ISSN: 2321-8169
91 – 99

Time Series Least Square Forecasting Analysis and Evaluation for Natural Gas Consumption

Prabodh Kumar Pradhan
Assistant Professor
Regional College of Management
Chandrasekhar Pur, Bhubaneswar
751023 INDIA
pkpradhan.rcm@gmail.com

Sunil Dhal
Associate Professor
Sri Sri University, Cuttack 754006
INDIA
sunildhal@srisriuniversity.edu.in

Nilayam Kumar Kamila
Principal Associate
Capital One, Wilmington, DE 19801
USA
nilayam_kamila@yahoo.com

Abstract: Power and mechanical energies are now a down main stream of energy to produce any other type of product in any nation in the world. Currently natural resources are the main stream of resources to produce the power energy or mechanical energy. We aware that the natural resource energy is limited.It's the prime focus now a day to optimize the natural resource consumption. Natural Gas is an integral part of the natural resource and hence needs a planning and forecasting to optimize the natural gas usage and wastage. Without the proper planning and forecasting, the natural gas consumption will be highly usage and wasted and hence the limited resource will be drained off. In this paper, we analyze the time series least square method and compare its results with the actual data with absolute error percentage. This paper provides a clear analytic results and comparisons which serves as a base forecasting model for natural gas consumption.

Keywords: *Natural Gas, LPG Consumption, Time Series, Natural Gas Forecasting*

*****

1.0 Introduction:

Petroleum Gas e.g Liquified Petroleum Gas(LPG) is the prime constituent under the Natural Gas. The extraction of LPG from Natural resources is a very complex process. This Liquified Petroleum gas is usage in different way in different countries in the world. It's used as domestic and commercial purpose. In a study [3] in 2012, the Residential/Domestic usage is 19% and Commercial usage is 13%. Most power is consumed in Electric Power 31% and Industrial sector as 33%.

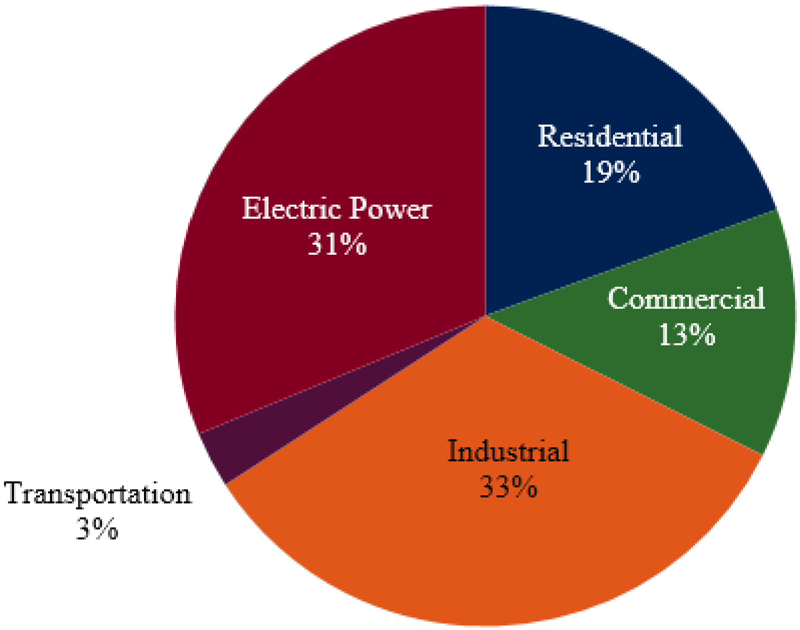

Figure 1: Natural Gas Consumption in 2012 in US

Electric Power and Industrial Power is a major player in natural gas consumption in US. In India [4], the domestic usage of LPG is the second largest in the world. Domestic LPG usage in India [4] reaches to 19 million ton in 2017.

The proper demand forecasting is important as without forecasting results, the natural gas supply will be either over supplied or under supplied. There are multiple number of time series models used for demandforecasting. In this paper, we focused on Least Square method.

The organization of this article is a s follows. In Section 2, we discuss briefly about the moving average method. The data analysis and moving average calculation is shown in section 3. In Section 4, we show the results through graphical representation. Section 5 concludes our work.

2.0 Time Series Analysis

Time series analysis is a method for forecasting the future data from a time series data. The time series calculates the future data from the existing input data. For example, if there are sales data for a number of months or number of years, then the next successive month(s) or successive year(s) could be calculated through the time series methods. There are multiple number of time series methods, and here we enlisted few of them.

a. Moving Average
b. Smoothing Analysis
c. Least Square Method

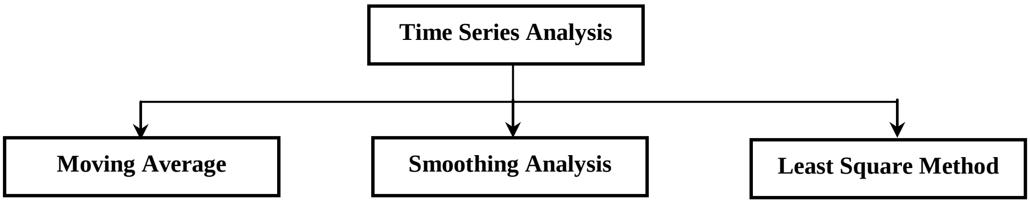

Figure2: Time Series Analysis Methods

Here in this article we'll briefly discuss each method and in next section, we'll apply the methods on our natural gas consumption data so as to find a best fit model.

A. Moving Method:

Moving average and Smoothing Analysis are basic model where the past data are used to forecast the future data. The calculation is simple and periodic average is taken to forecast the future data in case of Moving Average model. For example, if there are first four-month data available, then average of four months data is considered as the future data for fifth month.

B. Smoothing Analysis:

In Smoothing analysis, the weight is given to the past data. The basic formula used to forecast Smoothing analysis is as follows. The forecast at period t+1 could be calculated as the below formula.

$$f_{t+1} = \propto y_t + (1-\propto) f_t$$

Where $0 < \propto < 1$, actual data at time t is y_t.

C. Least Square Method:

Another popular mathematical regression method is known as least square method which evaluates a best fit line for the given data set. There are two set of data e.g. independent data set and dependent data set. For each value of independent data, there exists a dependent data and hence makes a data set pair. Here in case of natural gas consumption data set,

 ISSN: 2321-8169
91 – 99

month numbers could be taken as an independent variable and the actual consumption for the respective moth is taken as dependent variable. The line equation evaluated in the least square method is the regression line of best fit and the given data set pair is in the close proximity of the straight-line equation. The main objective of least square method is to find a straight line that minimizes the sum of squares of errors evaluated by the results of the equation. We have summarized the flow of the least square method through the flow model as shown in figure 3.

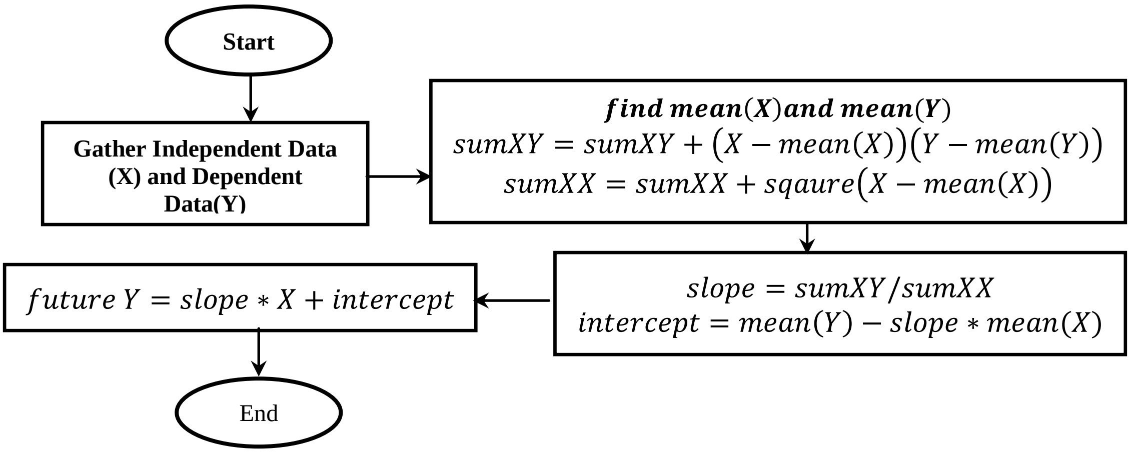

Figure3: Least Square Method Flow model

The method to calculate the least square method is as follows. Let's we have the ordered pairs data $(x_1, y_1), (x_2, y_2), (x_3, y_3) \dots \dots (x_n, y_n)$.

Step 1. Calculate the mean value of x and y

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Step 2. Calculate the slope of the line

$$m = \sum_{i=1}^{n}\{(x_i - \bar{x})(y_i - \bar{y})\} \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2$$

Step 3. Calculate the y-intercept of the line

$$b = \bar{y} - m\bar{x}$$

Step 4. Evaluate the future value of y based on x value using the following formula

$$y = mx + b$$

We have presented the below algorithm to show the algorithmic process for evaluation the forecasting data based on the past data of domestic LPG usage. In initialization process the data is gathered for the domestic natural gas usage as dependent variable with respect to the month number as independent variable.

Algorithm: leastSquareAnalysis

Initialization: *Gather the data for which forecasting to be evaluated. Here X is independent variable e.g. time data.*

$X = \{x_1, x_2, x_3, \dots \dots \dots x_t\}$

$Y = \{y_1, y_2, y_3, \dots \dots \dots y_t\}$

$\bar{x} = 0, \bar{y} = 0$

$\boldsymbol{for}\, i = 1\ \boldsymbol{to}\ t\, \boldsymbol{do}$

$\bar{x} = \bar{x} + x_i$

$\bar{y} = \bar{y} + y_i$

ISSN: 2321-8169
91 – 99

done
$evaluate\ \bar{x} = \bar{x}/t, and\ \bar{y} = \bar{y}/t$
$sumXY = 0$
$sumXX = 0$
$\boldsymbol{for} i = 1\ \boldsymbol{to}\ t\boldsymbol{do}$
$sumXY = sumXY + (x_i - \bar{x})(y_i - \bar{y})$
$sumXX = sumXX + (x_i - \bar{x})(x_i - \bar{x})$
done
$evaluate\ m = sumXY/sumXX$
$evaluate\ b = \bar{y} - m\bar{x}$
$\boldsymbol{for} i = 2\ \boldsymbol{to}\ m\boldsymbol{do}$
$y_i = mx_i + b$
done

3.0 Forecasting Data Error Evaluation

We used the above mentioned most popular methods to forecast the consumption data based on historical data. We've taken the data sets captured from petroleum agencies and applied the time series regression model methods to forecast the future data sets. The same future data sets are again compared with the actual consumption data through the graphical representation.

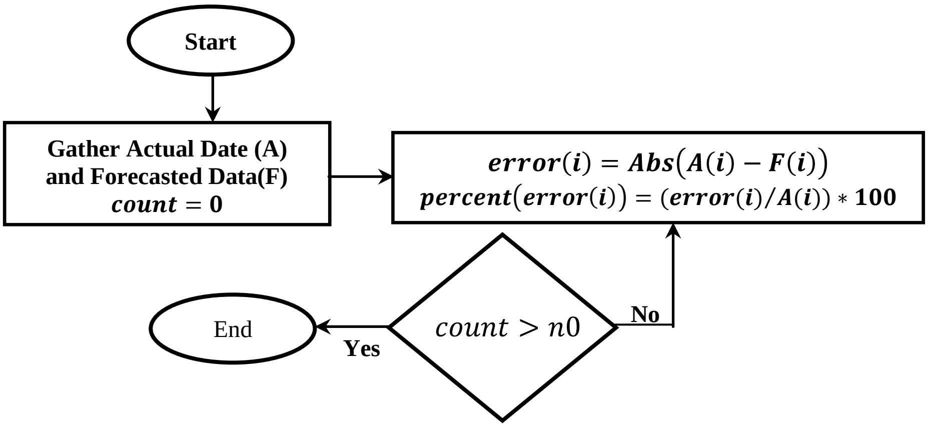

Figure4: Time Series Absolute Percent Error Flow

Algorithm: evaluate Absolute Percentage Error

Initialization: Gather the Actual Data (Data set A) and forecasted data (Data set F)
$A = \{a_1, a_2, a_3, \dots \dots \dots a_n\}$
$F = \{f_1, f_2, f_3, \dots \dots \dots f_n\}$
for $i = 1\ to\ n$ do
$e_i = [||(a_i - f_i)||]$
$p_i = [e_i/a_i] * 100$
done

In each of the model, we used the above algorithm to evaluate the errors which are presented in the graph for the visual presentation.Let's illustrate an example of the year 2012 for the domestic usage of liquified petroleum gas for the state of Odisha, India.

A. Natural Gas Consumption Least Square Regression Forecasting & Error Analysis

In case if Least Square method, we create the regression line equation with the formula $y = mx + b$. Using the value of x, the value of y is evaluated as a forecasted value. The method to caluclate the value of m and b is discussed in the earlier section of this article.

Based on the line equation, the forecasted value along with the actual value is listed in table 1. The absolute error and the absolute percentage error is calculated. The results shown in the below table for the absolute percentage error is ranging from 0.01 – 12.32%.

Months	Year 2012	y=mx+b	Absolute Error	Abs % Error
(Jan)1	16425	16680.15681	255.1568104	1.553466121
(Feb)2	16293	16680.1586	387.1586025	2.376226616
(Mar)3	16682	16680.16039	1.839605487	0.011027488
(Apr)4	14850	16680.16219	1830.162187	12.32432449
(May)5	17098	16680.16398	417.8360214	2.443771326
(Jun)6	16348	16680.16577	332.1657706	2.031843471
(Jul)7	17413	16680.16756	732.8324373	4.208536365
(Aug)8	17848	16680.16935	1167.830645	6.543201733
(Sep)9	17123	16680.17115	442.8288532	2.58616395
(Oct)10	16610	16680.17294	70.17293882	0.422474045
(Nov)11	16824	16680.17473	143.8252691	0.854881533
(Dec)12	16648	16680.17652	32.17652291	0.193275606
6.5	16680.16667		Mean Absolute % Error	2.962432729

Table 1: Actual Data and LS ForecastEvaluationfor 2012LPG Consumption Data Set

The actual vs forecasted value evaluated from least square method is shown in figure 5. The forecasted value attains a straight line as per the evaluation through least square method. The mean absolute percentage error is 2.96%.

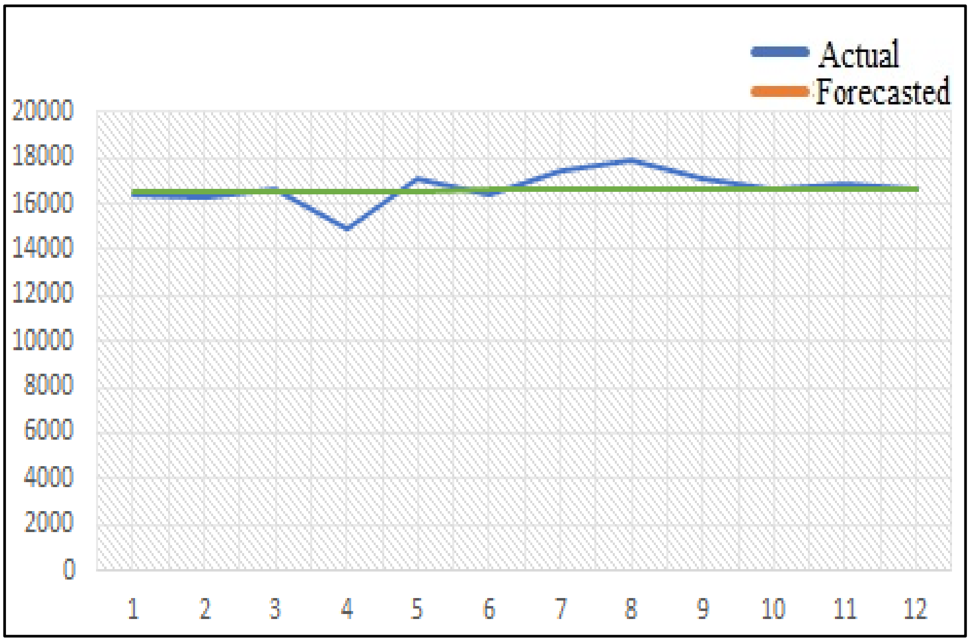

Figure5: Comparison of Actual Vs LS Forecasted for 2012 LPG Consumption Data Set

Similarly, the least square method is applied for the Domestic LPG consumption data for the year 2013. The absolute error and absolute percentage error is evaluated for all the months in the year. The absolute percentage error is ranging from 1.38 – 13.26%.

 ISSN: 2321-8169
91 – 99

Months	Year 2013	y=mx+b	Absolute Error	Abs % Error
(Jan)1	16449	17282.98942	833.9894226	5.07015273
(Feb)2	15409	17282.99135	1873.991346	12.1616675
(Mar)3	17732	17282.99327	449.0067311	2.532183234
(Apr)4	16092	17282.99519	1190.995192	7.401163262
(May)5	16428	17282.99712	854.9971153	5.204511293
(Jun)6	15259	17282.99904	2023.999038	13.26429673
(Jul)7	18287	17283.00096	1003.999038	5.490233709
(Aug)8	17715	17283.00288	431.9971153	2.438595062
(Sep)9	17047	17283.00481	236.0048079	1.384436018
(Oct)10	18412	17283.00673	1128.993269	6.131833961
(Nov)11	19536	17283.00865	2252.991346	11.53251098
(Dec)12	19030	17283.01058	1746.989423	9.180186141
6.5	17283		Mean Absolute % Error	6.815980886

Table 2: Actual Data and LS ForecastEvaluationfor 2013LPG Consumption Data Set

Figure 6 is shown with Actual consumption data against forecasted data through least square method for the year 2013. The forecasted value is the equation of straight line and hence the straight line is gathered through data. The mean percentage absolute error is 6.81%.

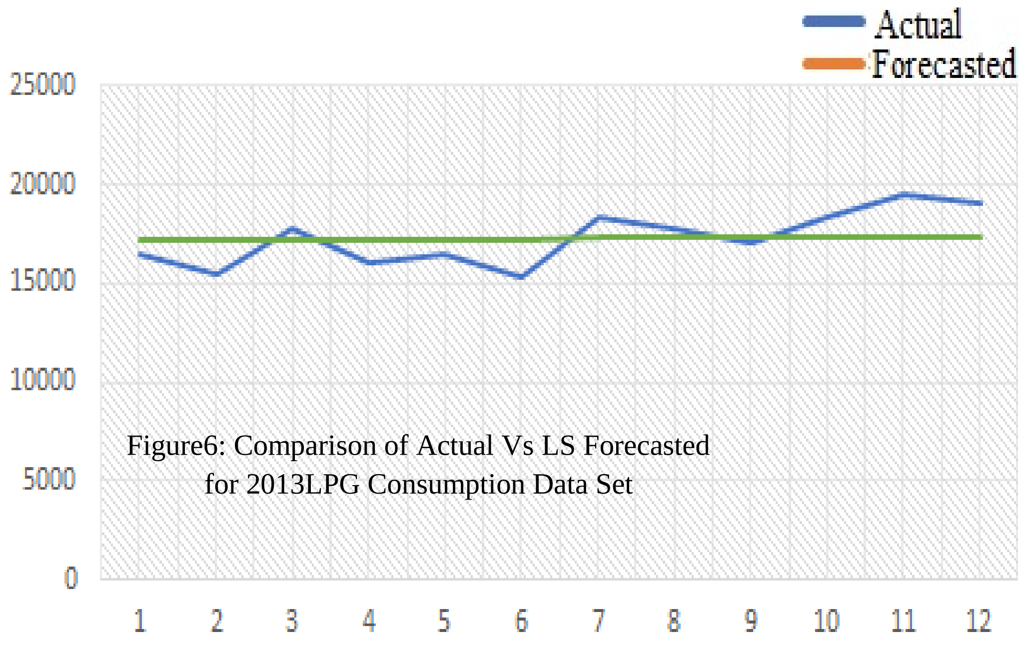

Figure6: Comparison of Actual Vs LS Forecasted for 2013LPG Consumption Data Set

The below table is the list of actual and forecasted data for the year 2014. The absolute error and absolute percentage error is evaluated through the difference between the actual data and forecasted data least square method. The absolute percentage error is ranged from 1.91 – 18.19%. The year 2014 has the highest deviation between actual and forecasted values.

Months	Year 2014	y=mx+b	Absolute Error	Abs % Error
(Jan)1	20174	20737.24327	563.2432717	2.791926597
(Feb)2	18531	20737.24449	2206.244495	11.90569583
(Mar)3	21205	20737.24572	467.7542817	2.205867869
(Apr)4	18250	20737.24694	2487.246942	13.62875037
(May)5	18831	20737.24816	1906.248165	10.12292584
(Jun)6	18039	20737.24939	2698.249388	14.95786567
(Jul)7	20233	20737.25061	504.2506117	2.49221871
(Aug)8	21142	20737.25184	404.748165	1.914427041
(Sep)9	22570	20737.25306	1832.746942	8.120278873
(Oct)10	21581	20737.25428	843.7457183	3.909669238
(Nov)11	22941	20737.25551	2203.744495	9.606139641
(Dec)12	25350	20737.25673	4612.743272	18.19622592
6.5	20737.25		Mean Absolute % Error	8.3209993

Table 3: Actual Data and LS Forecast Evaluation for 2014LPG Consumption Data Set

In below figure (figure 7) is the graph between actual consumption vs forecasted consumption. The forecasted values plot shows a straight line and the deviation of the actual data from the forecasted data is correctly captured as shown in figure 7. The mean absolute percentage error is 8.32%.

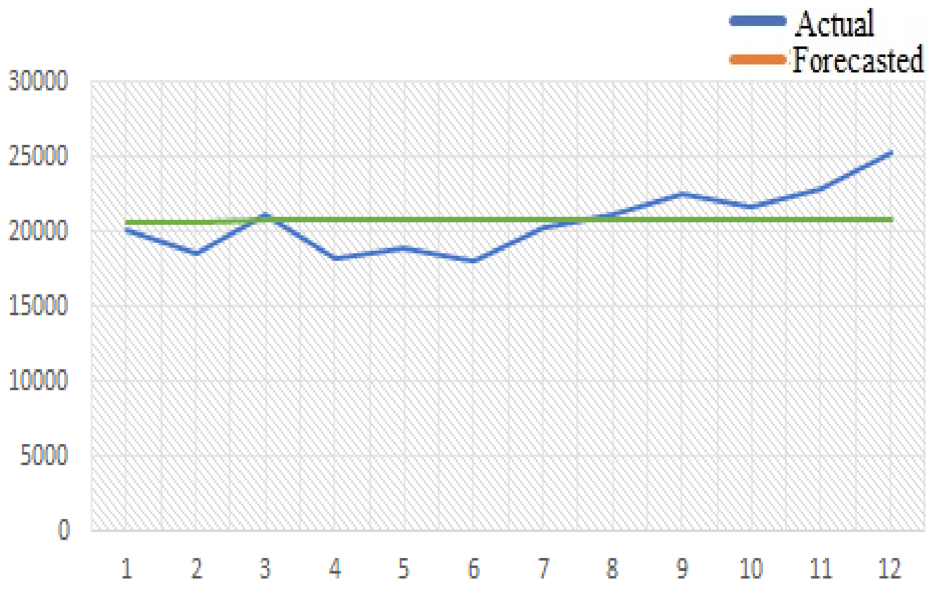

Figure7: Comparison of Actual Vs LS Forecasted for 2014LPG Consumption Data Set

4.0 Time Series Forecasting Method Comparisons

In below figure (refer figure 8, 9and 10), we have shown the comparison of absolute error percentage of different time series methods in different years. In 2012, least square method has shown a better result over other two methods. The least square methods though few times have a higher error percentage value, but most of the times it has the lower absolute percentage values.

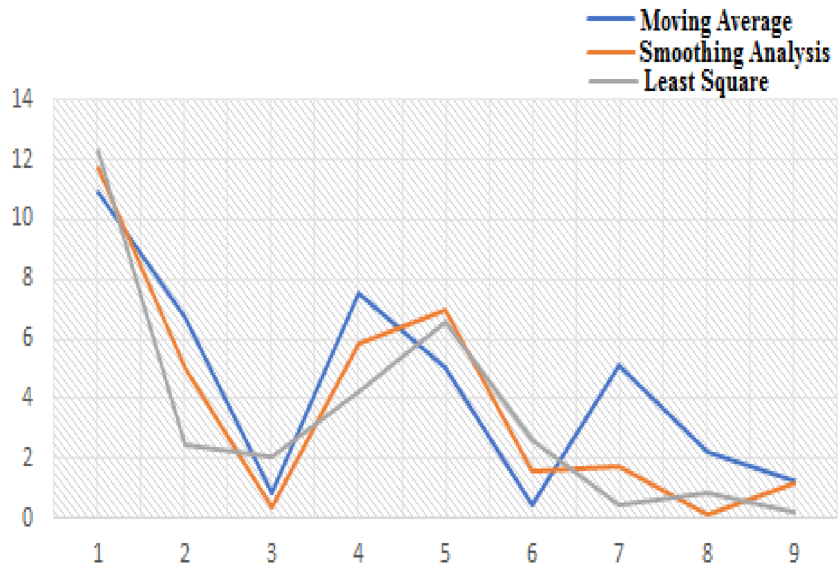

Figure8: Comparison of Absolute Percentage Error for MA Vs SA Vs LS Forecasted for 2012 LPG Consumption Data Set

In year 2013, we observed that the moving average is the best fit for the Natural Gas Consumption data. The least square method has higher percentage of absolute error in initial period of the year as well as towards the end of the year.

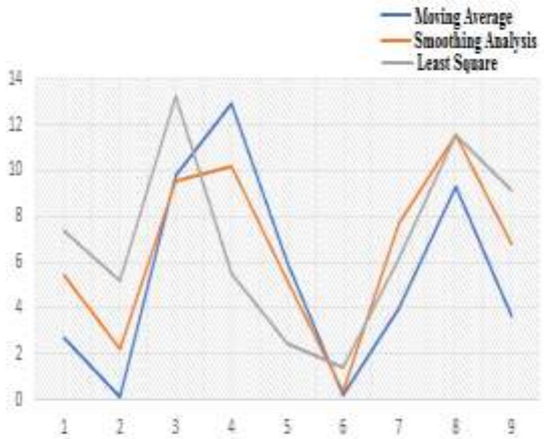

Figure9: Comparison of Absolute Percentage Error for MA Vs SA Vs LS Forecasted for 2013LPG Consumption Data Set

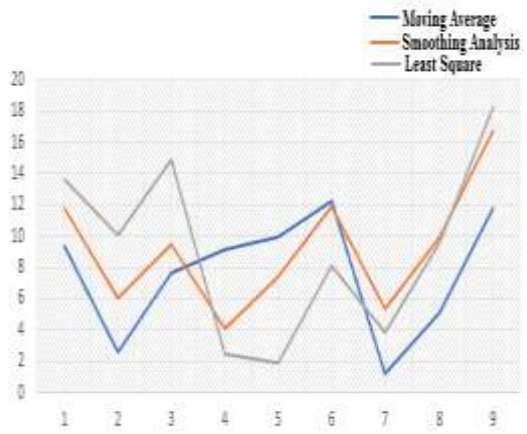

Figure 10: Comparison of Absolute Percentage Error for MA Vs SA Vs LS Forecasted for 2014LPG Consumption Data Set

Moving average is still best fit for the time series future evaluated data for the LPG Domestic consumption for year of 2014. Though moving average has shown a little higher error percentage in the middle of the year, but most of other times, it has the low percentage of absolute error.

5.0 Conclusion

Least Square method is a popular method for time series regression analysis for forcasting the future data. In this article, we discussed about different popular time series regression models and applied the Domestic LPG consumption data for 3 years based on the Least Square Method. The Moving average and the Smoothing analysis forecasting model data is evaluated along with absolute percentage error calculation separately and shown a comparison for the 3 years in figure 8, 9 and 10. We found that the

Moving average shows better results in year 2013 and 2014 whereas least square method shows better result in year 2012. All methods though show a fit to this natural gas consumption data set and the percentage of error varies between 0.01 – 18.19%. The forecasting through the time series model is showing a random behavior with moving average, smoothing analysis and least square model. The forecasted data though approximating to an error with <20%, but the consistency is not achieved with these models. The consumption data set requires still a better suit model to forecast the future consumption data with more accuracy and consistency. The best fit model with more accuracy and consistency is the future scope of the work.

References:

[1] Prabodh Pradhan, sunil dhal," A survey of different prediction models & role of artificial neural networks for Natural gas consumption", International journal of science and research, vol 4, issue 11, pg 1325-1328

[2] prabodh Pradhan, bhagirathinayak, sunil dhal," Time series data modeling and prediction of liquid petroleum gas"," International journal of development and research", volume 06, issue 10, pg: 9809-9812

[3] prabodh Pradhan, bhagirathinayak, sunil dhal," Time series data prediction of natural gas consumption using arima model", International journal of information technology and management information system", volume 7, issue 3, pg01-07

[4] prabodh Pradhan, bhagirathinayak, sunil dhal," Prediction of liquid petroleum gas for domestic consumption", " International journal of business, management and allied sciences, volume-4, issue 3, Pg: 4320-4325

[5] Ronald H. Brown, Steven R. Vitullo, George F. Corliss, Detrending daily natural gas consumption series to improve short-term forecasts in Power & Energy Society General Meeting, 2015 IEEE

[6] https://www.numbeo.com/cost-of-living/country_price_rankings?itemId=105

[7] https://www.studentenergy.org/topics/natural-gas

[8] Sun Licheng, Hu Ronghua Study on the relationship between the energy consumption and economic system of Jiangsu Province base on grey relational analysis in Grey Systems and Intelligent Services, 2009. GSIS 2009. IEEE International Conference on 08 February 2010 DOI : 10.1109/GSIS.2009.5408328

[9] https://en.wikipedia.org/wiki/List_of_countries_by_natural_gas_consumption

[10] http://ahnutritiontherapy.com/338/foods-to-fight-humidity

[11] https://www.epa.gov/climate-impacts/climate-impacts-agriculture-and-food-supply