

Comparative Analysis of Univariate Forecasting Techniques for Industrial Natural Gas Consumption

Iram Naim

Department of Polymer and Process Engineering, Indian Institute of Technology, Roorkee, 247667, India.
Email:iram.naim03cs@gmail.com

Tripti Mahara

Department of Polymer and Process Engineering, Indian Institute of Technology, Roorkee, 247667, India.
Email:triptimahara@gmail.com

Received: 10 January 2018; Accepted: 24 February 2018; Published: 08 May 2018

Abstract—This paper seeks to evaluate the appropriateness of various univariate forecasting techniques for providing accurate and statistically significant forecasts for manufacturing industries using natural gas. The term "univariate time series" refers to a time series that consists of single observation recorded sequentially over an equal time interval. A forecasting technique to predict natural gas requirement is an important aspect of an organization that uses natural gas in form of input fuel as it will help to predict future consumption of organization. We report the results from the seven most competitive techniques. Special consideration is given to the ability of these techniques to provide forecasts which outperforms the Naïve method. Naïve method, Drift method, Simple Exponential Smoothing (SES), Holt method, ETS (Error, trend, seasonal) method, ARIMA, and Neural Network (NN) have been studied and compared. Forecasting accuracy measures used for performance checking are MSE, RMSE, and MAPE. Comparison of forecasting performance shows that ARIMA model gives a better performance.

Index Terms—Forecasting, Natural gas, Simple Exponential Smoothing, Holt method, ETS method, ARIMA, Neural Network (NN).

I. INTRODUCTION

Forecasting is the process of making future predictions based on past and present data and most commonly by analysis of trends [1]. It has applications in a wide range of fields where estimates for future are useful like product forecasting, sales forecasting, flood forecasting, electricity forecasting, natural gas forecasting etc. There is a continuous rise in the requirement of energy demand worldwide, this includes electricity, coal, natural gas or

any other form of energy. According to [2] world consumption of natural gas increases by 1.7% annually for industrial use and the annual rise in consumption of natural gas is 2.2% for the electric power sector. The total share of both industrial and electric power sectors contribute to 73% of the total increase in world's natural gas consumption, and they expected to rise this to about 74% of total natural gas consumption by 2040.

Natural gas is a fossil source of energy [3]. Natural gas is widely used as a fuel in the manufacturing processes or in transportation because it is a clean fuel having pollution free emissions [4]. Another important feature is its slow cost in comparison with other fuels such as oil and electricity. Generally, the natural gas forecast is used by local distribution companies on the basis of demand & sale/consumption of natural gas. Natural gas accounts for around 10% share of the overall energy consumption in India [5]. In this research paper, we are dealing with the problem of natural gas consumption forecast at the organization level, linked with their production.

This paper attempts to evaluate the precise forecast of existing univariate forecasting techniques by providing accurate and statistically significant forecasts for natural gas consumption at the organization level. Various time series based forecasting techniques like Naïve method, Drift method, Simple Exponential Smoothing (SES), Holt method, Holt method with drift, ETS (Error, trend, seasonal) method, ARIMA and Feedforward Neural Network are analyzed.

This paper reports the results from the seven most competitive forecasting techniques. Comparison between these methods is performed by error analysis and time series cross-validation by taking the forecasting horizon of one and three months ahead. On comparison of univariate models used in this article, we find that ARIMA model provides most promising results for the particular dataset under consideration.

II. RELATED WORK

The Naïve method represents predicted values as the last observed value in the time series. Drift method represents the amount of average change over time on past values. This change can be decreasing or increasing. All these methods i.e. Naïve and Drift are available as established standard methods of forecasting. A new method of forecasting can be compared with these established methods and its performance can be evaluated [6].

Exponential smoothing method has been around since the 1950s and is the most popular forecasting method used in business and industry [7]. In reference [8] the analysis of seasonal time series data using Holt-Winters exponential smoothing method is being produced with a discussion of multiplicative seasonal model and additive seasonal models. In reference [9], formulae are given for evaluating mean and variances for lead time demand of exponential smoothing method over a wide range. In reference [10] researcher provided forecasting of inventory control using exponential smoothing.

The Holt-Winters forecasting procedure is a variant of exponential smoothing and suitable for producing short-term forecasts for sales or demand time-series data [11]. In reference [12] long lead-time forecasting by Holt-Winters method for UK air passengers is available. Reference [13] describe Holt-Winters method with different models of forecasting and prediction intervals for the multiplicative model.

ARIMA method is the most commonly used method for short-term forecasting. Reference [14] presented a time series method for daily basis forecast and used the concept of degree day for the prediction of gas demand of a city. Reference [15] presented a time series method for the annual forecast to make use of multivariable regression analysis and discover various parameters affecting gas demands. Reference [16] provide monthly forecast by partitioning years into heating & non-heating seasons and use degree day method for the forecast in time series method. Reference [17] studied ARIMA models for forecasting natural gas consumption in Turkey on monthly basis. Reference [18] shows the daily basis forecasting for local distribution companies using ARIMA models and Artificial Neural Network. Reference [19] generate forecasting models for dynamic relationships among several potentially relevant time series variables. Reference [20] studied about natural gas consumption of Turkey by using exponential smoothing and Box-Jenkins methods. Reference [21] define ARIMA forecasting model represented in terms of past values of itself including current and lagged values of a 'white noise' error term. Reference [42] analyzed the data of given parameters and noticed their predictions for a particular period by using the strategy of Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing (ETS). The performances of four different interval ARIMA-base time series methods are evaluated in financial markets forecasting [43].

In reference [22] researcher explain transit containers in Kaohsiung port by adopting artificial neural network(ANN). In reference [23] a model of neural networks (NN) for forecasting throughput of Hong Kong port cargo has been discussed. Reference [24] used the NN approach to forecast time series on a quarterly basis. Reference [25] and [26] applied error correction model to forecast Hong Kong's throughput. Reference [27] describe ANN with seven different algorithms, namely-Limited Memory Quasi-Newton Algorithm, Conjugate Gradient Descent Algorithm, Quasi-Newton Algorithm, Levenberg-Marquardt Algorithm, Quick propagation Algorithm, Incremental Back-Propagation Algorithm and Batch Back-Propagation Algorithm to achieve daily and monthly forecast of natural gas consumption in Istanbul. Reference [28] describe Artificial Neural Network and artificial bee colony (ABC) algorithm to study consumption demand forecasting on monthly basis. In Reference [40] Artificial Neural Network (ANN) technique has been used to develop one-month and two month ahead forecasting models for rainfall prediction using monthly rainfall data of Northern India. A Neural Network structure has been used to forecast the stock values of different organizations of Banking and Financial sectors in NIFTY 50 [41].

III. METHODOLOGY

In this section, we presented the details about all the univariate forecasting models considered for natural gas consumption forecasting. Considering y_1, y_2, \dots, y_T represents previous data. $\hat{y}_{T+h|T}$ represent forecasted values based upon y_1, y_2, \dots, y_T , the various techniques of forecasting are as discussed below:

A. Univariate Forecasting Techniques

1) Naïve Method

In this method, forecast value is equal to last measured value. If y_T is the last measured value of time series data then future values represent the same value [29, 30]. As naive method provides information about the most recent values, the forecast value generated from the Naive model is not linked with previous values present in Time series [29, 30].

Forecasts Equation:

$$\hat{y}_{T+h|T} = y_T \quad (1)$$

2) Drift Method

In Drift method, forecast value is the sum of last measured value and the average change in time series. This method allows a change in forecast values to increase and decrease over time depending upon the average change within the time series.

Forecasts Equation:

$$\begin{aligned} \hat{y}_{T+h|T} &= y_T + \frac{h}{T-1} \sum_{t=2}^T (y_t - y_{t-1}) \\ &= y_T + \frac{h}{T-1} (y_T - y_1) \end{aligned} \quad (2)$$

3) Simple Exponential Smoothing Method

Forecasts values generated are weighted averages of past observations, with the weights decaying exponentially as the observations get older. The value of α falls between zero and unity, the pattern of weights attached to the observations shows exponentially decreasing trend. The value of α towards zero indicates more weight of observations of long past. However larger values of α towards one signifies more weight to recent observations [29, 30].

Forecast equation :

$$\begin{aligned} \hat{y}_{t+h|t} &= \sum_{j=1}^t \alpha(1-\alpha)^{t-j} y_j(1-\alpha)^t \iota_0, \\ \text{Where } (0 \leq \alpha \leq 1) \end{aligned} \quad (3)$$

ι_0 is an unobserved "state".

4) Holt's Method

Holt makes an extension to Simple exponential smoothing method to allow forecasting of data with trends. Two smoothing parameters are being introduced namely: α and β [29, 30].

Forecast Equation:

$$\begin{aligned} \hat{y}_{t+h|t} &= l_t + hb_t \\ \text{Level } l_t &= \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \\ \text{Trend } b_t &= \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \end{aligned} \quad (4)$$

l_t indicates the estimate of level of the series at time t .

b_t indicates the estimate of the slope of the series at time t .

5) ETS Method

Each model has equations for states level, trend and season. It is also known as state space models [29, 30]. Two models for each method: namely additive and multiplicative errors. A combination of 30 models is found. Additive and multiplicative methods provide the similar point forecasts with different forecasting intervals.

ETS(Error, Trend, Seasonal):

- Error={A,M}
- Trend={N,A,A_d,M,M_d}
- Seasonal={N,A,M}

ETS (M,A,N)

$$\begin{aligned} y_t &= (l_{t-1} + b_{t-1})(1 + \epsilon_t) \\ l_t &= (l_{t-1} + b_{t-1})(1 + \alpha\epsilon_t) \\ b_t &= b_{t-1} + \beta(l_{t-1} + b_{t-1})\epsilon_t \end{aligned} \quad (5)$$

In this model error is multiplicative, the trend is additive and no season is available.

6) ARIMA Method

For any given time series of data X_t , the ARMA model is a technique for predicting future values in the given

time series. The model has two components, first is autoregressive (AR) and another is moving average (MA) part. The AR components include regressing the variable on its own lagged or past values. The MA part includes modeling the error term as a linear combination of error terms. An autoregressive integrated moving-average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. In ARIMA differencing is performed to make time series stationary [31, 32].

for zero differencing $x_t = X_t$

for differencing only once $x_t = X_t - X_{t-1}$

And the general equation is given as

$$X_t = c + \sum_{i=1}^p \phi_i x_{t-i} - \sum_{i=1}^p \theta_i \epsilon_{t-i} \quad (6)$$

where the moving average term is negative as the error is minimized each time differencing is done.

7) Neural Network

It is a class of statistical methods for information processing consisting of a large number of simple processing units (neurons), which exchange information of their activation via directed connections [33, 34, 35].

Activation Function

$$y_t = \begin{cases} 1 & \text{if } \sum_j w_{ji} x_j - \theta_i \\ 0 & \text{if } \sum_j w_{ji} x_j - \theta_i \end{cases} \quad (7)$$

Where y_i is output function, x_i is input and w_{ji} represents the intermediate nodes.

B. Accuracy measures

Any forecasting technique needs to be evaluated for prediction accuracy. Accuracy measures check the errors between obtained values and actual values during prediction. Let y_t denote the t^{th} observation and $\hat{y}_{t|t-1}$ indicate its forecasted value based on previous values, where $t=1, \dots, T$. The following measures are used to measure the accuracy of forecasting method. Mean absolute error (MAE) is the average absolute error value. The effectiveness of forecast is considered good for the value of MAE close to zero. Mean squared error is defined as the sum or average of the squared error values [48]. MAE ("Eq. (8)"), MSE ("Eq. (9)"), RMSE ("Eq. (10)") are all scale dependent. MAPE ("Eq. (11)") is scale independent but is only sensible if $y_t \gg 0$ for all t , and y has a natural zero [36]. The formulas for errors are as follows:

$$MAE = T^{-1} \sum_{t=1}^T |y_t - \hat{y}_{t|t-1}| \quad (8)$$

$$MSE = T^{-1} \sum_{t=1}^T (y_t - \hat{y}_{t|t-1})^2 \quad (9)$$

$$RMSE = \sqrt{T^{-1} \sum_{t=1}^T (y_t - \hat{y}_{t|t-1})^2} \quad (10)$$

$$MAPE = 100 T^{-1} \sum_{t=1}^T \frac{|y_t - \hat{y}_{t|t-1}|}{|y_t|} \quad (11)$$

Along with the accuracy measures, Akaike information criterion(AIC) represents the criteria to select the appropriate model."AIC" is an estimate of a constant plus the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model. Lower AIC means a model is considered to be closer to the truth [37-38].

C. Time series cross-validation

In time series cross-validation [39], many different training sets, each containing one more observation than the previous one are used. Assuming that there is a requirement of k observations for producing a strong reliable forecast, the process can be described as below steps.

1. Pick the test set observation for time $t = k + i$ and use the observations at times $1, 2, \dots, k + i - 1$ to estimate the forecasting model. Calculate the error on the forecast for time $k + i$.
2. Repeat the above step for $i = 1, 2, \dots, T - k$ where T is the overall number of observations.
3. Based on the errors evaluated, calculate forecast accuracy measures.

D. Proposed Framework

To perform a comparative analysis, first data collection is carried out. The forecasting techniques (Naïve, Drift, SES, Holt method, ETS, ARIMA, and Neural network) have been applied to collected data and forecasted values have been obtained for each technique. The error components RMSE, MAE, MAPE, and MASE have been worked out for each forecasting technique. Time series

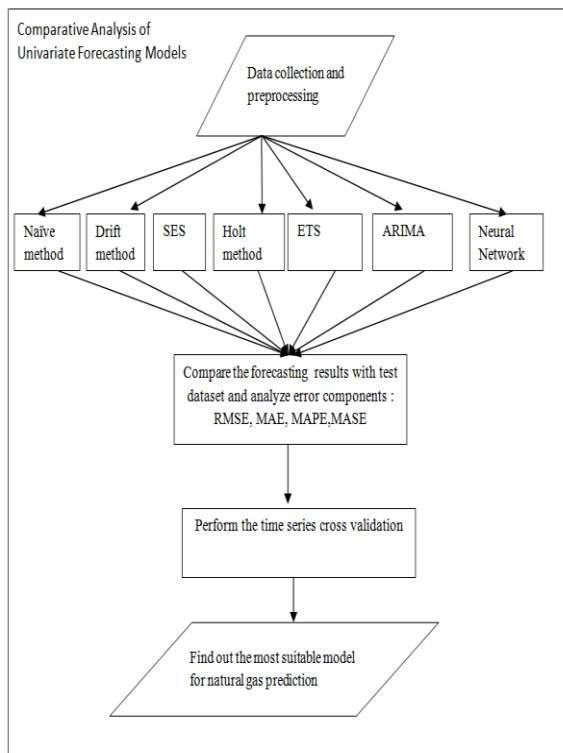


Fig.1. Framework for comparison of Univariate Forecasting Models

cross-validation is also performed for various forecasting techniques. The complete framework for forecasting of the univariate model is shown in figure-1.

IV. EXPERIMENT AND RESULT

This section is divided into three parts. The first section describes the data related to industrial natural gas consumption, second will show the results from different forecasting techniques and third contains the cross-validation and error analysis to check the most suitable forecasting techniques for this natural gas data.

A. Industrial natural gas monthly consumption data

A monthly dataset is prepared by collecting natural gas consumption data from CFFP, a manufacturing unit of Bharat Heavy Electricals Limited, India for their consumption of natural gas production of Castings products from year April 2014 to January 2017. The consumption data is measured in scm (standard cubic meter) for this period. This is the total sum of consumed natural gas in all the operations related to the casting process. In order to examine the pattern and behavior of data with respect to time, a time series plot of considered data has been drawn in figure 2.

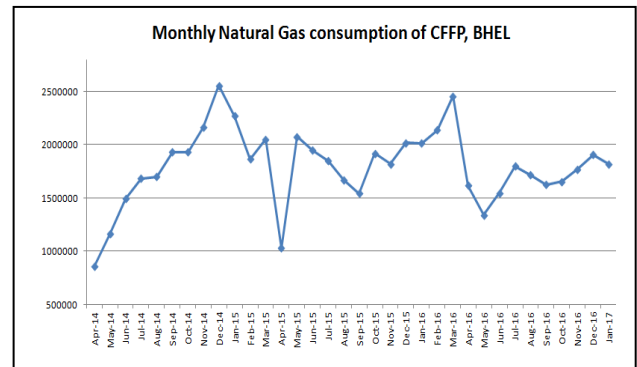


Fig.2. Monthly natural gas consumption in CFFP,BHEL

To perform the forecast of natural gas consumption on monthly basis for CFFP, BHEL, various univariate forecasting models have been implemented and their comparative analysis is done. This study identifies the most suitable model for short-term forecasting on the basis of error analysis for 1- month forecast, and 3-month forecast. For any forecasting method, forecasting period should be predefined that requires the division of data in two parts i.e. in sample period and out-of-sample period. The dataset used for identifying the relationship between the data is termed as training set or in-sample period. On the other hand, test set or out-sample period is the data set that is used to examine the strength of the predicted relationship.

1. Training set: Month1 to Month 30 used for model development.
2. Test set: Month 31 to Month 34 used for training Purpose.

To further analyze the randomness of the considered data, autocorrelation (ACF) plot as shown in figure 3 was created. This is used to measure the relationship between lagged values of time series. At lag 1 the value is significant, representing the strong correlation between successive values.

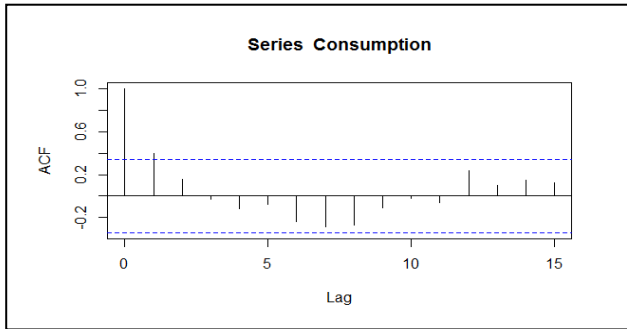


Fig.3. Autocorrelation Plot

B. Results from forecasting Techniques

This section contains the details of all fitted model for 1- month forecast of natural gas consumption. The forecast values are generated from the Random walk, SES, Holt, ETS, ARIMA and Neural Network.

1) Naïve method (Random Walk) Method

Naïve or random walk method is the most basic method for forecasting. The Forecasted values generated from this method is equal to the last observed value i.e. in our training set, we consider 1 to 30 months data than for 1- point forecast 31th observation will represent forecasted value. Figure 4 and Figure 5 are the graphical representation of forecasted values from a naïve random walk with drift method respectively. Table 1 represents both of these forecasted values. Random Walk with drift method provides forecast values equal to the sum of final observed value and average change in time series.

Table 1. Forecasted Values from Random Walk Method

S. No.	Test set	Naïve	Drift
1	1653243	1625847	1652319

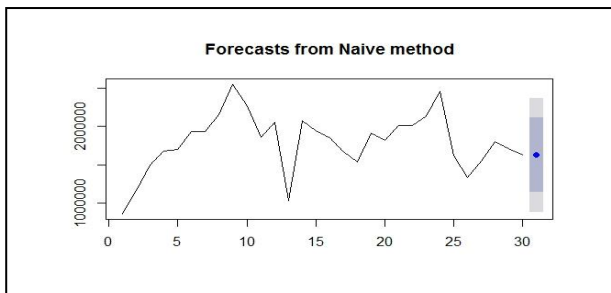


Fig.4. 1-Month ahead Forecast from Random Walk/Naïve method

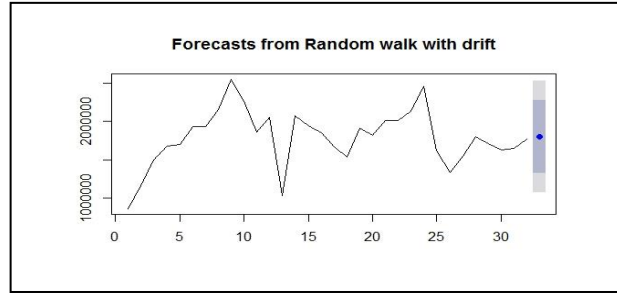


Fig.5. 1-Month ahead forecast from Random Walk with Drift method

2) Simple Exponential Smoothing Method

Forecast with Simple Exponential Smoothing (SES) method requires initialization of smoothing parameter. In SES only one parameter, alpha, is used for Smoothing. The parameters alpha has values between 0 and 1. If the value is close to 0 it represents that less weight is assigned to a most recent measured value. The estimated value of alpha for 1- month natural gas consumption forecast is 0.61. This represents the current value relation with last measured value. Table 2 depicts forecasted value from SES. Figure 6 shows the forecasted value from SES graphically and figure 7 is also graphical representation for level decomposition by SES method.

Table 2. Forecasted Value from SES Method

S. No.	Test set	SES
1	1653243	1666915

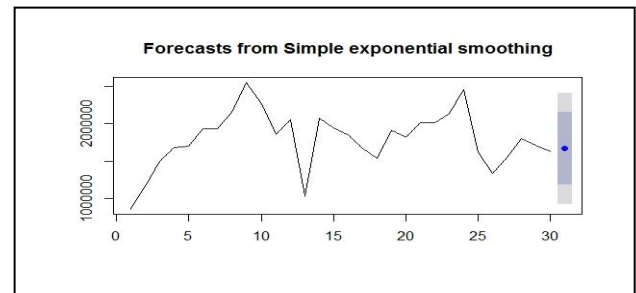


Fig.6. 1-Month ahead Forecast from SES method

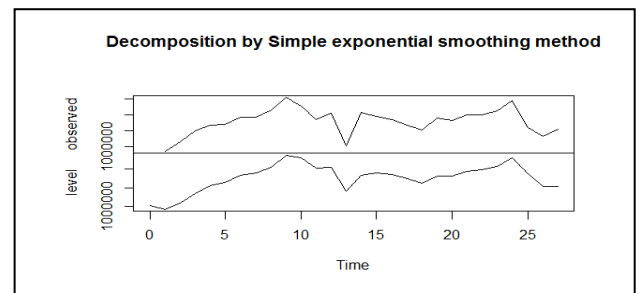


Fig.7. Decomposition from SES method

3) *Holt's Method*

The smoothing parameters for Holt's method are denoted by α and β with values ranging from 0 to 1. For 1-month ahead, forecast of natural gas consumption has $\alpha = 0.6014$ and $\beta = 0.0733$. Damped Holt's method has smoothing parameter with one damping parameter ϕ , with constraint $0 < \phi < 1$. For 1-month forecast of natural gas consumption $\alpha = 0.4473$, $\beta = 0.0001$ and $\phi = 0.8588$. Figure-8 and figure-9 graphically represents the forecasted values from Holt's method and Damped Holt's method with drift respectively. Table 3 represents both the forecasted values.

Table 3. Forecasted Values from Holt's Method

S. No.	Test set	Holt	Damped Holt
1	1653243	1634483	1674545

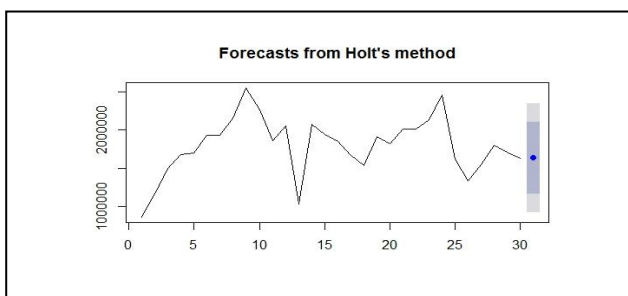


Fig.8. 1-Month ahead Forecast from Holt's method

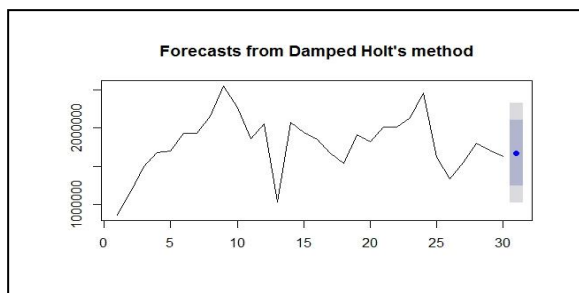


Fig.9. 1-Month ahead Forecast from Holt's with Damped method

4) *ETS (Error, Trend & Seasonality) Method*

ETS method for this forecast contains multiplicative level, additive trend, and absence of seasonal component given by N. The Smoothing parameters of fitted ETS (M,A,N) model are $\alpha = 0.4486$ and $\beta = 0.0648$. Table 4 represents a forecasted value from ETS method. Figure 10 graphically shows the forecasted value from ETS and Figure 11 is graphical representation for level decomposition by ETS method.

Table 4. Forecasted Value from ETS Method

S. No.	Test set	ETS(M,A,N)
1	1653243	1620304

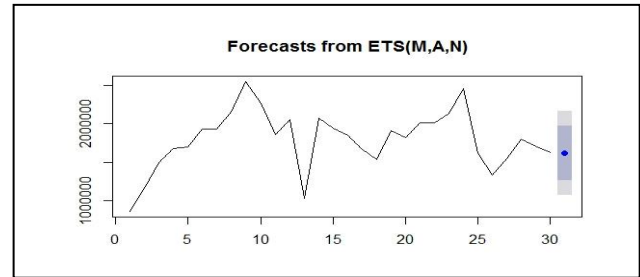


Fig.10. 1-Month ahead Forecast from ETS method

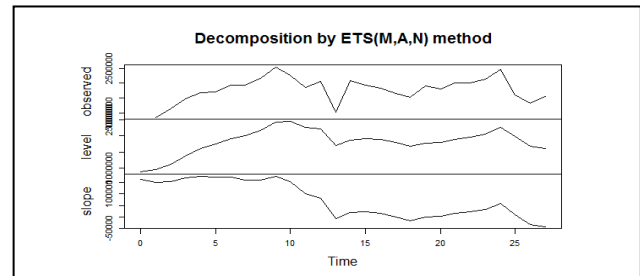


Fig.11. Decomposition from ETS method

5) *ARIMA Method*

The forecast with ARIMA is done using Hyndman-Khandakar algorithm that minimizes AIC values to arrive at an appropriate model with p, d, and q parameter. It performs changes in p and q values by a maximum of ± 1 . ARIMA model selected for 1-month forecast is of the order ARIMA(1,0,0). In which p=1, represent Auto-Regressive (AR) component of the first order. The value of d define the order of differencing, here d=0 represent series is stationary and auto-correlated. The q=0 shows the absence of Moving Average term. Table 5 depicts forecasted value from ARIMA method. Figure 12 is graphical representation of the forecasted value from ARIMA method.

Table 5. Forecasted Value from ARIMA Method

S. No.	Test set	ARIMA (1,0,0)
1	1653243	1695431

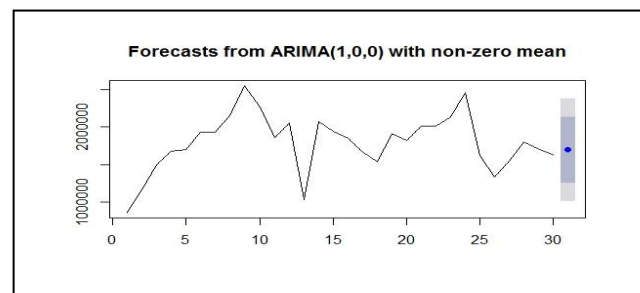


Fig.12. 1-Month ahead Forecast from ARIMA method

6) Neural Network

This forecast uses feed-forward networks denoted by NNAR(p,k) to indicate that there are p lagged inputs and k nodes in the hidden layer. Fitted neural network NNAR(1,1) model denotes one input layer, one hidden layer, and one output layer. The value p=1 tells that only previous value act as input to the model. The value of k represents that only single neuron is present at hidden layer. The generated model is the average of 20 networks, each of which is a 1-1-1 network with linear output units. In table 6 forecasted value from Neural Network method is presented. Figure 13 shows a graphical representation of the forecasted value from Neural Network method.

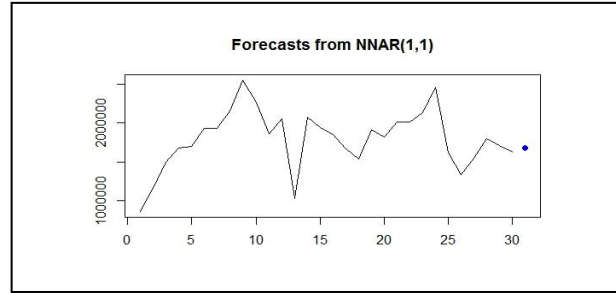


Fig.13. 1-Month ahead Forecast from Neural-Network method

Similarly, forecast values generated for the 3-month forecast for the training set 1 is represented in figure 14.

Table 6. Forecasted Value from Neural Network Method

S. No.	Test set	Neural Network
1	1653243	1678049

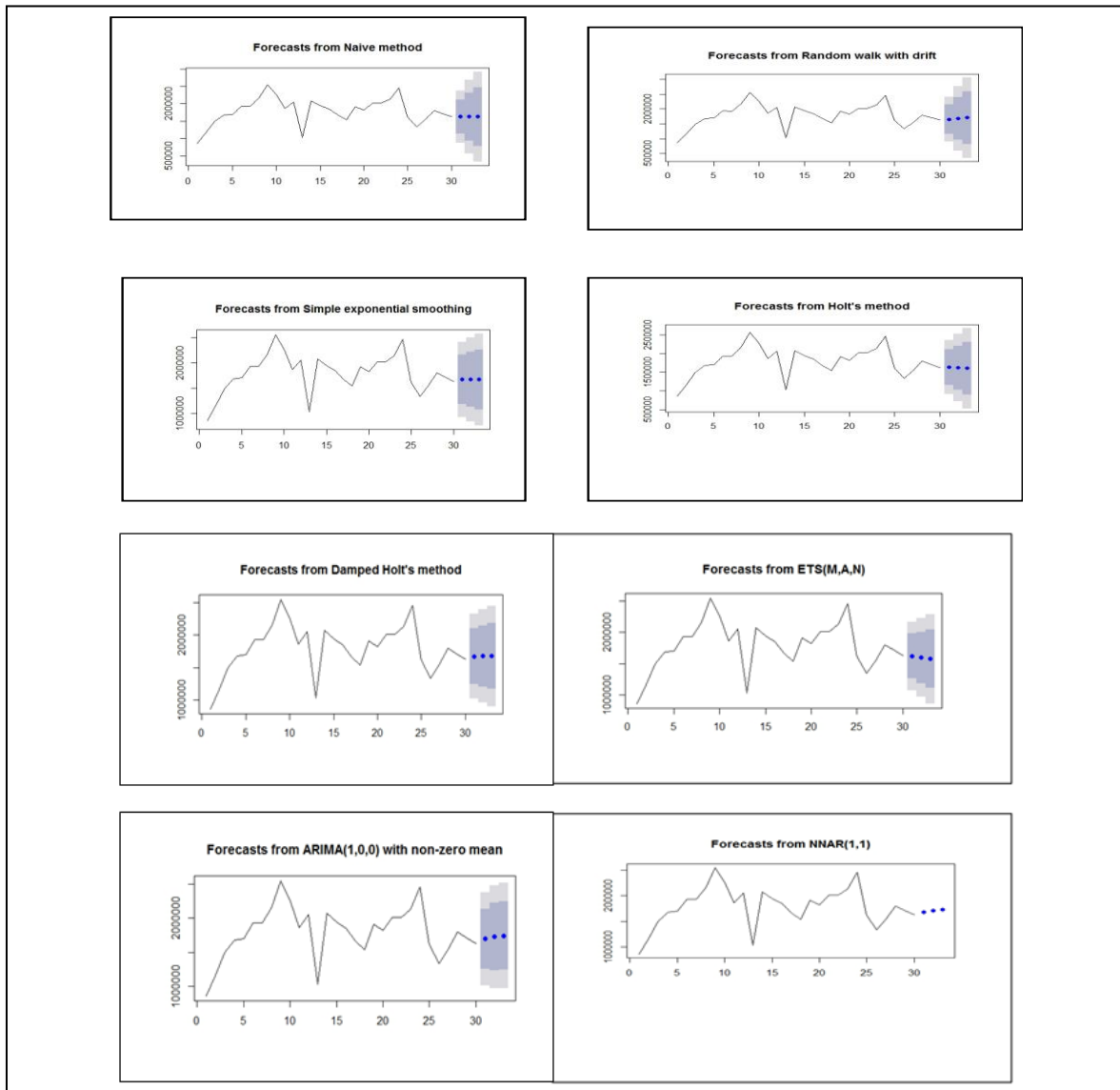


Fig.14. 3-Month Forecast values from Different Models for training set 1

C. Cross-validation and Error Analysis

In time series cross-validation, the procedure for validation is to use different training sets, each one containing one more observation than the previous one.

In our 1 -month forecast we start with 30 months data as a training set and predict for 31stmonth value. In the next step, the value of training set has been incremented by one i.e. now 31 values are there in training set and prediction is made for 32nd-month value. Similarly, we move to the 3rd and 4th step. Figure 15 defines the procedure for time series cross-validation where blue dots represent training set, red dot represent test data and black dot is not considered in forecasting. In 3 -month forecast we start with 30 months data as a training set and predict for three points ahead i.e. 31th, 32nd and 33rd-month value. In the next step, the value of training set has been incremented by one i.e. now 31 values are there in training set and prediction is made for next three months value explained in figure 16.

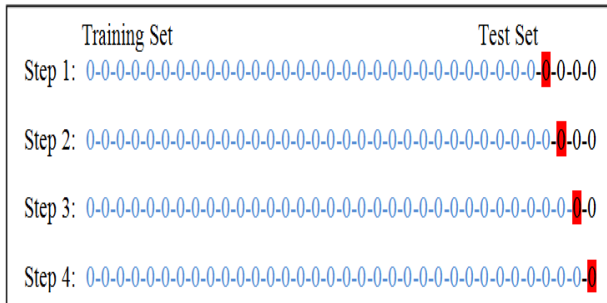


Fig.15. Time-Series Cross-validation for 1-month forecast

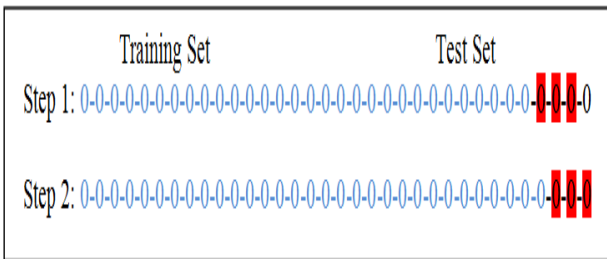


Fig.16. Time-Series Cross-validation for 3-month forecast

All these steps have been performed using all the method defined in section III. The forecasted values are generated from these methods with the different training set for 1-month and 3-month forecast. Table 7 summaries obtained forecast values at each step.

Tr1, Tr2, Tr3, and Tr4 are the four training set for the different step of cross-validation. We check the accuracy of each model using AIC parameter of chosen techniques that are shown in table 8. The AIC value for 1-month forecast for all four training sets Tr1, Tr2, Tr3, and Tr4 are 854.83, 882.11, 909.39, 936.78 respectively for ARIMA technique. It is clear from table 8 that these AIC values for ARIMA are lowest as compared to all other

forecasting techniques. Similar findings are observed and recorded for 3-month forecast as shown in table 8. Therefore this cross-validation established that ARIMA (1,0,0) is the most suitable forecasting technique for such time series data.

Table 7. Forecasted Values from Different Models

1-Month ahead Forecast										
Training set	Test set	Naïve	Drift	SES	Holt	Holt Damped	ETS	ARIMA (1,0,0)	Neural Network	
Tr1	1653243	1625847	1652319	1666915	1634483	1674545	1620304	1695431	1678049	
	1770841	1653243	1679746	1660174	1632118	1666128	1613835	1706892	1690050	
Tr3	1907709	1770841	1800283	1725934	1714168	1713981	1720585	1765997	1786147	
	1820379	1907709	1940508	1838449	1847284	1802269	1804614	1836123	1904373	
3-Month ahead Forecast										
Training set	Test set	Naïve	Drift	SES	Holt	Holt Damped	ETS	ARIMA (1,0,0)	Neural Network	
Tr1	1653243	1625847	1652319	1666915	1634483	1674545	1620304	1695431	1678037	
	1770841	1625847	1678792	1666915	1619156	1675757	1596556	1729028	1708932	
	1907709	1625847	1705264	1666915	1603824	1676797	1572808	1745249	1730739	
Tr2	1770841	1653243	1679746	1660174	1632118	1666128	1613835	1706892	1690051	
	1907709	1653243	1706249	1660174	1618324	1667162	1592344	1732953	1713907	
	1820379	1653243	1732752	1660174	1604531	1668049	1570853	1745614	1731434	

Table 8. AIC Parameters for Different Model

Method	AIC for 1-month forecast				AIC for 3-month forecast	
	Tr1	Tr2	Tr3	Tr4	Tr1	Tr2
SES	879.044	908.143	932.6278	961.8721	879.0435	908.1429
Holt	880.661	909.686	938.8584	968.1831	880.661	909.6857
Holt Damped	877.451	906.304	935.2558	964.4701	877.4514	906.304
ETS	872.505	901.006	928.7165	957.7302	872.5053	901.0062
ARIMA	854.83	882.11	909.39	936.78	854.83	882.11

Furthermore, the analysis of different types of errors occur during forecasting at each step of validation is performed. To compare the accuracy of various models the average value is evaluated. Table 9 represents the error components occur during forecasting from different models. Now we plot the error terms from each model into a single plot. Figure 17 and 18 depicts the error terms for each model and it is clear that all the concerned error components have minimum values for ARIMA method in comparison to other methods. Thus, it is concluded that ARIMA (1,0,0) is good techniques to forecast for the available training and test data set.

Table 9. Error Components for 1-Month and 3- Month Forecast

Method	Errors	Errors for 1-month Forecast					Errors for 3-month forecast		
		Tr1	Tr2	Tr3	Tr4	Average	Tr1	Tr2	Average
Naïve	RMSE	27396	117598	136868	87330	92298	183684.5	188429.2	186056.85
	MAE	27396	117598	136868	87330	92298	151417.3	179733.3	165575.3
	MAPE	1.657107	6.640799	7.174469	4.797353	5.067432	8.206621	9.720337	8.963479
	MASE	0.098576	0.436246	0.517135	0.335018	0.346744	0.544826	0.666745	0.605785
Drift	RMSE	923.6897	91094.9	107426.4	120128.7	79893.4224	128397.8	137310.8	132854.3
	MAE	923.6897	91094.9	107426.4	120128.7	79893.4224	98472.71	126727.1	112599.91
	MAPE	0.05587138	5.14416	5.631171	6.599105	4.35757684	5.288626	6.839371	6.0639985
	MASE	0.003324	0.337929	0.405894	0.460842	0.301997	0.354322	0.470112	0.412217
SES	RMSE	13672.19	110666.9	181774.9	18070.08	81046.0175	151623.6	181829.9	166726.75
	MAE	13672.19	110666.9	181774.9	18070.08	81046.0175	119463.9	172802.3	146133.1
	MAPE	0.8269921	6.2494	9.528439	0.992655	4.39937152	6.439288	9.341848	7.890568
	MASE	0.049195	0.410534	0.686809	0.069321	0.303965	0.429852	0.641034	0.535443
Holt	RMSE	18759.79	138723.2	193541.2	26904.84	94482.2575	196387.5	223292.2	209839.85
	MAE	18759.79	138723.2	193541.2	26904.84	94482.2575	158108.8	214652.2	186380.5
	MAPE	1.134727	7.833749	10.14522	1.47798	5.147919	8.54319	11.622011	10.082601
	MASE	0.675009	0.514613	0.731660	0.103213	0.506124	0.568903	0.796282	0.682592
Holt Damped	RMSE	21302.47	104712.7	193728.3	18109.93	84463.35	144701	175149.4	159925.2
	MAE	21302.47	104712.7	193728.3	18109.93	84463.35	115766.2	165863.1	140814.65
	MAPE	1.288526	5.913161	10.15502	0.994844	4.58788775	6.254037	8.963462	7.6087495
	MASE	0.076650	0.388446	0.731973	0.069474	0.316636	0.416547	0.615292	0.515919
ETS	RMSE	32938.86	157006.2	187123.9	15764.83	98208.4475	218799	249245	234022
	MAE	32938.86	157006.2	187123.9	15764.83	98208.4475	180708	240632	210670
	MAPE	1.992379	8.866193	9.808829	0.8660192	5.38335505	9.796487	13.03488	11.415684
	MASE	0.118520	0.582436	0.707019	0.060478	0.367113	0.650220	0.892659	0.771439
Arima (1,0,0)	RMSE	42187.98	63949.36	141712.2	15743.83	65898.3425	99868.76	115785.6	107827.18
	MAE	42187.98	63949.36	141712.2	15743.83	65898.3425	82153.59	104490.1	93321.845
	MAPE	2.551832	3.611242	7.428396	0.8648654	3.61408385	4.47633	5.626288	5.051309
	MASE	0.151800	0.237229	0.535438	0.060397	0.246216	0.295603	0.387620	0.341611
Neural Network	RMSE	24806.5	80791.01	121562.5	83993.7	77788.4275	109187.7	131652.9	120420.3
	MAE	24806.5	80791.01	121562.5	83993.7	77788.4275	87890.88	121179	104534.94
	MAPE	1.500475	4.562296	6.372172	4.614077	4.262255	4.75743	6.535732	5.646581
	MASE	0.098258	0.299971	0.459306	0.322220	0.294938	0.316247	0.449530	0.382888

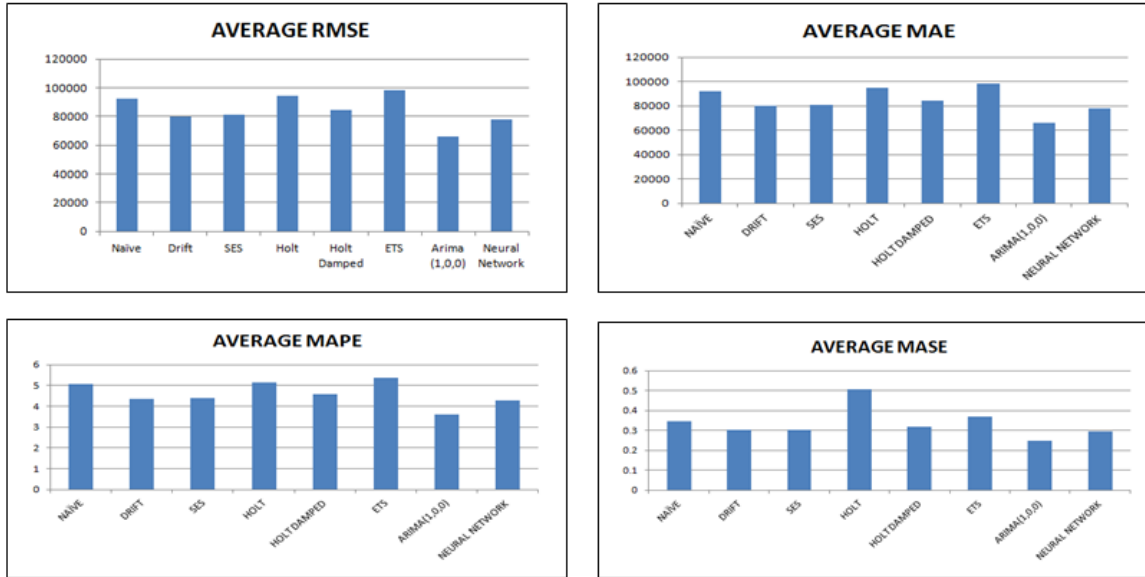


Fig.17. Comparison of error terms for 1-month forecast from Different Models

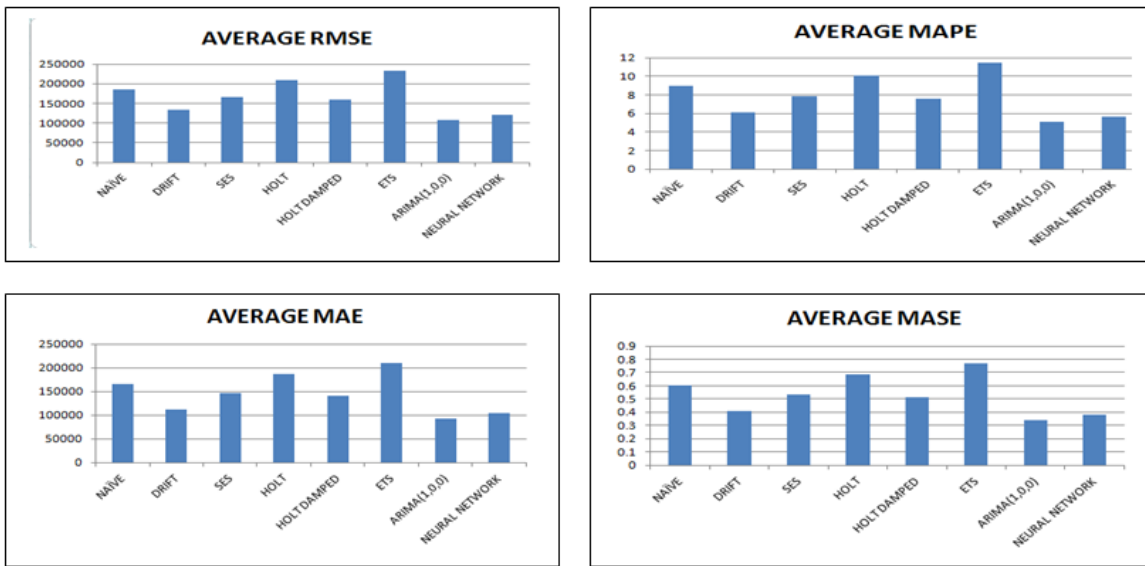


Fig.18. Comparison of error terms for 3-month forecast from Different Models

The result show forecasted values obtained data from different techniques used for prediction. The forecasted values obtained from Naïve, SES technique, Holt, ETS and neural network techniques have a significant difference as compared to actual test data set. For 1-month ahead forecast the average values of RMSE and MAE for Naïve, Drift, SES, Holt, Holt Damped, ETS, ARIMA and neural network are 92298, 79893.42, 81046.02, 94482.26, 84463.35, 98208.45, 65898.34 and 77788.43 respectively. The average value of error MAPE for Naïve, Drift, SES, Holt, Holt Damped, ETS, ARIMA and neural network is 5.06 %, 4.35%, 4.39%, 5.41%, 4.58%, 5.38%, 3.61% and 4.26% respectively. The another error component MASE for Naïve, Drift, SES, Holt, Holt Damped, ETS, ARIMA and neural network are 0.346744, 0.301997, 0.303965, 0.506124, 0.316636, 0.367113, 0.246216 and 0.294938 respectively. It is evident that value of all the error components RMSE, MAE, MAPE and MASE is less for ARIMA (1,0,0) in

comparison to other forecasting techniques. Similarly for 3-month ahead forecast the value of all the error components are less for ARIMA (1,0,0) forecast as can be seen from table no-8 and figure 18. The significance of low error values is that forecasting results obtained from ARIMA (1,0,0) are closer to the test data set as compared to other forecasting techniques.

V. CONCLUSION

This paper begins with an introduction of the importance of forecasting in relation to consumption of natural gas in industry. The foremost aim was to evaluate and compare the seven univariate forecasting techniques for forecasting the consumption of natural gas in industry. Due to limited data availability, forecasting horizon of one month and three month ahead forecast is selected. The out-of-sample forecasts are calculated by applying

various forecasting techniques. Error components RMSE, MAE, MAPE, and MASE are used to compare the forecasted values of different forecasting techniques. It has evaluated from the comparison that all the error components are less for ARIMA(1,0,0) technique and the forecasted values obtained are closer to the actual consumption of natural gas.

ACKNOWLEDGMENT

The corresponding author is grateful to organization CFFP, BHEL, Haridwar for sharing their valuable contents with us.

REFERENCES

- [1] Robinson, *Planning and forecasting techniques*. London: Weidenfeld& Nicolson, 1972.
- [2] "EIA - International Energy Outlook 2017", *Eia.gov*, 2018. [Online]. Available: http://www.eia.gov/forecasts/ieo/nat_gas.cfm. [Accessed: 05- Aug- 2017].
- [3] Wheeler, *Fossil fuels*. Edina, Minn.: Abdo Pub., 2008.
- [4] G. C. Leung "Natural Gas as a Clean Fuel." *Handbook of Clean Energy Systems*.
- [5] "India Natural Gas Market Size by Type 2019- TechSci Research", *Techsci research.com*, 2017. [Online]. Available: <http://www.techsci research.com/report/india-natural-gas-market-forecast-and-opportunities-2019/497.html>. [Accessed: 10- Aug- 2017].
- [6] B. Abraham and J. Ledolter, *Statistical Methods for Forecasting*. New York, NY: John Wiley & Sons, 2009..
- [7] R. Hyndman, *Forecasting with exponential smoothing*. Berlin: Springer, 2008.
- [8] P. S. Kalekar, Time series forecasting using holt-winters exponential smoothing. In :*Kanwal Rekhi School of Information Technology*, pp. 1-13, 2004.
- [9] R. Snyder, A. Koehler, R. Hyndman and J. Ord, "Exponential smoothing models: Means and variances for lead-time demand", *European Journal of Operational Research*, vol. 158, no. 2, pp. 444-455, 2004.
- [10] R. Snyder, A. Koehler, and J. Ord, "Forecasting for inventory control with exponential smoothing", *International Journal of Forecasting*, vol. 18, no. 1, pp. 5-18, 2002.
- [11] C. Chatfield and M. Yar, "Holt-Winters Forecasting: Some Practical Issues", *The Statistician*, vol. 37, no. 2, p. 129, 1988.
- [12] H. Grubb and A. Mason, "Long lead-time forecasting of UK air passengers by Holt–Winters methods with damped trend", *International Journal of Forecasting*, vol. 17, no. 1, pp. 71-82, 2001.
- [13] A. Koehler, R. Snyder, and J. Ord, "Forecasting models and prediction intervals for the multiplicative Holt–Winters method", *International Journal of Forecasting*, vol. 17, no. 2, pp. 269-286, 2001.
- [14] F. GÜmrah, D. Katircioglu, Y. Aykan, "Modeling of Gas Demand Using Degree-Day Concept: Case Study for Ankara", *Energy Sources*, vol. 23, no. 2, pp. 101-114, 2001.
- [15] F. GORUCU, "Evaluation and Forecasting of Gas Consumption by Statistical Analysis", *Energy Sources*, vol. 26, no. 3, pp. 267-276, 2004.
- [16] H. ARAS and N. ARAS, "Forecasting Residential Natural Gas Demand", *Energy Sources*, vol. 26, no. 5, pp. 463-472, 2004.
- [17] M. Akpınar, and N. Yumusak, Forecasting household natural gas consumption with ARIMA model: A case study of removing cycle. In: *7th IEEE International Conference on Application of Information and Communication Technologies (AICT)*, pp. 1-6, 2013.
- [18] C. Villacorta Cardoso and G. Lima Cruz, "Forecasting Natural Gas Consumption using ARIMA Models and Artificial Neural Networks", *IEEE Latin America Transactions*, vol. 14, no. 5, pp. 2233-2238, 2016.
- [19] L. Liu and M. Lin, "Forecasting residential consumption of natural gas using monthly and quarterly time series", *International Journal of Forecasting*, vol. 7, no. 1, pp. 3-16, 1991.
- [20] M. Akkurt, O. F. Demirel, and S. Zaim, Forecasting Turkey's natural gas consumption by using time series methods. *European Journal of Economic and Political Studies*, Vol. 3, no. 2, pp. 1-21, 2010.
- [21] F. Faisal, Time series ARIMA forecasting of natural gas consumption in Bangladesh's power sector, *Elixir Prod. Mgmt*, 2012.
- [22] C. H. Wei, and Y. C. Yang, A study on transit containers forecast in Kaohsiung port: applying artificial neural networks to evaluating input variables. *Journal of the Chinese Institute Transportation*, vol. 11, no. 3, pp. 1-20, 1999.
- [23] W. Lam, P. Ng, W. Seabrooke and E. Hui, "Forecasts and Reliability Analysis of Port Cargo Throughput in Hong Kong", *Journal of Urban Planning and Development*, vol. 130, no. 3, pp. 133-144, 2004.
- [24] G. Zhang and D. Kline, "Quarterly Time-Series Forecasting With Neural Networks", *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1800-1814, 2007.
- [25] M. Fung, "Forecasting Hong Kong's container throughput: an error-correction model", *Journal of Forecasting*, vol. 21, no. 1, pp. 69-80, 2001.
- [26] E. Hui, W. Seabrooke and G. Wong, "Forecasting Cargo Throughput for the Port of Hong Kong: Error Correction Model Approach", *Journal of Urban Planning and Development*, vol. 130, no. 4, pp. 195-203, 2004.
- [27] R. Kizilaslan, and B. Karlik, Combination of neural networks forecasters for monthly natural gas consumption prediction. *Neural network world*, vol. 19, no. 2, 2009.
- [28] M. Akpınar, M. F. Adak, and N. Yumusak, Forecasting natural gas consumption with hybrid neural networks—Artificial bee colony. In: *2nd IEEE International Conference on Intelligent Energy and Power Systems (IEPS)*, pp. 1-6, 2016.
- [29] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman, (1998). *Forecasting methods and applications*. John Wiley& Sons. Inc, New York, 1998.
- [30] R. Hyndman and G. Athanasopoulos, *Forecasting*. [Heathmont]: OTexts, 2016 E. Hannan, *Multiple Time Series*. Hoboken: John Wiley & Sons, Inc., 2009.
- [31] S. MAKRIDAKIS and M. HIBON, "ARMA Models and the Box-Jenkins Methodology", *Journal of Forecasting*, vol. 16, no. 3, pp. 147-163, 1997.
- [32] C. Grillenzoni, "ARIMA Processes with ARIMA Parameters", *Journal of Business & Economic Statistics*, vol. 11, no. 2, p. 235, 1993.
- [33] Freeman and D. Skapura, *Neural networks*. Reading, Mass: Addison-Wesley Publishing Company, 1991.
- [34] S. Haykin, *Neural Networks*. Prentice Hall Upper Saddle River, 1999.
- [35] B. White and F. Rosenblatt, "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms", *The American Journal of Psychology*, vol. 76, no. 4, p. 705, 1963.

- [36] Robjhyndman.com, 2018. [Online]. Available: <http://robjhyndman.com/talks/RevolutionR/2-Toolbox.pdf>. [Accessed: 15- Jan- 2017].
- [37] "Understanding the construction and interpretation of forecast evaluation statistics using computer-based tutorial exercises The Economics Network", Economicsnetwork.ac.uk, 2018. [Online]. Available: https://www.economicsnetwork.ac.uk/showcase/cook_for_ecast. [Accessed: 14- Apr- 2017].
- [38] K. Aho, D. Derryberry and T. Peterson, "Model selection for ecologists: the worldviews of AIC and BIC", *Ecology*, vol. 95, no. 3, pp. 631-636, 2014.
- [39] R. Hyndman and A. Koehler, "Another look at measures of forecast accuracy", *International Journal of Forecasting*, vol. 22, no. 4, pp. 679-688, 2006.
- [40] Neelam Mishra, Hemant Kumar Soni, Sanjiv Sharma, A K Upadhyay, "Development and Analysis of Artificial Neural Network Models for Rainfall Prediction by Using Time-Series Data", *International Journal of Intelligent Systems and Applications(IJISA)*, Vol.10, No.1, pp.16-23, 2018. DOI: 10.5815/ijisa.2018.01.03.
- [41] Vikalp Ravi Jain, Manisha Gupta, Raj Mohan Singh, "Analysis and Prediction of Individual Stock Prices of Financial Sector Companies in NIFTY50", *International Journal of Information Engineering and Electronic Business(IJIEEB)*, Vol.10, No.2, pp. 33-41, 2018. DOI: 10.5815/ijieeb.2018.02.05.
- [42] Er. Garima Jain, Bhawna Mallick, "A Study of Time Series Models ARIMA and ETS", *International Journal of Modern Education and Computer Science(IJMECS)*, Vol.9, No.4, pp.57-63, 2017. DOI: 10.5815/ijmeecs.2017.04.07.
- [43] Mehdi Khashei, Mohammad Ali Montazeri, Mehdi Bijari, "Comparison of Four Interval ARIMA-base Time Series Methods for Exchange Rate Forecasting", *International Journal of Mathematical Sciences and Computing(IJMCS)*, Vol.1, No.1, pp.21-34, 2015. DOI: 10.5815/ijmsc.2015.01.03.

Authors' Profiles



Iram Naim was born on May 26, 1984. She received the B. Tech. degree from M. J. P. Rohilkhand University, Bareilly, in year 2007. She did M. Tech from Aligarh Muslim University, Aligarh in the year 2011. Currently, she is a Ph. D. Scholar in Indian Institute of Technology (IIT), Roorkee. In 2008, she joined the M.J.P.

Rohilkhand University, Bareilly as an Assistant Professor. Her published articles list is as follows: User feedback based metasearching using neural network (International Journal of Machine Learning and Cybernetics (2015) 6(2)), Metasearching using modified rough set based rank aggregation (International Conference on Multimedia, Signal Processing and Communication Technologies, IMPACT 2011 (2011)). Her recently presented articles include Effect of Energy Resource Consumption Forecasting on Procurement Strategy: (In 11th ISDSI International Conference, Indian Institute of Management (IIM) Tiruchirappalli, Tamil Nadu, India.) Forecasting of Energy for Manufacturers using Arima and Neural network (5th international conference on business analytics and intelligence, Institute of Management (IIM) Bangalore) etc. Her research interests include Data Mining, Machine Learning, Neural Network, and Time series forecasting methods.



Tripti Mahara received the B. E degree in Computer Science & Engineering from Sardar Patel University in 1999, M.Tech and Ph.D. from Industrial and Management Engineering, I.I.T Kanpur in 2004 and 2009, respectively. Currently, she is Assistant Professor with the Department of Polymer and Process Engineering IIT Roorkee, India.

She has published over 18 referred journals and conference papers. She works in the area of Recommender System and Predictive Analytics and Data Mining.

How to cite this paper: Iram Naim, Tripti Mahara, "Comparative Analysis of Univariate Forecasting Techniques for Industrial Natural Gas Consumption", *International Journal of Image, Graphics and Signal Processing(IJIGSP)*, Vol.10, No.5, pp. 33-44, 2018. DOI: 10.5815/ijigsp.2018.05.04