International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST218375

A Review on Forecasting Models of Natural Gas

Dr. Meenakshi Thalor[1], Ritesh Choudhary[2], Ajay Jangid[2], Deep Gandhecha[2], Rishab Bhat[2]
[1]HOD of Information Technology Department, AISSMS Institute of Information Technology, Pune, Maharashtra, India
[2]Department of Information Technology, AISSMS Institute of Information Technology, Pune, Maharashtra, India

Article Info
Volume 8, Issue 3
Page Number : 251-258

Publication Issue
May-June-2021

Article History
Accepted : 18 May 2021
Published : 24 May 2021

ABSTRACT

This work gives away a survey of different published papers related to forecasting techniques in natural gas demand and consumption. Demand forecasting plays an important role in the company's decision making and inventory management process. It is important for local natural gas distributors to accurately predict the natural gas needs of their customers. While classifying the recent natural gas demand forecasting techniques, we have taken into account their methodologies, approach, data size, performance, results and limitations. Aim of this survey paper is to present a classified study of liquefied petroleum gas and natural gas related assets forecasting algorithms. This study provides readers with an understanding of the recent research in the natural gas supply-demand forecasting techniques.

Keywords — Natural gas, Artificial Neural Networks, Time Series, Statistical Models

I. INTRODUCTION

Natural Gas is an important natural resource and is likewise the essential source of cooking fuel. In the Liquefied Petroleum Cylinder business, one of the foremost critical resources is the LPG cylinder. Due to increasing competition in the market, it is necessary to make the right decision and plan future business related events such as expansion and production. Demand forecasting plays a vital role in the decision making of a business, as it reduces risk related to business activities and helps it to take effective decisions. For firms having production at the mass level, the importance of future demand forecasting is more. Many researchers have proposed forecasting models of liquefied petroleum gas future sales, demand, daily consumption, and ups and downs in prices. There are different forecasting ranges in natural gas, like year ahead monthly, month-ahead daily and day-ahead demand forecasting approaches. Good forecasting helps a firm in better planning related to business targets.

In this work, the criteria for classifying the papers include methodologies or algorithms, approach, magnitude of data, results and comparative performance. In specific, we compare different models proposed by the researchers wherever possible, along with their specific advantages and limitations. Aim of this paper is to present a classified study of liquefied petroleum gas or natural gas related assets forecasting algorithms and scope for future research. The Second Section presents an overview of published papers in last 6 years. Section 3 describes the models and application areas. Finally we end this survey with major conclusions and directions for future research in Section 4.

II. LITERATURE SURVEY

Several researchers have proposed forecasting and prediction models to estimate fossil fuel demand or consumption. For several years, solely common classical models are used for natural gas forecasting by distributors in native regions. The foremost common strategies to forecast demand and sales are statistical approaches, either exponential smoothing strategies or autoregressive models. Auto regressive moving average is a technique that calculates the trends in data and is extremely helpful for forecasting short-run trends. In this a new demand number is obtained for a fixed time range, while keeping the time period locked.[5] Exponential smoothing focuses a lot on most up-to-date data, giving more weight to the foremost recent observations. Whereas time series models try and forecast future values taking think into consideration solely the patterns in the historical data, they don't consider any factors that may influence the future. For more details on time series see, [1,2].

One of the meta-learning applications is to select the best methodology from the set of accessible ones. Such method is termed as algorithm selection and it was discussed, among others, in [7]. Using this approach, it is possible to settle on from totally different categories of machine learning algorithms, e.g. Support Vector Machines, Neural

Networks, Random Forests, Decision Trees, and Logistic Regression the one that's expected to relinquish the best results on a given information set. Wanshuai Hu and Tao (2014) presented a combined model that is predicted to show improved results using wavelet transform and back-propagation neural network to forecast gas load in a pipeline.[9] The authors in [1] proposed a multi-model techniques and practices to overcome the drawbacks of individual algorithms for forecasting gas demand. Researchers concluded that using this technique factors such as overfitting, underfitting are minimised and additional strong results were achieved. Viewing monthly consumption, Horacio Paggi and Franco Robledo [4] used a Neural Network model to predict consumption of 2

different kinds of LPG cylinders. Jointly considering deterministic demand, reverse logistics, temperature, wind, gas load researchers[3,6,8] proposed NN models differing within the training dataset size and forecasting horizon.

Zlatan Sicanica and Zdravko Oklopcic [10] have shown a comparative study of Neural Networks models and classical Machine learning models for daily gas demand prediction in Croatia.

III. EXISTING MODELS

Predictive problems arise in many areas, and existing predictive literature offers similar applications in many different fields. This section covers the various approaches and applications used to predict natural gas supply demand and related activities. Let's look at the details of each model used in the literature and also areas where natural gas forecasting applies, highlighting the features of some of the methods

A. Statistical Models

1) ARIMA

ARIMA models show that there is a linear relationship between serial data. It is the most common class of models for predicting time series data. Random variables that are time-variate series and if their statistical characteristics are all constant over time then ARIMA can be the most suitable choice. These models have been applied in various applications of natural gas forecasting till date. Junhui Guo (2019) used the WTI oil worth data of the last thirty-three years to train the ARIMA model. But the training data had variations in price of oil between 2000 and 2013, creating it troublesome to train with primitive ARIMA models. The researchers remarked that a time variate data is often called stationary if for a given T period, the log(WTI oil price) should have the same distribution over the time T. Additionally he added, stationarity is often checked by conducting a Dickey-Fuller test. Further he decomposes the log oil price data into trend, seasonality by setting the seasonality decomposition frequency to 360 days. Analyzing the results ARIMA model had lowest R squared value when compared to different prediction algorithms, and is least capable of handling data with high variations.

2) Multiple Linear Regression

MLR is used to determine a numerical relationship among a number of random features or variables. In simple terms, MLR analyses how closely different independent variables are associated or related to one dependent variable. MLR assumes that there exists a linear relationship between both dependent and independent variables, also there is no major correlation between the independent variables. M. Kozielski and Z. Laskarzewski (2019) proposed a meta-model that learns the association between the models and the characteristics of customers LPG usage, which can be a composition of many factors. The meta-model can be

expressed as a linear combination of base-models and it has a form like: $y = a1 * x1 + \ldots + am * xm + b$, where: m is the total number of base models considered.

The information set provided was the hourly measurements of three hundred LPG cylinders. Every information entry included date and time , percent of gas level of the cylinders and room temperature. The researchers concluded that this technique works well when the prediction horizon is of weekly or monthly basis.

3) Exponential Smoothing

Exponential smoothing is a time-series forecasting approach used for univariate data with a systemic or seasonal trend. This forecasting method is similar in terms that it uses a weighted sum of past observations, but the model can explicitly use an exponentially decreasing weighted sum for learning through past observations. Aldina Correia (2017) used a time series exponential smoothing model to predict monthly consumption of a Portuguese Company. The dataset provided by the corporation includes details like sales, returns of LPG cylinders, stocks, assets of the 2 varieties of cylinders differing in terms of weight. The Researchers assume that propane gas cylinders are principally used for cooking, boiling water, thus whenever there's a decrease in temperature, the consumption of fossil fuel can increase in Portugal. They used centered moving averages technique, and calculated twelve seasonal indexes which specified the quantity of fuel gas consumption in every month. The seasonal coefficients obtained were used to deseasonalize the data and then Holt's method was applied to forecast fuel gas cylinder sales and returns. The model predicts best once the training information is massive or else the seasonality coefficients will be far away from actual values.

B. Artificial Neural Network Models

1) Conventional Neural Network

ANN is an analytical method based on learning processes driven by complex biological systems, specifically by the neural function of the human brain. ANN is a distributed information processing system that consists of several simple computational elements that interact between each other. Artificial Neural Networks betrays distinctive characteristics and possess the ability to learn through examples and create micro functional relationships with data. Horacio Paggi (2014) used a weekly prediction Artificial Neural Network model for time series prediction of propane gas cylinders. The training information was provided by the state owned petroleum company and it collected the daily consumption of natural gas measured in liters, atmospheric temperatures, wind, humidity, rains and max-min temperature. Inorder to predict the weekly consumption, the daily information was summed up from Sunday to Saturday. The researches claimed that the linear correlation between the variables max-temperature and sales is larger than with min-temperature. They used MLPs as network Topology because it is repeatedly used by researchers and the period length of fifty two was set for training. Prabodh Kumar Pradhan (2018) proposed an integrated Conventional Fusion Artificial Neural network model for predicting the consumption of natural gas. They provided 5 different inputs for training specifically daily demand, temperature, humidity, rainfall to the NN as an associate input layer. The neural network is trained using a back propagation algorithm by the researchers. The analyzed results show that projected CFANN performed higher than over different classical prediction and ANN models.

2) Fuzzy Cognitive Neural Network

Another delicate computing method which could be exceptionally capable of developing complex and nonlinear Neural Network models is neuro-fuzzy approach. Katarzyna Poczeta (2018) proposed a process model, which mixes FCM technique with a common ANN to construct a cascaded model. Structure optimization genetic algorithms pick out the foremost vital concepts (nodes) for fuzzy cognitive maps[6]. These concepts (nodes) in FCM are used as the inputs for training artificial neural networks. After providing the selected data attributes in a type of nodes as the inputs for artificial neural networks, it is allowed to learn using back propagation methods. The presented FCM-ANN hybrid model proves to have adequate fitting and forecasting capacity, on comparing results with the standard ANN models.

3) LSTM

The LSTM network is a type of recurrent neural network (RNN) that can learn from the long-term dependency between sequential data time points. LSTMs are often applied to a variety of deep learning tasks that largely require prediction based on past information. LSTM is mostly used as a key tool for time series prediction issues and has pertinency in many alternative scientific domains. Athanasios Anagnostis (2019) applied the LSTM for the needs of a day-ahead natural gas demand prediction in 3 distribution center's in Greece nation. The historical information of natural gas demands within the distribution center of over 5 years, was used as associate input data. For building a LSTM model previous 4 years of data was fed and in this case, they set the timestamp to be 364 values so as to predict consecutive 364 days. The results showed addressing a forecasting problem with low degree of input variables can be done constructively with LSTM models however with increasing variables performance of LSTM drops.

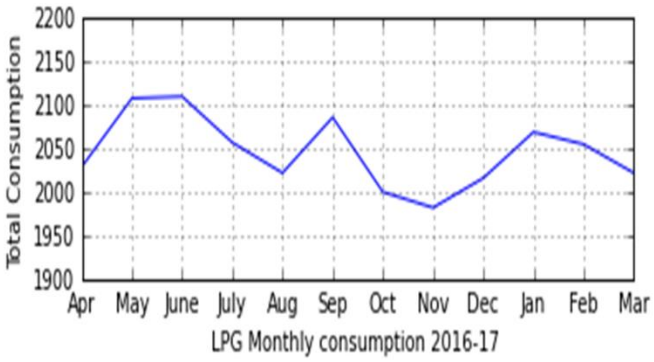
C. Combinational or Multi Models

In addition to the traditional models mentioned above, some researchers have chosen to use two or more models to predict natural gas in their work. These attempts are made to achieve good results, test a hypothesis model and compare results with conventional models. Aldina Correia (2020) planned an ensemble methodology using time series prediction, MLRs models and ANNs for forecasting monthly consumption. The information provided by a Portuguese petroleum company enclosed details like sales, returns of LPG cylinders. The Holt's formula is employed in Exponential smoothing so as to notice trends, MLR models are used to incorporate many variable factors and single layer architecture with four hidden nodes was used as ANN. These models are collaborated into a single model by adding a definition of weights to each one of them. The researcher's concluded that multi-model methodology derived by combining different approaches, allowed to get an accurate forecasting even with meagre information. Zlatan Sicanica (2018) revealed a relative study of Machine learning and neural network algorithms for natural gas demand forecasting. Researchers used the three years of daily consumption data of natural gas measured in cu metre, Croatia. Models like neural networks, decision trees, support vector regression, linear regression, lasso linear regression were compared. The final results showed large differences between the validation and test errors, they described this could be due to bias or overfitting in the training phase.

IV. RESULTS AND D ISCUSSIONS

For comparing the performance of the listed methodologies we have used the monthly LPG consumption data of Indian Market of the year 2016-17. The dataset consists of monthly LPG consumption in a total of 35 regions in India, majorly labeled as per states and union territory. The total consumption in financial year 2016-17 accounts for a sum of 22.5 million ton of LPG.

Figure. 1 Monthly LPG consumption of year 2016-17



The models were trained against input parameters like temperature, region(state & union territory), consumption in region(LPG sales), timestamp and season label. Below Table 1 shows the performance comparison of different models with parameters RMSE and R-squared evaluation parameters.

Table. I Comparison of Existing models against LPG sales data 2016-17

Paper No	Model	RMSE	R-Squared
[1],[7]	Multiple Linear Regression	36.0585	0.6819

[1],[10]	Multi Model	4.8621	0.9437
[2],[12]	Exponential Smoothing	21.7884	0.8863
[3],[13]	LSTM	4.1843	0.9501
[4],[9], [11],[14]	Artificial Neural Networks	26.6530	0.8159
[5]	ARIMA	31.0923	0.7517
[5],[8], [15]	Conventional Neural Network	19.2087	0.9029
[6]	Fuzzy Cognitive Neural Network	18.5296	0.9177

The evaluation parameters we have used to measure the performance metrics and efficiency of different models are RMSE and R-squared values and are shown in Figure 2 and 3.

Figure. 2 Comparison of R-squared

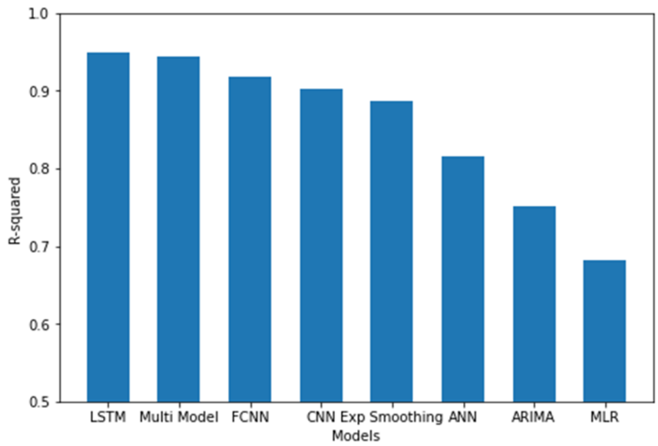
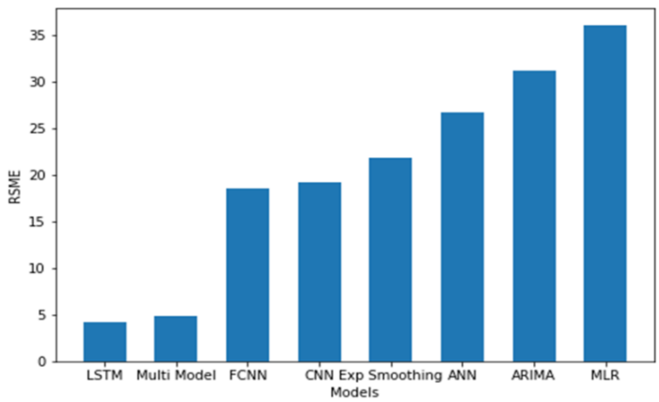


Figure. 3 Comparison of RMSE



Due to the monthly cyclic nature of input dataset LSTM model outperformed here. Also the multi model approach consisting of Neural Networks, Lasso Linear Regression,

SVR and Decision Tree gave nearly the same results as of LSTM. The Multi model approach being adaptive to different input coefficients gave RMSE of 4.8621 and R-squared value 0.9437. If the input training data would have been even large, then the Multi model approach would be the best performer due to its adaptive nature of learning.

Fuzzy Cognitive Neural Network and Conventional Neural Network showed nearly the same performance and can be preferred over typical Artificial Neural Networks. However there was a good difference between FCNN and the top performing LSTM, this is due to the number of hidden layers and the aberration in the dataset. Exponential smoothing gave average results due to its nature of weighted sum approach for near past observations. Forming a multi model using Exponential Smoothing combined with LSTM can give more appropriate results. Arima being the moving averages technique works best with limited input features, as we can see with increased parameters the results are in the bottom list. Multiple linear regression being the simplest technique alone gave the least 0.6819 R-squared value. This experiment shows that with increasing dimensionality of input parameters the Multi Model and enhanced models achieve better clutch over the classical models.

V. CONCLUSION

Forecasting the consumption of natural gas with low degree of input variables can be done effectively with LSTM models, however the performance of LSTM starts to decline greatly with increasing input variables. The combined or multi models eliminates the drawbacks of individual methods like overfitting, underfitting, along with maintaining their benefits and thus providing more accurate forecasts. Moreover, if the input data contains tons of nonlinearity and seasonality coefficient then Exponential Smoothing models deal alright with them. CFANN or FCM-ANN have adequate fitting and forecasting capability, when comparing results with the standard ANN models. ARIMA models in comparison to different forecasting algorithms are among the least capable models of handling data with high variations and nonlinearity. Multiple linear regression being the simplest technique requires no irregularity or anomaly in the input data of consumption and seasonality, also the prediction horizon is weekly or monthly basis.

Another space to be explored is the usage of additional sophisticated models. Meta modelling can be the best match if one wishes to include usage characteristics of consumers for weekly or monthly data. Moreover, convolution neural networks (CNN), that has been extensively employed in computer vision and image recognition, hasn't been explored a lot for its application on time series. Gated Recurrent Unit(GRU), may provide with higher insight of the time series modelling and achieve higher prediction accuracy. This could be a possible focus for future analysis, finding an additional sophisticated model that would improve the accuracy of forecasting models. The forecasting results may be improved using heuristic optimizations, ensemble regressors or RNNs, these models have capability of improving accuracy and robustness of basic machine learning models.

VI. REFERENCES

[1]. Aldina Correia, Cristina Lopes, Eliana Costae Silva, Magda Monteiro,Rui Borges Lopes.,"A multi-model methodology for forecasting sales and returns of liquefied petroleum gas cylinders",Springer-Verlag London Ltd.,2020.

[2]. Aldina Correia, Eliana Costae Silva, Cristina Lopes, Claudio Henriques, Fabio Henriques, Mariana Pinto, Magda Monteiro, Rui Borges Lopes and Ana Sapata., "LPG Demand Forecast using Time Series",Proceedings of the 17 th International Conference on Computational and Mathematical Methods in Science and Engineering,CMMSE 2017 4–8 July,2017,Page 656 of 2288

[3]. Athanasios Anagnostis, Elpiniki Papageorgiou, Vasileios Dafopoulos, Dionysios Bochtis.,"Applying Long Short-Term Memory Networks for natural gas demand prediction", 10th International Conference on Information, Intelligence, Systems and Applications (IISA) 2019.

[4]. Horacio Paggi,Franco Robledo.,"A Neural Networks Based Model For The Prediction Of The Bottled Propane Gas Sales",International Conference on Mathematics and Computers in Sciences and in Industry,2014,Page 69-74.

[5]. Junhui Guo.,"Oil price forecast using deep learning and ARIMA", 2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI),2019,Page 241-247.

[6]. Katarzyna Poczeta,Elpiniki I. Papageorgiou., "Implementing Fuzzy Cognitive Maps with neural networks for natural gas prediction",IEEE 30th International Conference on Tools with Artificial Intelligence,2018,Page 1026-1032.

[7]. Michal Kozielski, Zbigniew Laskarzewki, "Matching a Model to a User - Application of Meta-Learning to LPG Consumption Prediction",2019.

[8]. Prabodh Kumar Pradhan, Sunil Dhal, Nilayam Kumar Kamila., "Artificial Neural Network Conventional Fusion Forecasting Model for Natural Gas Consumption",2018,Page 2200-2205

[9]. Wanshuai Hu,Zeyuan Tao,Dongyu Guo,Zixiao Pan.,"Natural Gas Prediction Model Based on Wavelet Transform and BP Neural Network",The 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC) May 18-20, 2018, Nanjing, China,2018,Page 952-955.
[10]. L Zlatan Sicanica, Zdravko Oklopcic.,"Countywide Natural Gas Consumption Forecast, a Machine Learning Approach",MIPRO 2018 , May 21-25, 2018, Opatija Croatia,2018,Page 1070-1073.
[11]. Fernandez J, Cruz-Ramirez M, "Sensitivity versus accuracy in ensemble models of artificial neural networks". Neural Computer Application(2018) ,Page:289–305
[12]. Domino M, Laskarzewki Z, "Classification of LPG clients using the hurst exponent and the correlation co-efficient." Theor. Appl. Inf. (2015) Page:27-39
[13]. M. Ganjkani, S. N. Fallah, "Computational intelligence on short-term load forecasting: ," Energies. 2019, Page:109-131
[14]. Szoplek, J ."Forecasting of natural gas consumption with artificial neural networks." Energy.2015, Page:208-220.
[15]. Prabodh K, Sunil D, and Nilayam K."Neural Network based Forecasting for Natural Gas Consumption." Communications on Applied Electronics 2017,Page:45-80
[16]. Bendat I, Moscar R,"A new stochastic multi source approach to improve the accuracy of the sales forecasts." Foresight,2017,Page:48–64

Cite this Article

Dr. Meenakshi Thalor, Ritesh Choudhary, Ajay Jangid, Deep Gandhecha, Rishab Bhat, " A Review on Forecasting Models of Natural Gas", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 3, pp. 251-258, May-June 2021. Available at doi : https://doi.org/10.32628/IJSRST218375
Journal URL : https://ijsrst.com/IJSRST218375