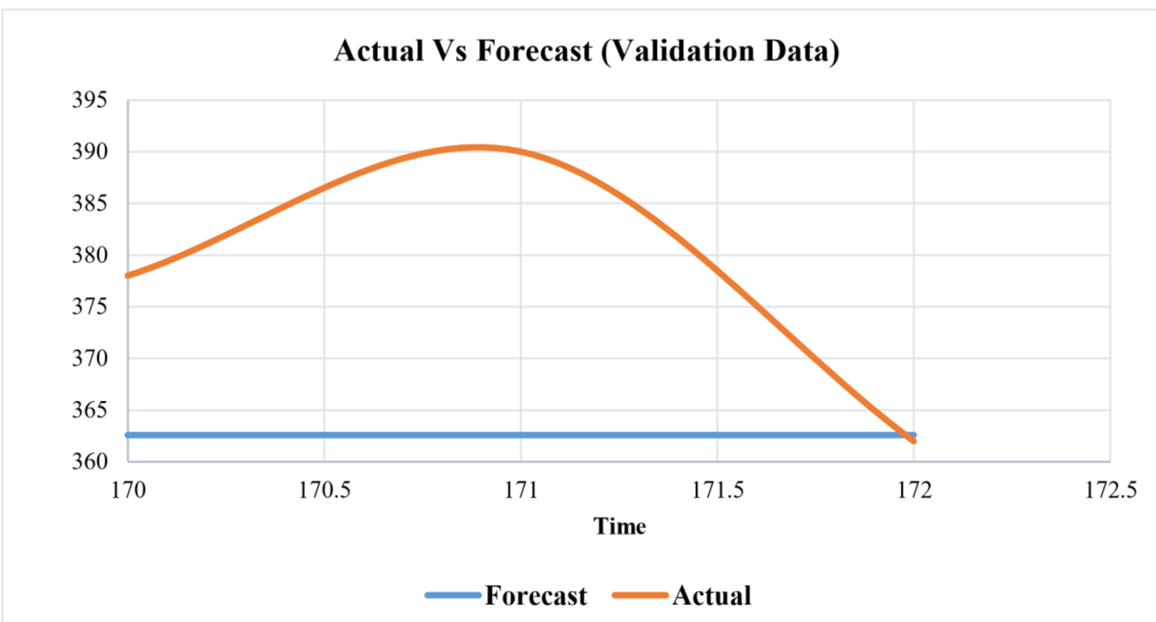


Figure 3.4 Single Exponential Validation Data

The above two graphs plotted shows forecasting performed using single exponential smoothing. Since single exponential forecasting works well for data which does not involve trend nor seasonality, therefore the result is not surprising. Moreover, single exponential smoothing only forecasts levels therefore we obtain forecast just as a single line which will give same forecast at any time.

3.2 Holt's Method

Double Exponential Smoothing is concept which works well for data which includes trend only but no seasonality. Hence the further obtained forecast would work well for the current dataset.

Time	Actual	Forecast	Error	LCI	UCI
170	378	370.3452	7.654811	245.067	495.6234
171	390	372.1348	17.86517	244.0941	500.1755
172	362	373.9245	-11.9245	243.1049	504.744

Upon calculation of confidence interval i.e. upper and lower limit, it is seen that the original values fall within the confidence limits.

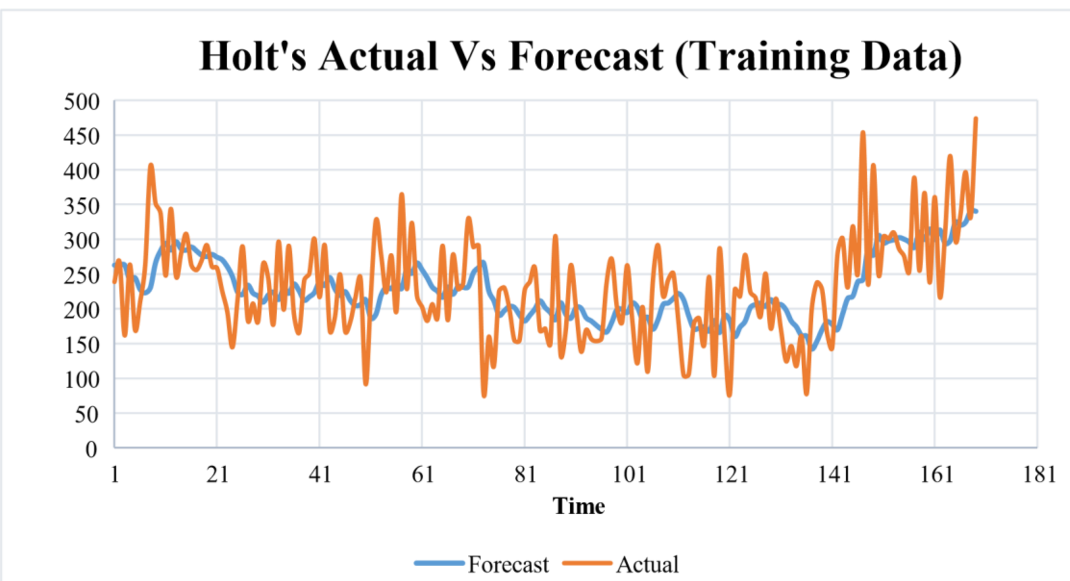


Figure 3.5
Holt's
Training Data

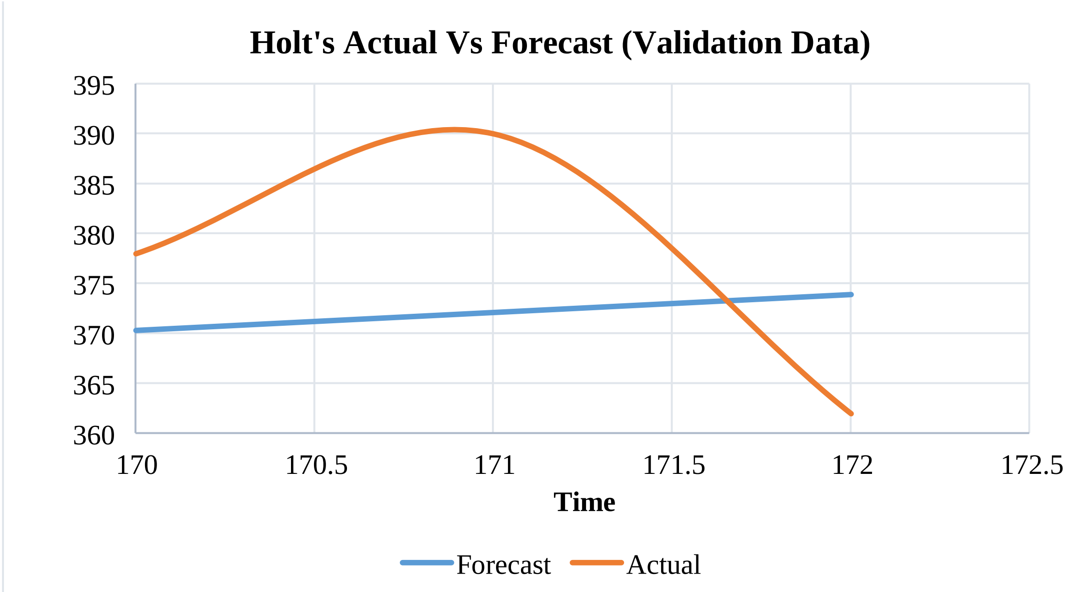


Figure 3.6 **Holt's Validation Data**

The forecast obtained from double exponential is a linearly increasing line. Inspite of giving very low error on forecast this model fails to identify any fluctuation that might occur in forecast. The linear forecast line does not depict any pattern of the validation data which was clearly modelled by ARMA in spite of having slightly high prediction error.

3.3 Holt-Winter's Method

Holt-Winter's approach suits well for data which include trend as well as seasonality in it. There are two variations, additive as well as multiplicative. The following approach use additive HoltWinter's method.

Time	Actual	Forecast	Error	LCI	UCI
170	378	363.2672	14.73281	232.2068	494.3275
171	390	420.2251	-30.2251	285.7537	554.6965
172	362	348.2153	13.78473	210.1317	486.2988

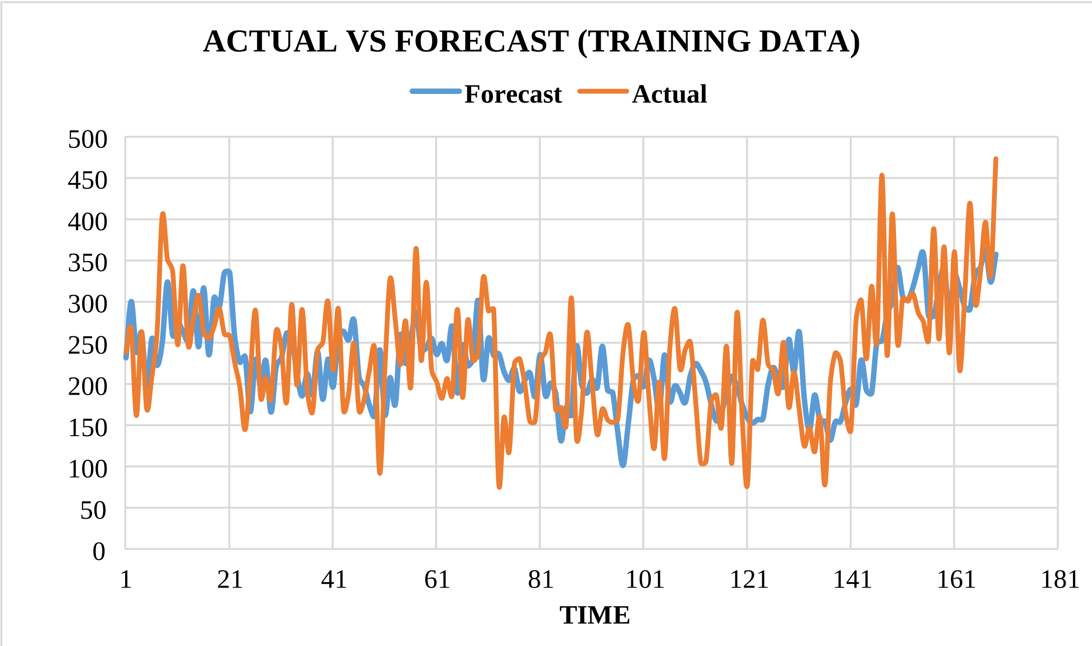


Figure 3.7 Holt-Winter's Training Data

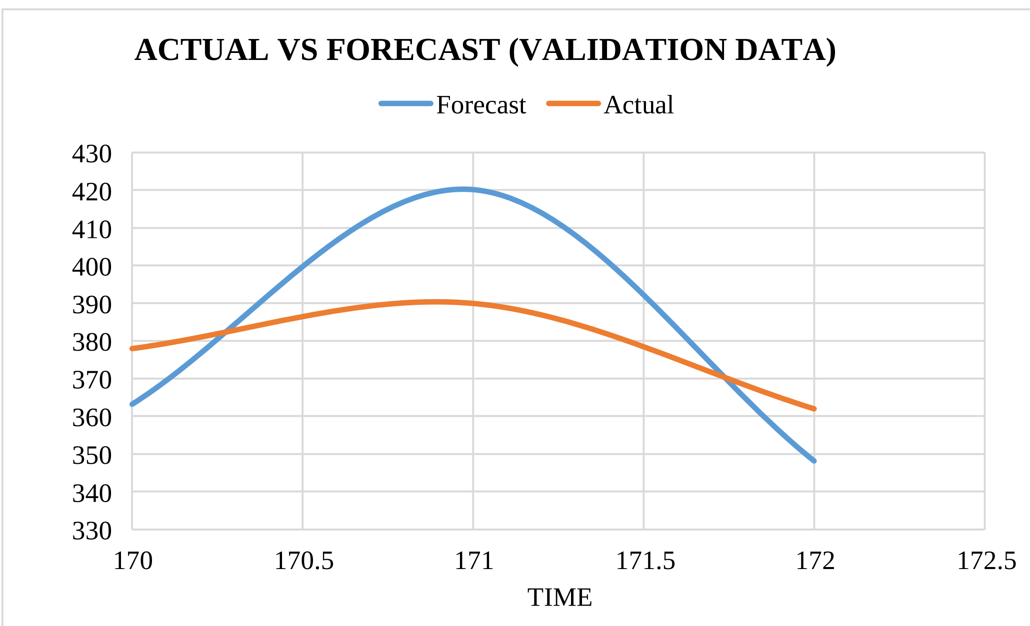


Figure 3.8 Holt-Winter's Validation Data

The forecast of Holt-Winter's is dependent on trend and seasonality. Clearly we don't have any seasonality which can be seen explicitly in our data. Even if there is any seasonality it is not regular.

Therefore it can be expected that Holt-Winter method won't perform up to the mark, but it provides better forecast than exponential and double exponential. Although the forecast lies between the confidence interval and the forecast looks promising, it can be seen that ARMA modelling provides better prediction with good recall of pattern of forecast.

Conclusion

The given dataset portrays a deterministic trend of 4^{th} order polynomial and AR(2) model was found adequate. On performing Joint Optimization on the combined adequate polynomial and ARMA model, it was seen that there was no significant difference in the parameters. Moreover, the characteristic roots indicated that there was no stochastic seasonality. Forecasted values obtained gave a fair result and were compared with other smoothing techniques namely Single, Double (Holt) and Triple (Holt-Winter's) Exponential Smoothing Methods. The values obtained initially fared better when compared to latter approaches. Proper forecasting for this particular topic requires high precision as it would affect even the exchange rates and as well as other trading of commodities. Imports and Exports are a main part of country's economy. The forecast that we obtained has some deviation from the actual value which might be due to the presence of white noise, but the results obtained after implementing our modelling procedure were found to be acceptable, given they were within the confidence intervals.