Prediction of Petroleum Price in India

[1]M.S. Khan, [2]A. Saikia, [3]Dr. Sushanta Tripathy
[1,2,3]School of Mechanical Engineering, KIIT University, Bhubaneswar, Odisha, INDIA

Abstract
Forecasting of oil prices has always been a matter of great importance due to its influence in driving a country's economy. As a matter of fact, the petroleum industry is considered to be the biggest contributor in the industrial sector in terms of providing raw materials to the other industries and generating revenues. Due to the non-linear and unpredictable nature of the oil prices, a lot of forecasting techniques have been developed and used to check whether they are capable of forecasting the oil prices satisfactorily. In this paper, two widely used techniques, ARIMA analysis and GMDH Neural network has been used to forecast the prices of four petroleum products such as Petrol, Diesel, LPG & Kerosene in India for a three month period (February 2015 to April 2015). The results obtained are compared with the actual prices for the above time period. It is observed that the overall accuracy considering all the four petroleum products shows promising results thus justifying them capable of forecasting the prices of the different petroleum products in India. When both the techniques are compared, ARIMA modelling shows better results (97.99% accuracy) as compared to that of the GMDH neural network method which is 96.97%.

Keywords
Petroleum Products, Prices, ARIMA Modelling, GMDH Neural Network

I. Introduction

Petroleum can be described as a type of fossil fuel, which is found below the earth's surface. It is formed when dead organisms decompose in large quantities due to the high pressure and temperature they are subjected to when they are pressed and buried under the rock bodies. Mainly consisting of compounds of carbon and hydrogen, petroleum can be processed and refined to obtain a large number of fuels and other related products. Popular petroleum derivatives are LPG, diesel, petrol, jet fuel, kerosene, paraffin, wax, tar, etc.

The prices of petroleum products always make the front page in India. Fluctuation of prices are always met with a strong response in this nation. The price structure of these products is the sole reason for this kind of response. Depending on the type, each product produced from crude oil has a different pricing structure. It is because of these different pricing structures that some product's price over a period of time remains stable whereas for another product, the price keeps on changing within a short period of time.

Petrol, in India, is already a deregulated commodity since 2010. This means that the government does not provide any compensation to the oil marketing companies for losses obtained on petrol, unlike diesel, LPG and other regulated items. On the other hand, it is not possible for diesel to undergo deregulation since more than 96 percent of its usage is in the manufacturing and industrial sectors while the remaining part of it is used as a form of fuel in automobiles. Hence, any increase in the price level of diesel will lead to increase in prices of the goods that are produced due to the increasing levels of cost of transportation and production. The other products concerned include LPG and kerosene. These products are referred to as controlled products as their prices remain stable over the course of time. This is because these products are consumed by the urban/rural poor and hence their price absorption capacities are less.

Increase or decrease in oil prices can have major political, social and economic impacts which can drastically affect the direction of a country's growth. The biggest example of this was seen in the late 80s when it was widely believed that decrease in oil prices was one of the biggest contributors in causing the disintegration of the USSR. Decrease in price of oil does not necessarily mean negative impact. It is observed that oil importing nations like India in fact benefit from fall in oil prices with the GDP of the country increasing by a small margin of one to two percent during that time period. Given the importance of oil prices in driving inflation and economic activities, prediction of prices of oil and its different derivatives becomes a vital criterion in analysing and assessing the future economy of the country, which brings this topic down to this paper. In this journal, a three month forecast (February 2015 to April 2015) of the prices of four different petroleum products in India is carried out using a time series modelling technique (ARIMA modelling) and a neural network technique (GMDH modelling).

II. Literature review

A number of research papers have been published on prediction of oil price considering various methodology. Primarily, two techniques, neural network and time series models individually or with various combinations are employed to forecast oil price. The review of some of the research works related to crude oil prices which used artificial neural networking (ANN) have been presented in this part. Neha Sehgal and Krishan K. Pandey, in the year 2014, proposed a method, comprising of two stages, called the MI3 Algorithm. The algorithm was used to determine the major parameters which affect the oil prices. The results confirmed that the proposed algorithm, which used cascaded neural network, multi layered perception neural networking, and general regression neural network for forecasting purposes, produced better results when compared to other traditional methods [1]. In 2013, semi supervised learning was used by Hyunjung Shin, et.al. to study what effect do the impact of economic factors have in determining the price of oil. The change in price trend was studied using this algorithm [2]. I. Haider, S. Kulkarni and H. Pan presented a feed forward kind of ANN, consisting of three layers, meant for forecasting oil prices on a short term basis. Results showed high level of accuracy even for series containing non linearity or noise [3]. Edmundo G. de Souza e Silva, et. al. investigated a non-linear time series model to predict the price movements of oil in the future. Called the hidden Markov model (HMM), they used this to develop a new forecasting methodology which could be used as a decision aiding mechanism for factors which effect the oil market[4]. Hassan Mohammadi and Lixian Su examined various ARIMA-GARCH models used for forecasting of the volatility of weekly spot prices of crude oil in over several markets all over the world from the time period of Jan 1997 to October 2009. Results revealed that the AGARCH model showed better

performance when compared with the other models [5]. These techniques are also used to forecast other important commodities. In a study, Thomas Kriechbaumer, et. al. assessed an improved combined wavelet ARIMA approach for forecasting of monthly prices of different base metals like aluminium, copper, lead and zinc. The study showed that the ARIMA model's performance is similar to that of the plain and traditional models when it comes to forecasting the prices of base metals [6]. In 2012, Zhong-bing Zhou and Xiu-cheng Dong published a paper to examine the seasonal nature of oil imports in China to provide the necessary assistance that is required for production planning and inventory control to the stake holders. X-12 ARIMA modelling method was used and the results showed that the quarterly series showed better adjustment to the seasonality nature as compared to the monthly series [7]. Bahram Ghorbani,et.al. proposed a hybrid GMDH type neural network model, aided by generic algorithm, to find out the required polynomial correlation necessary to estimate the viscosity of oil which has a direct relationship with the oil price. The results approved the fact that the models were highly accurate in estimating the viscosity of the crude oils found in Iran [8]. In 2012 Mingzhu Zhang,et. al. analysed the traditional GMDH network and introduced the concept of diversity to improve its immunity towards noise. The new D-GMDH model was found out to have good immunity towards noise and performed well even in noisy atmospheres [9]. Again in 2013 they, along with Xin Gu, Bing Zhu proposed a forecasting model called the D-GMDH model to study the economy of a provincial territory in China. The results obtained from this model were compared to that of the different GMDH models and the highly popular ARIMA model and showed better consistency than those methods [10].

III. Methodology

From the extensive literature review, it is found that two techniques, namely, time series and ANN models are mostly used for forecasting various domains including crude oil prices. In the time series method, ARIMA model and GMDH models in ANN are widely used for forecasting. Therefore, it is proposed to forecast the price of petroleum products in India using this two techniques and compare the results.

A. Arima Modelling

In time series analysis, the ARIMA model is considered to be the most widely used and most effective model when it comes to forecasting any activity or event. Also known as Box-Jenkins method, this process can be used for forecasting activities as they have the ability to fit into the provided time series data. The integrated part of the ARIMA model helps in reducing non-linearity or non-stationarity. One of its main reasons of its popularity is its flexibility. It has the ability to fit almost any data. Only the previous data of the concerned activity is required, thus reducing the total number of parameters. ARIMA model is also useful when the forecast is short term in nature. A lot of softwares offer the Box-Jenkins or ARIMA modelling in their statistical packages. Some of the most commonly used softwares are IBM SPSS, Stata, SAS, R, Matlab etc.

B. GMDH-Type Neural Network

Artificial neural networks, commonly known as ANN, is a type of process which is strongly inspired from the human neuron. ANN can be used to forecast to complicated networks and data which are beyond the comprehensive ability of both humans and other forecasting techniques. A neural network has an adaptive learning ability allowing it to take decisions on the basis of the initial data.

The artificial neural network (ANN) reduces the complexity of an activity by dividing it into more simple parts with the help of a neuron. Even though the functions are simple, it is the entire setup of the network which takes a lot of time to set up and get accustomed to the training data. The setup can be made simpler if there is a method to effectively select and identify the structure and accordingly provide the required weights to the connections between the different layers. A GMDH type neural network can be used to overcome this limitation. In a GMDH network, the neurons present are more complex than a normal ANN neuron. A GMDH neuron contains a polynomial transfer function and the connections present between the different network layers are simplified. An automatic algorithm is used to provide the required weights and identify and select the necessary structure. This saves a lot of time as the user does not have to define weights. The neuron itself will provide the appropriate weights with the aid of the algorithm. As a result, GMDH is also termed as polynomial neural networks due to the polynomial nature of the neurons. The GMDH network analysis has been found to be accurate in many cases related to forecasting. Also, it is less time consuming that ANN as the weights of the connections between the neuron layers are automatically set due to the polynomial nature of the neuron. Some of the most popular softwares containing this feature are GMDH Shell, GEvom, KnowledgeMiner etc.

IV. Data Collection and Analysis

The four vital petroleum products in India whose prices will be forecasted are Petrol, Diesel, Kerosene and Domestic LPG. Short term forecasting consisting of a 3-month period will be done using ARIMA analysis and GMDH neural network. Long term forecasting is not preferred when oil prices are forecasted because when it comes to forecasting a time period with a range spreading years, the price of oil can change due to different economic and geo-political reasons which are beyond the reach of any existing forecasting technique. A lot of non-linearity is present in the trend of oil prices that is why short term forecasting of three months is preferred. The time period that will be forecasted is from February 2015 to April 2015. The prices of petrol, diesel, LPG and kerosene starting from January 2010 to January 2015 is considered as the sample data. The sample data was collected from the official website of Indian Oil Corporation Limited (www.iocl.com). The retail prices which is followed in Delhi is considered in this paper.

A. ARIMA Analysis

IBM SPSS Statistics Version 21 software is used to forecast oil price using ARIMA model. An ARIMA(1,0,0) model with no transformation is used for forecasting the price of the oil products. This is because among all the models, the ARIMA(1,0,0) model shows the best results when it comes to producing or fitting the price models of the different petroleum products. The following tables and graphs were obtained after analysis of the prices using IBM SPSS:

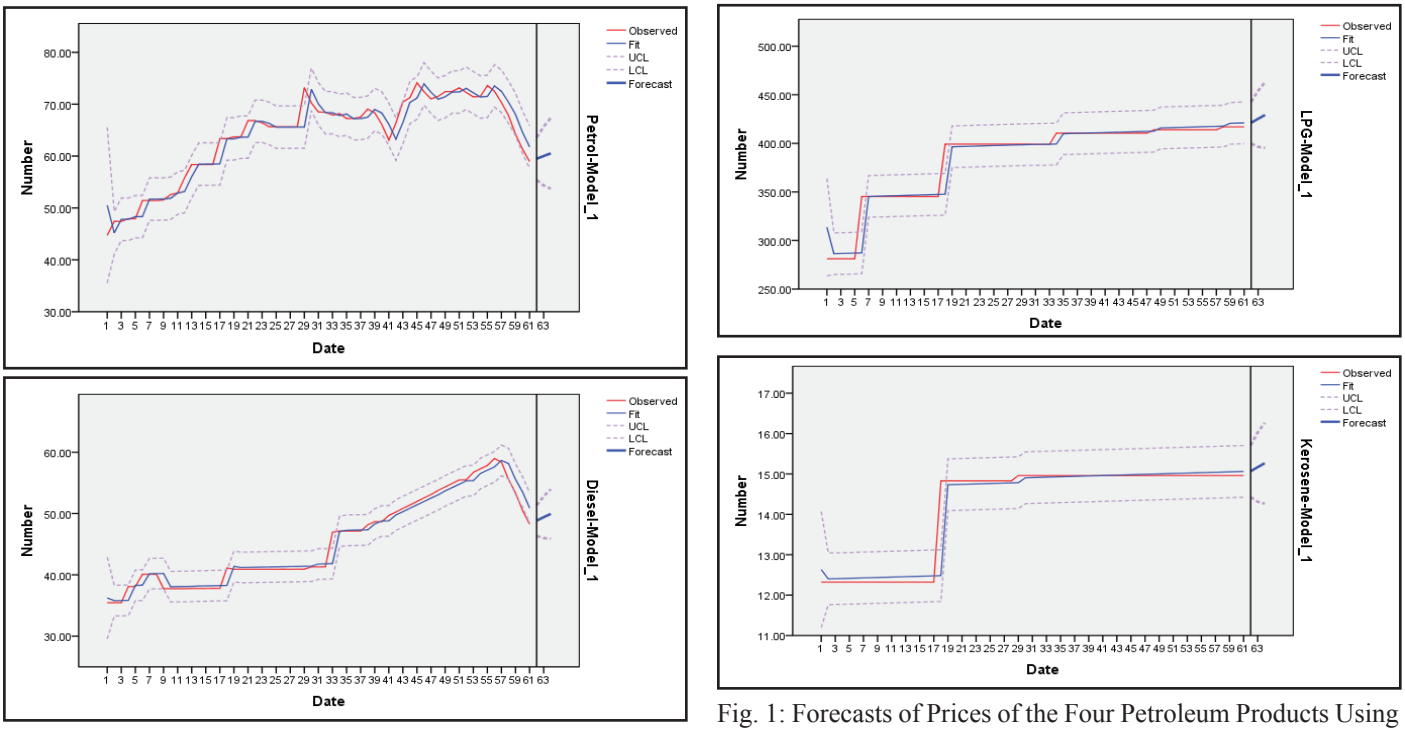

Fig. 1: Forecasts of Prices of the Four Petroleum Products Using ARIMA (1,0,0) Model

Table 1: The Errors and Accuracy of The ARIMA Model (Petrol)

Month	Foreca-sted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	59.48	57.31	2.17	3.786424708	96.213575	98.173351
Mar-15	60.01	60.49	0.48	0.79351959	99.20648	
Apr-15	60.54	60	0.54	0.9	99.1	

Table 2: The Errors and Accuracy of The ARIMA Model (Diesel)

Month	Forecasted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	48.84	46.62	2.22	4.761904762	95.238095	97.21513536
Mar-15	49.4	49.71	0.31	0.623616978	99.376383	
Apr-15	49.94	48.5	1.44	2.969072165	97.030928	

Table 3: The Errors and Accuracy of The ARIMA Model (LPG)

Month	Forecasted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	421.28	417	4.28	1.026378897	98.973621	98.01438
Mar-15	425.34	417	8.34	2	98	
Apr-15	429.22	417	12.22	2.930455635	97.069544	

Table 4: The Errors and Accuracy of The ARIMA Model (Kerosene)

Month	Forecasted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	15.07	14.96	0.11	0.735294118	99.2647059	98.59625
Mar-15	15.17	14.96	0.21	1.403743316	98.5962567	
Apr-15	15.27	14.96	0.31	2.072192513	97.9278075	

B. GMDH Analysis

GMDH, an ANN technique which is a highly valued software for forecasting any activity. In this paper GMDH Shell DS 3.6.2 software have been used to analyse and forecasting the prices of petroleum products in India. The following graphs show the results obtained from the GMDH analysis.

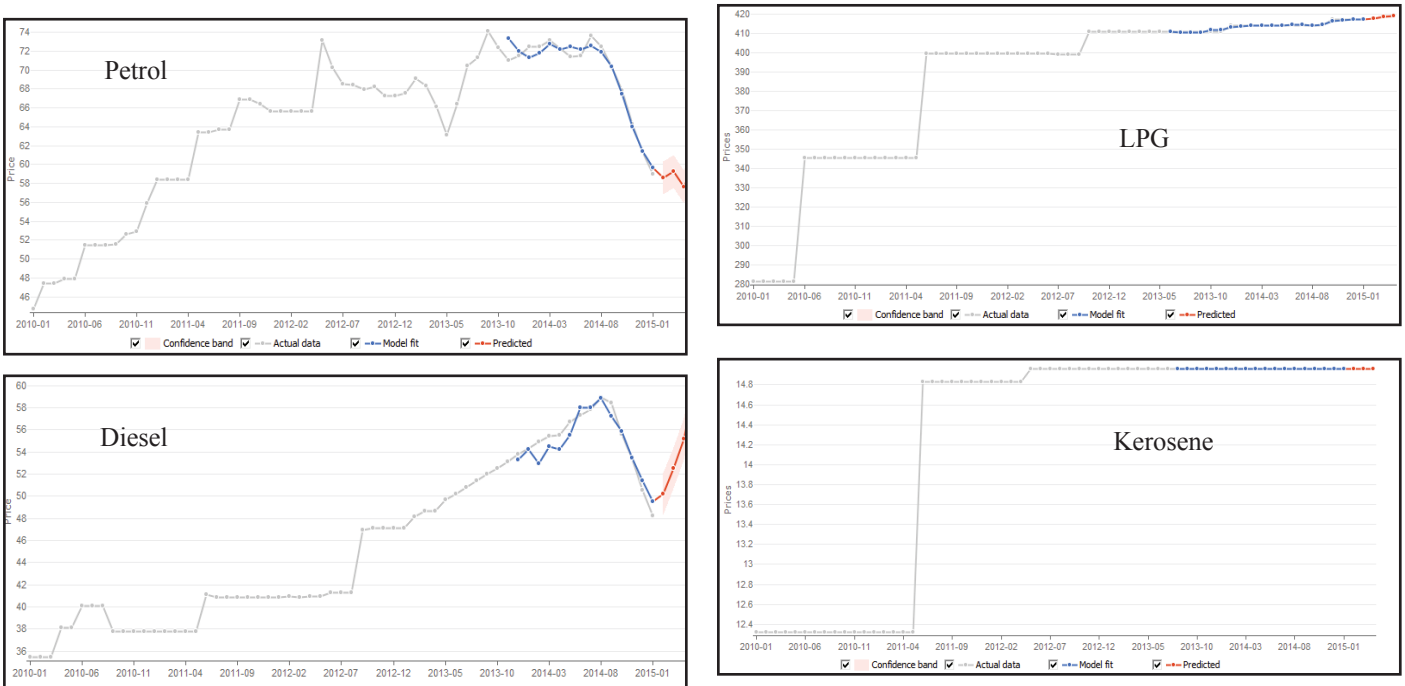

Fig. 2: Forecasts of the Prices of Four Petroleum Products (Petrol, Diesel, LPG, Kerosene) Using GMDH Shell (From Top Left To Bottom Right) Respectively

Table 5: The Errors and Accuracy of the GMDH Method (Petrol)

Month	Forecasted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	58.6069	57.31	1.2969	2.262955854	97.7370441	
Mar-15	59.3063	60.49	1.1837	1.956852372	98.0431476	97.27589
Apr-15	57.6285	60	2.3715	3.9525	96.0475	

Table 6: The Errors and Accuracy of the GMDH Method (Diesel)

Month	Forecasted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	50.2021	46.62	3.5821	7.683612184	92.3163878	
Mar-15	52.5655	49.71	2.8555	5.744317039	94.255683	90.92457
Apr-15	55.1922	48.5	6.6922	13.79835052	86.2016495	

Table 7: The Errors and Accuracy of the GMDH Method (LPG)

Month	Forecasted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	417.572	417	0.572	0.137170264	99.862829	
Mar-15	418.459	417	1.459	0.349880096	99.650119	99.68920
Apr-15	418.857	417	1.857	0.445323741	99.554676	

Table 8: The Errors and Accuracy of the GMDH Method (Kerosene)

Month	Forecasted Price	Actual Price	MAE	MAPE	Accuracy (%age)	Mean Accuracy
Feb-15	14.96	14.96	0	0	100	
Mar-15	14.96	14.96	0	0	100	100
Apr-15	14.96	14.96	0	0	100	

Note: Mean Absolute Error (MAE) = Forecasted Price – Actual Price
Mean Absolute Percentage Error (MAPE) = (MAE/Actual Price)*100
Accuracy (%age) = 100- MAPE
Mean Accuracy = Average of the accuracies in predicting the prices for the three months period

V. Results and Discussions

As evident from Table 9 below, mean accuracy has been used as the criterion to check whether the two methods are accurate in predicting the prices of petroleum products in India. Results show that both the methods were highly effective in forecasting the prices of all the four petroleum products with accuracy higher than 97% in all cases except one i.e. diesel prices using GMDH Shell

was 90.9246%. Another interesting observation can be seen from the table while predicting the prices of products like kerosene, LPG using the GMDH neural network. In both cases the accuracy clocked the 100% mark (99.68% and 100% for LPG and kerosene respectively). However, when both the methods are compared against one another, the ARIMA technique has a slightly higher overall accuracy than the GMDH neural network technique i.e. 97.99% for ARIMA and 96.97% for GMDH neural network).

Table 9: Comparison of Accuracies of ARIMA Model and GMDH Model

Product	Mean Accuracy of ARIMA(1,0,0) Model	Mean Accuracy of GMDH Neural Network
Petrol	98.1733519	97.2759
Diesel	97.2151353	90.9246
LPG	98.0144	99.689209
Kerosene	98.5963	100
Overall Accuracy	97.9997968	96.97242725

Note: Overall Accuracy for a method is the average of the mean accuracies obtained for each product using that particular method.

One of the main factors that determines the price trend of a petroleum product in a country is the economic state of that particular country. A stable economy ensures a stable price trend whereas any hike or change in the oil price can easily alter the price trend. Clearly from the results, it can be stated that during the time period February 2015 to April 2015, India's economy was very stable which reflects on the forecasted prices that show a high level of accuracy. Accuracy increases when the economy is stable because it becomes easier to predict the prices during such kind of situations. The same explanation also validates for the price trend of controlled products like LPG and kerosene whose prices do not vary over a course of time. This is the reason why the accuracy of forecasting the prices of these products are nearly 100% using both the techniques. The high level of accuracies clearly justify the effectiveness of the two techniques in predicting activities which are non-linear in nature. Both the techniques are able to adjust to the non- linearity present in the sample data and use their necessary algorithms to forecast the future activity. The GMDH technique showed higher accuracy while calculating prices of controlled products as compared to petrol and diesel. The ARIMA modelling showed consistent performances for all the four products with accuracy higher than 97% in each case.

VI. Conclusion

Predicting oil prices has never been an easy task because of its non-linearity nature. Over the years, a lot of forecasting techniques have been tried and tested to solve this problem. In this work, two of those techniques, ARIMA analysis and GMDH Neural networking, have been used to forecast the prices of four petroleum products in India for a three month period (February 2015 to April 2015). The results obtained show excellent levels of accuracy when compared with the actual prices for the same time period. The analysis that was done in this paper validated the effectiveness of both the techniques, ARIMA modelling and GMDH neural networking, in predicting the price of petroleum products in India for a period of three months. With an overall accuracy of around 97.99%, ARIMA modelling slightly outperformed the GMDH technique in predicting the required prices, which has an accuracy of 96.97%. However, the GMDH technique was found to be more accurate in calculating the prices of products whose prices are stable or controlled. Judging from the high level of accuracy in both the methods, it can be concluded that India's economy was quite stable during the forecasting time period as it is comparatively easier to forecast prices when the economy is stable.

References

[1] Neha Sehgal, Krishan K. Pandey,"The Drivers of Oil Prices – A MI3 Algorithm Approach", Energy Procedia, Volume 61, 2014, pp. 509-512.

[2] Hyunjung Shin, Tianya Hou, Kanghee Park, Chan-Kyoo Park, Sunghee Choi,"Prediction of movement direction in crude oil prices based on semi-supervised learning", Decision Support Systems, Vol. 55, Issue 1, April 2013, pp. 348-358.

[3] I. Haider, S. Kulkarni, H. Pan,"Forecasting Model for Crude OIL PRICES Based on Artificial Neural Networks", International Conference on Intelligent Sensors, Sensor networks and Information Processing, IEEE (2008), pp. 103–108.

[4] Edmundo G. de Souza e Silva, Luiz F.L. Legey, Edmundo A. de Souza e Silva,"Forecasting oil price trends using wavelets and hidden Markov models", Energy Economics, Vol. 32, Issue 6, November 2010, pp. 1507-1519.

[5] Hassan Mohammadi, Lixian Su,"International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models", Energy Economics, Vol. 32, Issue 5, September 2010, pp. 1001-1008.

[6] Thomas Kriechbaumer, Andrew Angus, David Parsons, Monica Rivas Casado,"An improved wavelet–ARIMA approach for forecasting metal prices", Resources Policy, Vol. 39, March 2014, pp. 32-41.

[7] Zhong-bing Zhou, Xiu-cheng Dong,"Analysis about the seasonality of China's crude oil import based on X-12-ARIMA", Energy, Vol. 42, Issue 1, June 2012, pp. 281-288.

[8] Bahram Ghorbani, Masoud Ziabasharhagh, Majid Amidpou. "A hybrid artificial neural network and genetic algorithm for predicting viscosity of Iranian crude oils", Journal of Natural Gas Science and Engineering, Vol. 18, May 2014, pp. 312-323.

[9] Mingzhu Zhang, Changzheng He, Panos Liatsis,"A D-GMDH model for time series forecasting", Expert Systems with Applications, Vol. 39, Issue 5, April 2012, pp. 5711-5716.

[10] Mingzhu Zhang, Changzheng He, Xin Gu, Panos Liatsis, Bing Zhu,"D-GMDH: A novel inductive modelling approach in the forecasting of the industrial economy", Economic Modelling, Vol. 30, January 2013, pp. 514-520.

ISSN : 2249-5762 (Online) | ISSN : 2249-5770 (Print)

Dr. Mohammed Sadique Khan has completed his Ph.D. from National Institute of Technology, Rourkela, India in 2008. He has 24 years of teaching experience. His area of interest includes Industrial Engineering, Operations management, Quality Engineering and non conventional manufacturing. Presently, he is serving as Associate Professor in the School of Mechanical Engineering, KIIT University, Bhubaneswar, India. More than 20 published journal papers and conferences papers are in his credit.

Mr. Abhinash Saikia is currently perusing his post graduation in the Department of Mechanical Engineering of KIIT university, Bhubaneswar, India.

Sushanta Tripathy is currently working as Professor in the School of Mechanical Engineering in KIIT University, Bhubaneswar, India. He has completed his PhD from the Department of Industrial Engineering and Management, Indian Institute of Technology, Kharagpur. He has over 25 years of teaching, research and industry experience in India and abroad. He has published a number of papers in national and international conferences and journals. His major areas of interest include Production operations management, Multivariate analysis, Service operations management, Supply chain management and Productivity management. He is a fellow of the Institution of Engineers (India).