IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 11, Issue 6 Ver. II (Nov. - Dec. 2015), PP 23-36
www.iosrjournals.org

Application of Regression and Neural Network Models in Computing Forecasts for Crude Oil Productions

Augustine Pwasong [I a], Caleb Nimyel [b]

[a] *Department of Mathematics, University of Jos, 2084 Jos, Nigeria* [b]*Plateau State Polytechnic Barkin-Ladi, Plateau State, Nigeria*

Abstract: *In this study, a quadratic regression model and a two layered layer recurrent neural network (TLLRNN) method were used to model forecasting performance of the daily crude oil production data of the Nigerian National Petroleum Corporation (NNPC). The two methods were applied on the difference series and log difference series of the NNPC series. The results indicates that the two layered layer recurrent neural network model have better forecasting per-formance greater than the quadratic regression method based on the mean error square sense. The root mean square error (RMSE) and the mean absolute error (MAE) were applied to ascertain the assertion that the two layered layer recurrent neural network method have better forecasting performance greater than the quadratic regression method. The outcome of the analysis also indicates that modeling forecasting performance of the NNPC data with the log dif-ference series of the data gives greater forecasting performances than modeling with the difference series of the NNPC data irrespective of the method used in modeling with the series. These results were achieved from 1 day ahead pre-dictions, 3 days ahead predictions and 5 days ahead predictions for 50 days sample length, 100 days sample length, 200 days sample length, 400 days sample length and 800 days sample length. Autocorrelation functions emerging from the increment series, that is, difference series and log difference series of the daily crude oil production data of the NNPC indicates significant autocorrelations and significant partial autocorrelations. The data used in this study is a time series data obtained from the daily crude oil production of the Nigerian National Petroleum Corporation (NNPC) for a period of six years (1st January, 2008 - 31st December, 2013). The analysis for this study was simulated using MATLAB software, version 8.03*

Keywords: *Regression, neural network, root mean square error, mean absolute error, forecasting.*

I. Introduction

A mineral resource product which is vital to global economy is crude oil. Strictly speaking, crude oil is a key factor for the economic advancement of industrialized and developing countries as well as undeveloped countries respectively. Besides, political proceedings, extreme meteorological conditions, speculation in fiscal market amidst others, are foremost events that characterized the eventful style of crude oil market, which intensifies the level of price instability in the oil markets.

The crude oil industry in Nigeria is the largest industry. Oil delivered around 90 percent of foreign exchange incomes, about 80 percent of federal government proceeds and enhances the progress rate of the country's gross domestic product (GDP). Ever since the Royal Dutch Shell discovered oil in the Niger Delta in 1956, specifically in Oloibiri, Bayelsa State, the crude oil industry has been flawed by political and economic discord mainly due to a long antiquity of corrupt military governments, civilian governments and collaboration of multinational corporations, particularly Royal Dutch Shell. About six oil firms namely - Shell, Elf, Agip, Mobil, Chevron and Texaco controls the oil industry in Nigeria. The aforementioned oil companies collectively dominate about 98 percent of the oil re-serves and operational possessions. There are three key players in the Nigeria oil industry which include the Federal

Ministry of Petroleum Resources, the Nigerian National Petroleum Corporation (NNPC) and the crude oil prospecting companies which comprises the multinational companies as well as indigenous companies as asserted by Baghebo [1].

In this study we intend to used data obtained from the daily crude oil production of the Nigerian National Petroleum Corporation (NNPC) for a period of six years (1st January, 2008 - 31st December, 2013). The data constitute a time series data in view of the daily pattern of its occurrence which is depicting a regular pattern. The daily crude oil production series is applied in this study to establish empirical instances of smearing the quadratic regression model and a two layered layer recurrent neural network in forecasting the daily crude oil production series of the NNPC.

Forecasting begins with assumptions based on the organization's experience, knowledge and judgment. These estimates are anticipated into the coming months or years using one or more methods such as Box-Jenkins models, Delphi method, exponential smoothing, moving averages, regression analysis, and trend projection. Since any variation in the assumptions will result in a similar variation in forecasting, the method of sensitivity analysis is used which assigns a range of values to the uncertain variables. A forecast should not be confused with a budget.

This study intends to provide a comparison study of the quadratic regression model and a two layer recurrent neural network technique in forecasting the daily crude oil production of the Nigeria National Petroleum Corporation (NNPC). The forecasting performances of the quadratic regression method and the two layered layer recurrent neural network method will be assessed through the determination of both the root of mean square error (RMSE) and the mean absolute error (MAE).

The remaining part of this study is organized as follows. Section 2 describes the quadratic regression technique and the two layered layer recurrent neural network method contain in this paper. Section 3 presents simulation results on data obtained from the NNPC for a period of 6 years (2008-2013) using the quadratic regression method and the two layered layer recurrent neural network method contain in this study. Section 4 reports concluding remarks and future work.

II. Regression and neural network techniques

This study will focus on regression and neural network techniques to model and forecast the daily crude oil production of the NNPC. The regression technique considered here is the quadratic regression method which is a subsidiary of polynomial regression models. Neter et al. [2] declared that polynomial regression models are amid the supreme commonly used curvilinear response models in practice, for the reason that they can be easily handled as a distinct item of the general linear regression model given in equation (1) below:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \quad + \beta_{p\ 1} x_{i,p\ 1} + \epsilon_i \qquad (1)$$

Polynomial regression models can encompass one, two, or more than two predictor variables. Additionally, every single predictor variable may be current in several powers.

The method of linear least squares is a widespread statistical method and publications are huge. We will only momentarily appraise important advances of the least squares method. In 1805, Legendre published his book on comet orbits in Paris. He deliberated the method of least squares. This is the earliest publication on the least squares method. However Legendre only proclaimed but did not prove his results. Farebrother [3] explained that Laplace (1816) produced a sequential fitting technique for the method of least squares. His method is quite close to a con-temporary practice called the modified Gram-Schmidt orthogonalization procedure [3]. Cauchy (1836) anticipated a modest method for parameter estimation of linear models. The Cauchy method was later identified as an orthogonalization process, thoroughly associated to the results of Laplace [3].

A distinctive presentation of the technique of least squares was prearranged by Galton in 1886. Galton scrutinized the heights of parents and their developing children to advance intuitions into a congenital characteristic. Galton instituted that the heights of the children of both tall and short parents seemed to return or regress to the mean of the cluster. He deliberated this predisposition as a regression to the mean. Because of this and other works, Galton was considered the pioneer of regression, while Legendre and Gauss were the architects of regression as asserted by Tobia [4]

Healy [5] asserts that Pearson (1894) presented his technique of moment as an unconventional method for linear fitting problems. Pearson (1900) as well established a goodness-of-fit test, and the chi-square test, to discover if an assumed probability model sufficiently designate data being used. Both the technique of moment and the chi-square test are precisely significant to the least squares regression technique. The chi-square test is a normality trial applied for model residuals. An advancement of the technique of moment, the generalized method of moment (GMM), turns out to be a tremendously powerful estimation tool. The ordinary least squares (OLS) technique can be perceived as a special case of the GMM.

Elman [6] defines a recurrent neural network as a class of neural network that has recurrent connections, which permit a form of retention. This makes them appropriate for chronological prediction responsibilities with indiscriminate temporal dimensions. Recurrent neural networks is characterize by the feature that particular neuron outputs are fed back into the same layer or foregoing layers so that information drifts in equally forward and backward orders to perceive and cause time-varying configurations. Recurrent networks have feedback links that allow chronological in-formation to be characterized. Recurrent neural network (RNN) architecture uses particular hidden neurons to present feedback. These neurons institute the context layer of a network. The context layer plays a significant part in holding chronological information and aids the memory of the system by holding the state of the system before the next set of data is administered.

Tsoi and Back [7] explains that a typical RNN consists of five components: (1)neuron introduced by an activation function, (2) layer or a group of neurons, running and stopping at the same time, (3) link between neurons, (4) architecture and the organization of neurons interrelated by connections, and (5) clock, the operation sequence.

In another development Hopfield [8] anticipated a network that is used to stockpile steady target vectors. These steady vectors can be regarded as memories. The network is recursive. The purpose of the Hopfield network is to recover a pattern stockpiled in memory when offering an incomplete configuration or a noisy form of the configuration.

Jordan [9] offered two types of RNN networks: The local recurrent global feed-forward network, and the global recurrent global feed-forward network.

Recurrent neural networks are not lacking inadequacies. Its feedback acquaintances entail extra neurons, extra connections, a huge quantity of calculation, and a huge training set to create the RNN function. The entirety of these phenomena cause tough training and slow convergence. These are extracts from Atya and Parlos [10]. In this study, we concentrate our devotion on a two layered layer recurrent neural network (TLLRNN) which will be clearly illustrated in subsection 2.2 as well as a quadratic response function which will also be vividly illustrated in section 2.1.

2.1. Quadratic response function

We instigate by allowing for our polynomial regression model with two predictor variables elevated to the first and second powers specified by equation (2) below:

$$Yt = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_{11} x^2_{t-1} + \phi_{22} x^2_{t-2} + \phi_{12} x_{t-1} x_{t-2} + \epsilon_i \quad (2)$$

Equation (2)is a quadratic regression model with two predictor variables. The response function is given by equation (3) below:

$$E\{Y\} = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_{11} x^2_{t-1} + \phi_{22} x^2_{t-2} + \phi_{12} x_{t-1} x_{t-2} \quad (3)$$

which is the equation of a conic section. Equation (2) comprises distinct linear and quadratic constituents for both the two predictor variables and a cross-product term. The last represents the interaction effect between x_{t-1} and x_{t-2}. The coefficient ϕ_{12} is frequently called the interaction effect coefficient. These are extracts from [1]. In this study, we applied the least squares method to estimate the parameters of the quadratic response function (3). For the quadratic polynomial model (2) the parameters in the model can be easily obtained from the normal equations given below:

$$\sum Y_t = nb_0 + b_1\sum x_{t-1} + b_2\sum x_{t-2} + b_{11}\sum x^2_{t-1} + b_{22}\sum x^2_{t-2} + b_{12}\sum x_{t-1} x_{t-2}$$

$$\sum x_{t-1} Y_t = b_0 \sum x_{t-1} + b_1\sum x^2_{t-1}+ b_2\sum x_{t-1} x_{t-2} + b_{11}\sum x^3_{t-1} + b_{22}\sum x_{t-1} x^2_{t-2} + b_{12}\sum x^2_{t-1} x_{t-2}$$

$$\sum x_{t-2} Y_t = b_0 \sum x_{t-2} + b_1\sum x_{t-1} x_{t-2} + b_2\sum x^2_{t-2} + b_{11}\sum x^2_{t-1} x_{t-2} + b_{22}\sum x^3_{t-2} + b_{12}\sum x_{t-1} x^2_{t-2}$$

$$\sum x^2_{t-1} Y_t = b_0 \sum x^2_{t-1} + b_1\sum x^3_{t-1}+ b_2\sum x_{t-1} x_{t-2} + b_{11}\sum x^4_{t-1} + b_{22}\sum x^2_{t-1} x^2_{t-2} + b_{12}\sum x^3_{t-1} x_{t-2}$$

$$\sum x^2_{t-2} Y_t = b_0 \sum x^2_{t-2} + b_1\sum x_{t-1} x^2_{t-2} + b_2\sum x^3_{t-2} + b_{11}\sum x^2_{t-1} x^2_{t-2} + b_{22}\sum x^4_{t-2} + b_{12}\sum x_{t-1} x^3_{t-2}$$

$$\sum x_{t-1} x_{t-2} Y_t = b_0 \sum x^2_{t-1} x_{t-2} + b_1\sum x^2_{t-1} x_{t-2} + b_2\sum x_{t-1} x^2_{t-2} + b_{11}\sum x^3_{t-1} x_{t-2} + b_{22}\sum x_{t-1} x^3_{t-2} + b_{12}\sum x^2_{t-1} x^2_{t-2}$$

The cross product term in (2) is taken to be a second-order term, the same as ϕ_{11} x2 or ϕ_{22} x^2The motive may be seen by writing the last terms as $\phi_{11} x_{t-1} x_{t-1}$ and $\phi_{22} x_{t-2} x_{t-2}$ separately [2].

2.2. Layer recurrent neural network

The neural network technique to be used in this study is the Layer-Recurrent Network (LRN) and to be specific, a two layered layer recurrent neural network (TLLRNN). A prior basic form of this network was presented by Elman [Elma90]. Zhang et al.[11]. In the layer recurrent network there is a feedback loop, with a single delay, around each layer of the network excluding the last layer. The novel Elman network had merely two layers, and adopts a tansig transfer function for the hidden layer and a Purelin transfer function for the output layer. The novel Elman network was trained by means of an approximation to the back propagation algorithm. The Layrecnet command generalizes the Elman network to require a subjective number of layers and to have subjective transfer functions in every layer. The toolbox trains the LRN using exact versions of the gradient-based algorithms in Multilayer Neural Networks and Back propagation Training. The resulting network architecture in figure 1 exemplifies a two-layered layer recurrent neural network (TLLRN).

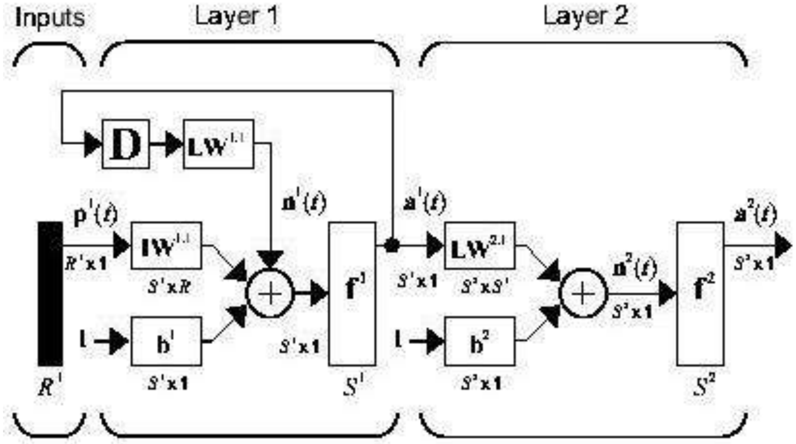

Figure 1: A two layered layer recurrent network architecture

Layer recurrent neural networks are alike to feedforward networks, apart from the fact that each layer has a recurrent connection with a tap delay linked with it. This permits the network to have an immeasurable dynamic response to time series input data. The layer recurrent network is analogous to time delay and distributed delay neural networks, which have predetermined input responses. These are assertions from Liu [12].

Basically, the two layered layer recurrent network possessed two major layers namely layer 1 and layer 2 as il-lustrated in figure 1. Layer 1 is a feedback layer that provides recurrence, while layer 2 is a feedforward layer that propagates information forward. In addition, the two layered layer recurrent network has input, output and hidden layers similar to other recurrent neural network. In this study, the input and output of the neural network is represented by x(t) and y(t) in that order. In figure 1, the bias terms are indicated by b^1 for layer 1 and b^2 for layer 2. The hidden layers for layers 1and 2 are indicated by $\mathbf{f^1}$ and $\mathbf{f^2}$ respectively. The notation $\mathbf{Lw}^{1,1}$ and $\mathbf{Lw}^{2,1}$ indicates feedforward connection weights from input to hidden layers and from hidden to output layers, respectively. In layer 1 which is the feedback layer, weighted sum of the delayed outputs of the hidden and output layers is served into the activation functions just as the feedforward layer neurons as explained by Aksu [13]. The output of layer 1 neurons are represented by $\mathbf{a}^1(t)$ and the output of layer 2 neurons by $\mathbf{a}^2(t)$ and applied to neurons of the hidden and output layers through $\mathbf{Lw}^{1,1}$ and $\mathbf{Lw}^{2,1}$ respectively. The notation **D** is the time delay or context layer. Following the time delay or context layer, we will use t to represent the time index.

The two layered layer recurrent neural network (TLLRNN), use dissimilar activation functions for its diverse types of neurons. It uses a sigmoid function or a hyperbolic tangent for usual hidden neurons, a time delay function for context neurons, and a linear function for output neurons. With these permutations, the TLLRNN network can estimate any function with any subjective precision just like the Elman network. The lone prerequisite is that the number of neurons in the hidden layer is sufficiently huge.

Training the two layered layer recurrent neural network (TLLRNN) can be done by the Levenberg-Marquardt algorithm, a backpropagation (BP) algorithm using second derivatives of errors short of computing the Hessian matrix. The calculations for two layered layer recurrent neural network (TLLRNN) network encompass additional equations than a feed-forward network due to feedback connections.

In the two layered layer recurrent neural network (TLLRNN), dissimilar categories of neurons have diverse activation functions given below:

$f^h(x(t)) = 1/(1 + e^{-x(t)})$ (activation function for hidden neurons) (4)

$f^0(x(t)) = x(t)$ (activation function for output neurons) (5)

$f^c(x(t)) = x(t-1)$ (activation function for time delay neurons)

For computations from t=1 to t=K, error of the system is defined as difference between reference output and the TLLRNN's output as in equation (7)

$$e(t) = y(t)\ \ \hat{y}(t) \qquad (7)$$

where $\hat{y}(t)$ is the reference output and y(t) is the TLLRNN's output. A fitness function using this error tries to optimize the network's weights[13]. At the commencement, outputs of the hidden and output layers are fixed to zero:

$$f(0) = 0, \quad y(0) = 0 \qquad (8)$$

The sum of the input weights $Lw^{1,1}$ and $Lw^{2,1}$ with the bias b^1 and b^2 are calculated by:

$$nwt_1(t) = [Lw^{1,1} f(t\ \ 1)] + b^1$$

$$nwt_2(t) = [Lw^{2,1} f(t\ \ 1)] + b^2 \qquad (9)$$

where $Lw^{1,1}$,$Lw^{2,1}$ and b^1, b^2, designate the input weights of the feedback layers and bias values of the feedback layer neurons, respectively. The net input to the outputs of the feedback layers are computed as:

$$f^1(t) = f^o(x(t))(nwt_1(t))$$

$$f^2(t) = f^o(x(t))(nwt_2(t)) \qquad (10)$$

where the f^o's are the activation functions of the feedback layer neurons. Computations of the net input of the outputs to the hidden layer neurons are as follows:

$$nwt_f(t) = [W_1 x(t)] + [Lw^{1,1} f(t)] + [Lw^{2,1} f(t)] + b_1$$

$$h(t) = f^h(x(t))(nwt_f(t)) \qquad (11)$$

where $nwt_f(t)$ is local field of the hidden layer neurons, W_1 is the weight between the input layer and hidden layer and b_1 is the bias applied to the hidden layer neurons. $Lw^{1,1}$ f (t) and $Lw^{2,1}$ f (t) represent the output weights of the feedback layer and $f^h(x(t))$ is the activation function of the hidden layer neurons. Computations of the net input to the context neuron of the context layer are as follows:

$$nwt_c(t) = [W_2h(t)] + [W_3y(t)] + b_3$$

$$c(t) = f^c(x(t))(nwt_c(t)) \qquad (12)$$

where $nwt_c(t)$ is local field of the context layer neurons, W_2 is the weight between the hidden layer and context layer and b_3 is the bias applied to the context layer neurons.W_3 represent the context weights of the feedback layer and $f^c(x(t))$ is the activation function of the context layer neurons.

III. Results and discussion

The data series for this study represents the daily crude oil production of Nigeria National Petroleum Corporation (NNPC). Figure 2 shows the whole picture of 2191 samples for a period of six years (1st January,2008 - 31st December, 2013) in barrels per day.

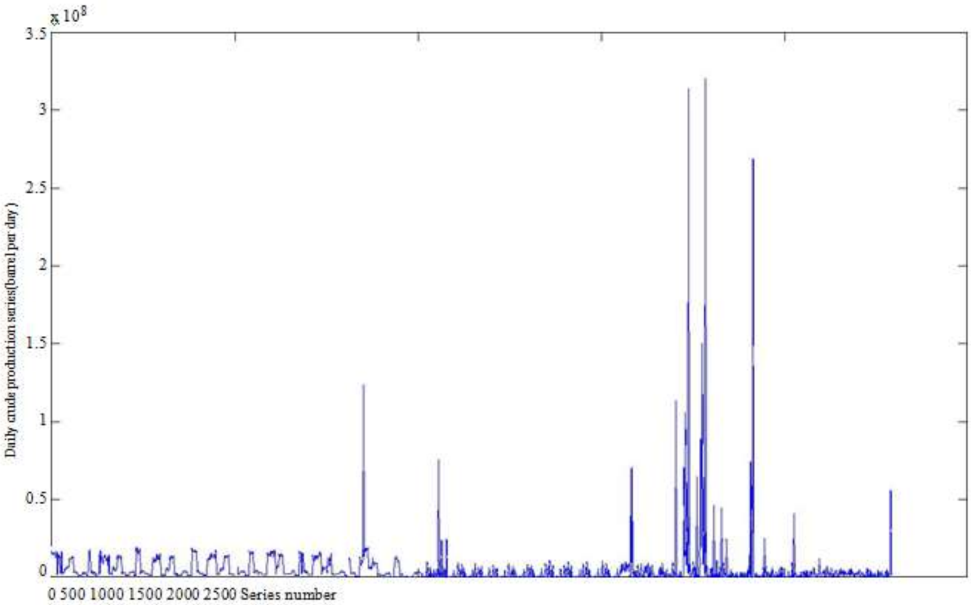

Figure 2: Daily Crude Oil Production of the NNPC

A sample autocorrelation picture of the daily crude oil production series is shown in figure 3. Figure 4 is the dia-gram of the sample partial autocorrelation function of the daily crude oil production series of the NNPC. Both figures points out that the daily crude oil production series of the NNPC are both autocorrelated and partially autocorrelated.

Sample Partial Autocorrelation Function

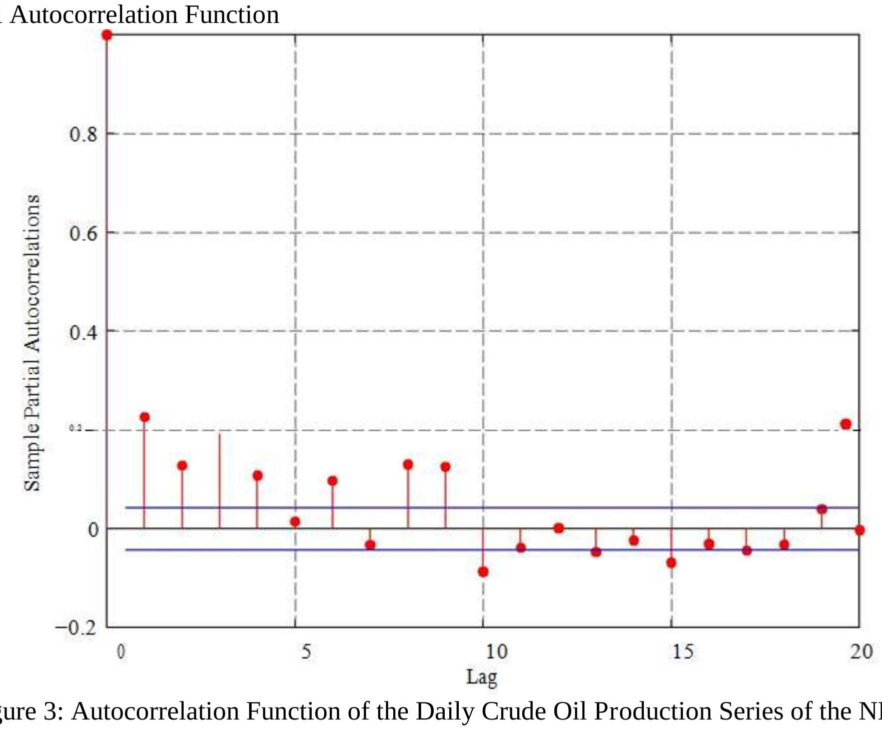

Figure 3: Autocorrelation Function of the Daily Crude Oil Production Series of the NNPC

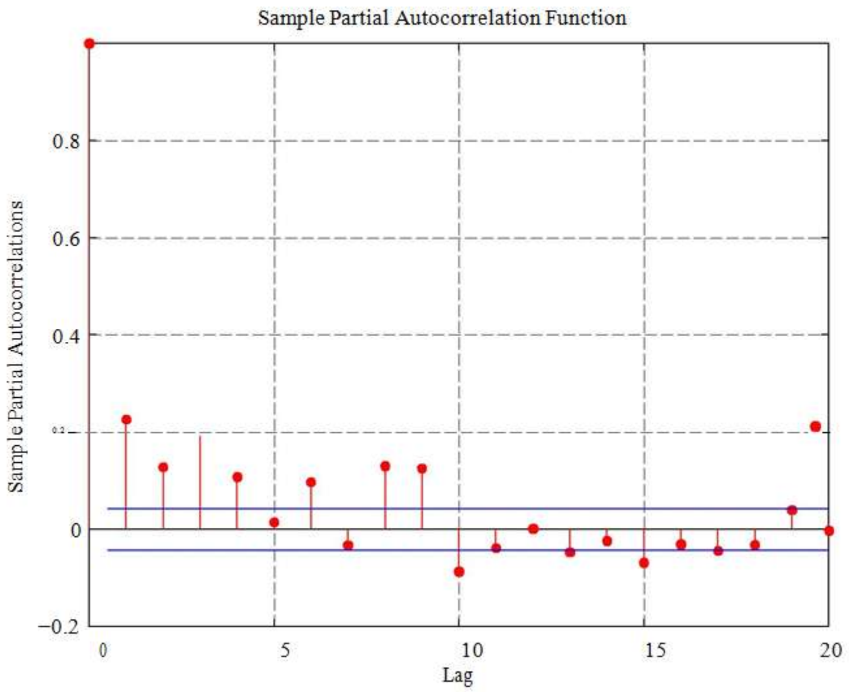

Figure 4: Partial Autocorrelation Function of the Daily Crude Oil Production Series of the NNPC

The increment series, that is, difference series and log difference series of the daily crude oil production series of the NNPC is used to conduct the Dickey-Fuller (DF) test to determine if the data series is stationary. In this study, we difference the original daily crude oil production series by using $X_t = Y_t \; Y_{t\;1}$ and $Q_t = \log(Y_t)$ $\log(Y_{t\;1})$ respectively, if Y_t is the original series. Figure 5 shows the picture of the daily crude oil production difference series of the NNPC. Figure 6 is the diagram of the sample autocorrelation function of the daily crude oil production difference series of the NNPC and figure 7 shows the diagram of the sample partial autocorrelation function of the daily crude oil production difference series of the NNPC. Both figures points out that the daily crude oil production difference series of the NNPC are both autocorrelated and partially autocorrelated.

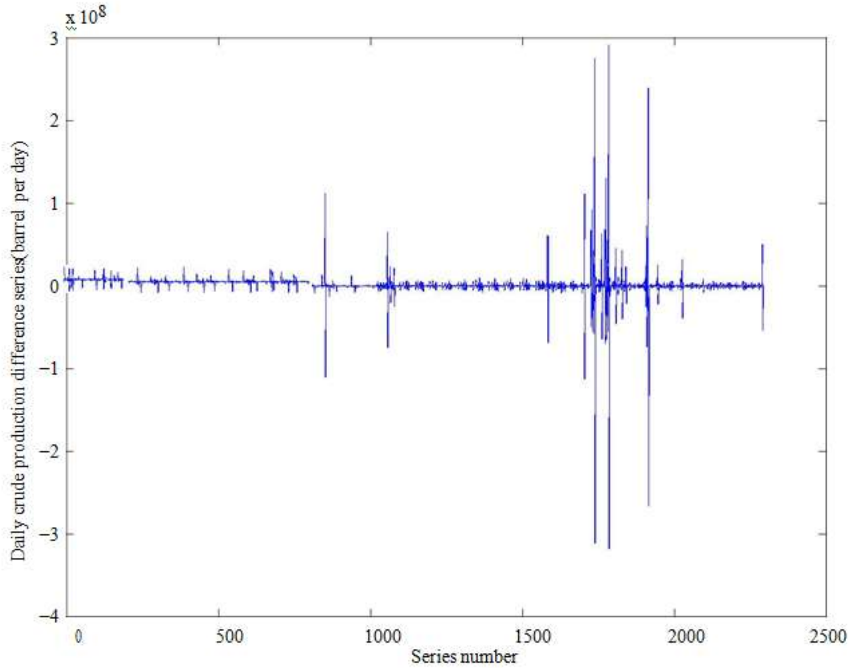

Figure 5: Daily Crude Oil Production Difference Series of the NNPC

Sample Autocorrelation Function

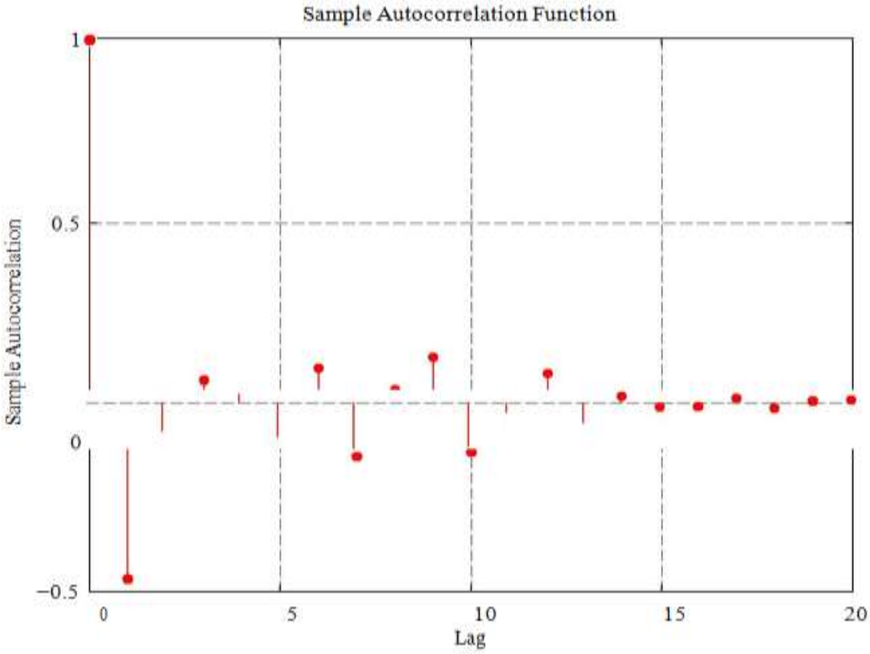

Figure 6: Autocorrelation Function of the Daily Crude Oil Production Difference Series of the NNPC

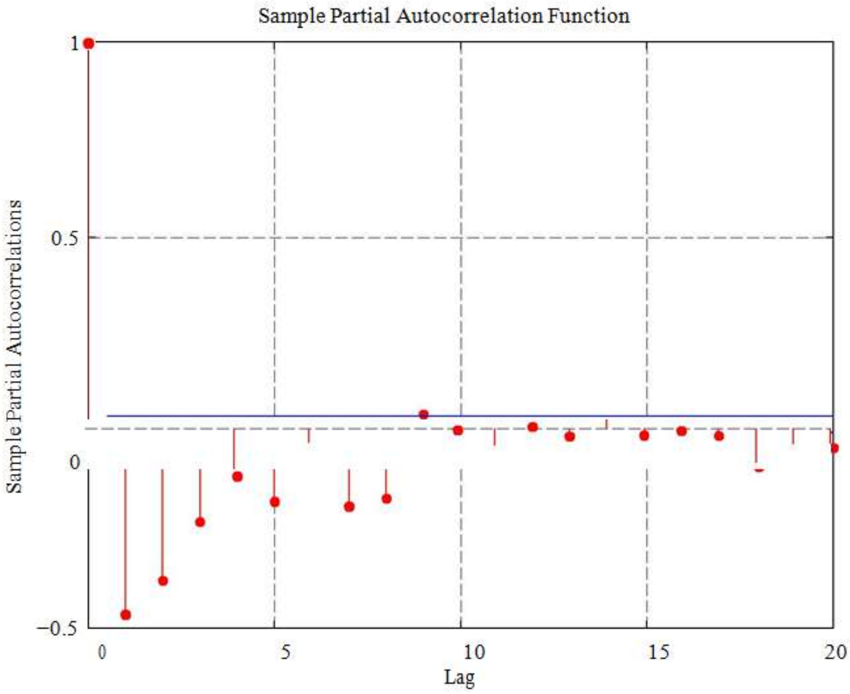

Figure 7: Partial Autocorrelation Function of the Daily Crude Oil Production Difference Series of the NNPC

The log difference series of the daily crude oil production is also illustrated pictorially in figure 8. Comparing figures 5 and 8 we notice that the log difference series of the daily crude oil production series of the NNPC has a smaller variance than the difference series of the daily crude oil production of the NNPC. The mean value of the difference series is 1.2790e+03, the median value is -11184 and the variance is 3.0162e+14. For the log difference of the NNPC series, the mean value is 0.0028, the median value is -0.0056 and the variance is 0.3889. The variances of the difference series and log difference series confirms our assertion in figures 5 and 8 that the difference series has a greater variance than the log difference series.

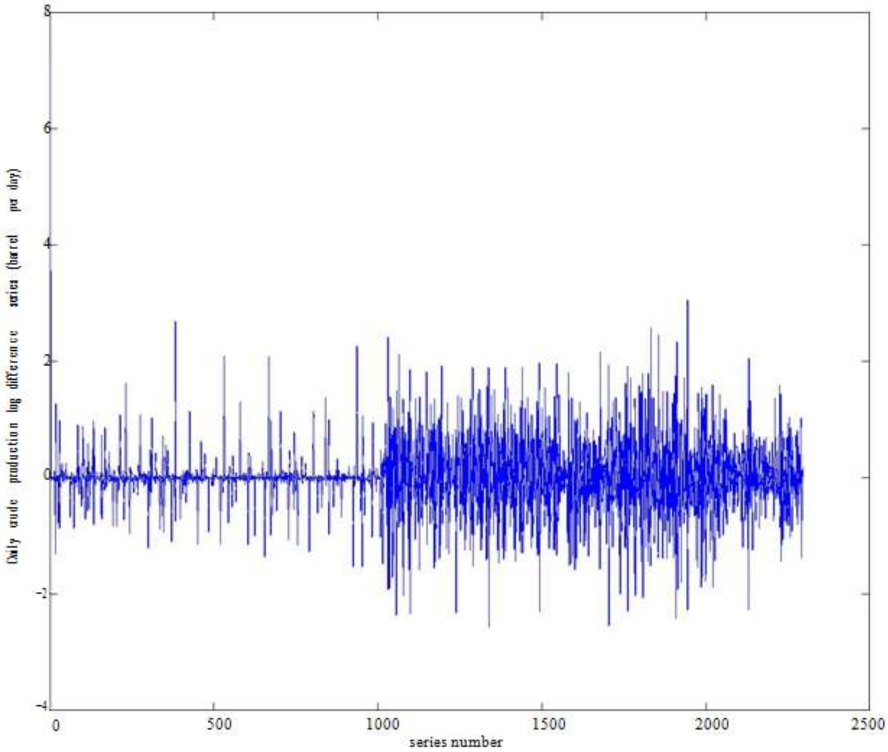

Figure 8: Daily Crude Oil Production Log Difference Series of the NNPC

Sample Autocorrelation Function 1

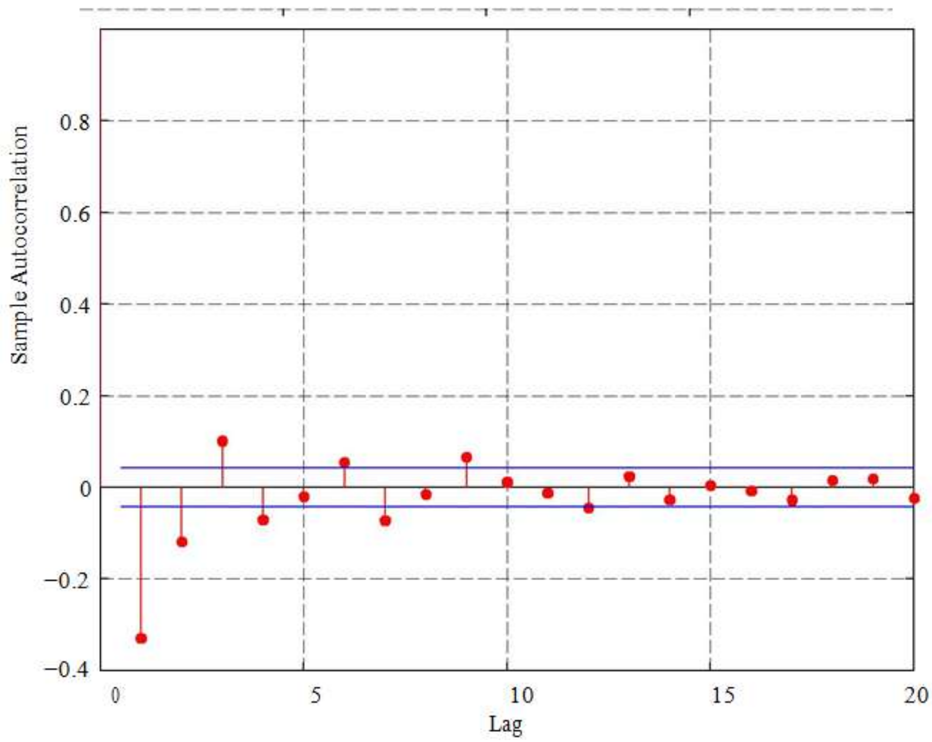

Figure 9: Autocorrelation Function of the Daily Crude Oil Production Log Difference Series of the NNPC

Sample Partial Autocorrelation Function 1

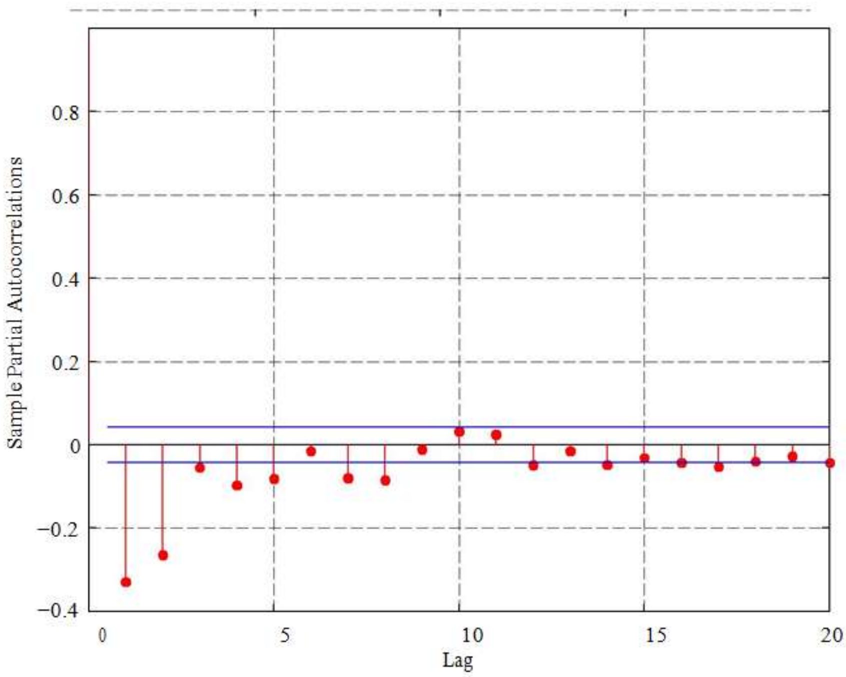

Figure 10: Partial Autocorrelation Function of the Daily Crude Oil Production Log Difference Series of the NNPC

Figure 8 shows the picture of the daily crude oil production log difference series of the NNPC. Figure 9 is the diagram of the sample autocorrelation function of the daily crude oil production log difference series of the NNPC and figure10 shows the diagram of the sample partial autocorrelation function of the daily crude oil production log difference series of the NNPC. Both figures points out that the daily crude oil production log difference series of the NNPC are both autocorrelated and partially autocorrelated.

An Augmented Dickey-Fuller (ADF) test of trend stationary is implemented by the MATLAB command ”kpsstest” on the difference and log difference series of the daily crude oil production data of the NNPC. The outcome points out that there is no statistical significant indication to accept the null hypothesis that unit roots occur for the difference series and log difference series for the daily crude oil production data of the NNPC. Hence, we conclude that the daily crude oil production difference series and the daily crude oil production log difference series are both stationary.

From the above analogy, one could deduce that the difference series and the log difference series for the daily crude oil production series of the NNPC will give better modeling and forecasting outcomes than the original daily crude oil production series of the NNPC. In view of this reason, we will use both the difference series and log difference series only for subsequent analysis, since it has demonstrated that its outcomes will produce better modeling and forecasting results than the original daily crude oil production series of the NNPC.

In this paper, we indicate that for the quadratic regression method, forecasting will hinge on two input variables, three input variables and four input variables. The procedure also applies to the two layered layer recurrent neural network (TLLRNN) method. For both quadratic regression and two layered layer recurrent neural network models, forecasts is made based on five varied sample measurement: 50 days data measurements, 100 days data measurements, 200 days data measurements, 400 days data measurements and 800 days data measurements. Furthermore, for both quadratic regression and two layered layer recurrent neural network models, our forecast is hinge on 1 day, 3 days and 5 days ahead predictions. This procedure will now result in the computation of 1 day, 2 days and 3 days ahead root mean square error (RMSE) and mean absolute error (MAE) for each model. This process determines the pattern in which RMSEs and MAEs are at variant from 1 day to 3 days predictions and from 3 days to 5 day predictions.

The fundamental function of the regression method is shown as equation (2) and the response function is shown as (3). Also, the normal equations for estimating the parameters in the regression model is shown as (??). In regression methods and neural network methods, appropriate input variables are fundamental to create credible models and predictions. The concept of stochastic decomposition is used here to decompose data series to get input variables for regression models and neural network models. We decomposed the difference series and log difference series of the NNPC series into two and three, as well as four independent series, which are then applied as input variables for the regression and neural network methods. As we have explained before the regression model used here is the quadratic regression model, while the neural network model used here is the two layered layer recurrent neural network.

Suppose a given set of observations are identically and independently distributed, the forecasts with the smallest mean square error in such observations are the best forecasts. This study evaluates forecasting performances by computing the root mean square error (RMSE) and mean absolute error for the difference series and log difference series of the daily crude oil production series of the NNPC.

Table 1: Difference series performance using quadratic regression model

Variable Name	RMSE		Variable name	MAE
50 days forecast v1 rmse	1.73674482475458e-08		50 days forecast v1 mae	1.84159500394969e-09
50 days forecast v3 rmse	33978164.4667824		50 days forecast v3 mae	6074665.17301324
50 days forecast v5 rmse	17560480.4098658		50 days forecast v5 mae	3433793.93657883
100days forecast v1 rmse	4.34609653126616e-09		100 days forecast v1 mae	1.19550250802951e-09
100days forecast v3 rmse	17756580.1859778		100 days forecast v3 mae	3493393.02284148
100days forecast v5 rmse	17756580.1856403		100 days forecast v5 mae	3493393.02284148
200days forecast v1 rmse	7.59859195247509e-09		200 days forecast v1 mae	1.53212636692569e-09
200days forecast v3 rmse	18138141.2077174		200 days forecast v3 mae	3572066.88347660
200days forecast v5 rmse	18138141.2077476		200 days forecast v5 mae	3572066.82425979
400days forecast v1 rmse	1.73674482475458e-08		400 days forecast v1 mae	1.84159500394969e-09
400days forecast v3 rmse	33978164.4667824		400 days forecast v3 mae	6074665.17301324
400days forecast v5 rmse	17560480.4098658		400 days forecast v5 mae	3433793.93657883
800days forecast v1 rmse	9.65870835286466e-09		800 days forecast v1 mae	2.83676968190029e-09
800days forecast v3 rmse	21399815.9870266		800 days forecast v3 mae	4577927.50803213
800days forecast v5 rmse	2.139981598711531e+07		800 days forecast v5 mae	4577927.43574297

The computational outcomes for root mean square error (RMSE) and mean absolute error for (MAE) for difference series of the daily crude oil production series of the NNPC is illustrated in table 1. It can be seen from the table that RMSEs and MAEs possessed better forecasting performances for 1 day prediction for 50, 100, 200 400 and 800 days data samples since their errors are very small. The MAE for 800 days data sample for 1 day prediction constitute better forecasting performance than the RMSE for 800 days data sample for 1 day prediction, since the error of the former is smaller than the error of the latter. The RMSEs and MAEs for 3 and 5 days predictions for 50, 100, 200, 400 and 800 days data sample are very large such they may not constitute good forecasting performance in view of their extremely large errors. Hence, for the difference series of the NNPC data, it is only the prediction of 1 day for all the data samples that can produce good forecasting performance as can be seen in table 1 for both RMSEs and MAEs.

For the log difference series of the NNPC series, we also considered log difference performances for four independent variable quadratic regression models. We used the concept of stochastic decompositions by Becker [14] to decompose the log difference series into two, three and four independent series. In this study we considered only the four independent series which is used as input variables for the quadratic regression model. Table 2 shows the RMSE and MAE forecasting performances for four input variables with 1 day, 3 days and 5 days predictions based on data sample length of 50, 100, 200, 400 and 800 days data samples respectively.

Table 2: log difference series performance using quadratic regression model

Variable Name	RMSE		Variable name	MAE
50 days forecast v1 rmse	0.700178629089724		50 days forecast v1 mae	0.555934449097355
50 days forecast v3 rmse	1.03116176568080		50 days forecast v3 mae	0.802793050841683
50 days forecast v5 rmse	0.905898723378657		50 days forecast v5 mae	0.693629758735138
100days forecast v1 rmse	0.700178629089724		100 days forecast v1 mae	0.555934449097355
100days forecast v3 rmse	1.03116176568080		100 days forecast v3 mae	0.802793050841683
100days forecast v5 rmse	0.905898723378657		100 days forecast v5 mae	0.693629758735138
200days forecast v1 rmse	0.634775169720679		200 days forecast v1 mae	0.516182205328637
200days forecast v3 rmse	0.918942871582563		200 days forecast v3 mae	0.726236904965330
200days forecast v5 rmse	0.780291436682125		200 days forecast v5 mae	0.616975456447820
400days forecast v1 rmse	0.647473770990421		400 days forecast v1 mae	0.521463449608953
400days forecast v3 rmse	0.940282341555137		400 days forecast v3 mae	0.736961346640355
400days forecast v5 rmse	0.794126063955765		400 days forecast v5 mae	0.623274280061115
800days forecast v1 rmse	0.668917570400212		800 days forecast v1 mae	0.527934559511973
800days forecast v3 rmse	1.023884192884413		800 days forecast v3 mae	0.796481075452469
800days forecast v5 rmse	0.861022230305001		800 days forecast v5 mae	0.656589744043977

In table 2 one could see that the MAEs outcomes possessed better forecasting performances than the RMSEs outcomes. This is because the MAEs values are smaller than the RMSEs values, thereby producing the optimal solution. Also, one could see that for the MAEs, predictions 1 day ahead, 3 days ahead and 5 days ahead based on a sample length of 50 and 100 days data sample, are equal and this trend also applies to the RMSEs.

Sequel to the foregoing statement, forecasting performance for 1 day ahead, 3 days ahead and 5 days ahead prediction is not affected by the sample length of the data in this situation. Hence, irrespective of whether the sample length of the data series is 50 days sample length or 100 days sample length, the forecasting performance remains the same. In comparing tables 1 and 2, it can be deduce from these tables that the log difference series, constitute better forecasting performance than the difference series of the NNPC series in the mean square error sense. Therefore, forecasting the daily crude production of the NNPC will produce better predictions for 1 day ahead, 3 days ahead and 5 days ahead based on a sample length of 50, 100, 200, 400 and 800 days data sample by using the log difference series of the NNPC data, rather than using the difference series of the NNPC data for 1 day ahead, 3 days ahead and 5 days ahead predictions, based on a sample length of 50, 100, 200, 400 and 800 days data sample.

Table 3: Difference series performance using two layered neural network model

Variable Name	RMSE		Variable name	MAE
50 days forecast v1 rmse	6.083876722734590		50 days forecast v1 mae	6.04192287906014
50 days forecast v3 rmse	6.084225084810980		50 days forecast v3 mae	6.04195986808638
50 days forecast v5 rmse	6.084388805428992		50 days forecast v5 mae	6.04214205409435
100days forecast v1 rmse	6.08346463571485		100 days forecast v1 mae	6.04111316157741
100days forecast v3 rmse	6.08356402153901		100 days forecast v3 mae	6.04091333694678
100days forecast v5 rmse	6.08364221194899		100 days forecast v5 mae	6.04099813697495
200days forecast v1 rmse	6.06368238779343		200 days forecast v1 mae	6.02069620950992
200days forecast v3 rmse	6.06404100020748		200 days forecast v3 mae	6.02074638620402
200days forecast v5 rmse	6.06421917455469		200 days forecast v5 mae	6.02092365138931
400days forecast v1 rmse	6.04147337002840		400 days forecast v1 mae	5.99783319665977
400days forecast v3 rmse	6.04190942302792		400 days forecast v3 mae	5.997934031498830
400days forecast v5 rmse	6.04174721105203		400 days forecast v5 mae	5.997754395562628
800days forecast v1 rmse	5.96390076943475		800 days forecast v1 mae	5.91875035452281
800days forecast v3 rmse	5.96446915704074		800 days forecast v3 mae	5.91890893826510
800days forecast v5 rmse	5.96427140967527		800 days forecast v5 mae	5.91867696454829

From table 3 one could see that the MAEs outcomes possessed better forecasting performances than the RMSEs outcomes. This is because the MAEs values are smaller than the RMSEs values, thereby producing the optimal solution. Although, a critical scrutiny of outcomes in table 3 shows that the outcomes are almost equal for the RMSEs and MAEs. For the MAEs, the outcomes for it in table 3 revealed that, predictions for 1 day ahead, 3 days ahead and 5 days ahead for 50 days sample length for the NNPC series are equivalent. This indicates that irrespective of predictions of 1, 3 and 5 days ahead, forecasting performance for the outcomes will remain the same or almost the same. Forecasting performance with regards to RMSEs for predictions 1 day ahead, 3 days ahead and 5 days ahead for 50 days sample length follows the same trend with the MAEs in this category. Table 3 also indicates that forecasting performance for 1 day, 3 days and 5 days ahead for 100 and 200 days sample length of the NNPC series are also similar, while forecasting performance for 1, 3 and 5 days ahead predictions for 400 and 800 sample length are similar. The similarities indicates that the number of days in the sample length of the data, does not adversely change the forecasting performances in forecasting the daily crude oil production of the NNPC using difference series, modeled by a two layered layer recurrent network. Table 3 also indicates that a similar analogy with the MAEs follows for the RMSEs.

Table 4: Log difference series performance using two layered neural network model

Variable Name	RMSE		Variable name	MAE
50 days forecast v1 rmse	6.08561017316472		50 days forecast v1 mae	6.04372480674281
50 days forecast v3 rmse	6.08591201187344		50 days forecast v3 mae	6.04391877340830
50 days forecast v5 rmse	6.08547424613058		50 days forecast v5 mae	6.04347184842881
100days forecast v1 rmse	6.08519394424764		100 days forecast v1 mae	6.04326763140537
100days forecast v3 rmse	6.08533086093488		100 days forecast v3 mae	6.04333624546782
100days forecast v5 rmse	6.08567949902457		100 days forecast v5 mae	6.04364116435706
200days forecast v1 rmse	6.06322832199181		200 days forecast v1 mae	6.02046328876359
200days forecast v3 rmse	6.063539242147152		200 days forecast v3 mae	6.02068801520373
200days forecast v5 rmse	6.063438837711934		200 days forecast v5 mae	6.02060682387445
400days forecast v1 rmse	6.04193867427260		400 days forecast v1 mae	5.99853725334160
400days forecast v3 rmse	6.04229216135372		400 days forecast v3 mae	5.99883410210333

400days forecast v5 rmse	6.04237594884981	400 days forecast v5 mae	5.99889006905902
800days forecast v1 rmse	5.96331306765217	800 days forecast v1 mae	5.91831585265713
800days forecast v3 rmse	5.96351821053871	800 days forecast v3 mae	5.91836902518220
800days forecast v5 rmse	5.96351499317985	800 days forecast v5 mae	5.91835653845694

Table 4 indicates that forecasting performance of the MAEs for log difference series of the NNPC data modeled by a two layerd layer recurrent neural network are slightly better than the RMSEs. This is because outcomes resulting from the MAEs are a little bit smaller than outcomes resulting from the RMSEs. The outcomes of the MAEs for predictions of 1 day ahead, 3 days ahead and 5 days ahead for 50 days, 100 days and 200 days sample length are almost identical. This shows that forecasting performance of the log difference series of the NNPC series modeled by a two layered layer recurrent neural network is not adversely affected by the different sample length of 50 days to 200 days, since the results are almost identical in this category. Table 4 also indicates that forecasting performance of the MAEs for 1 day ahead, 3 days ahead and 5 days ahead predictions for 400 days sample length and 800 days sample length are almost identical. This shows that forecasting performance of the log difference series of the NNPC series modeled by a two layered layer recurrent neural network for 400 days sample length and 800 days sample length are similar and possessed better forecasting performance than for 50 to 200 days sample length.

Table 4 also shows that forecasting performance resulting from analysis of the RMSEs as modeled by a two layered layer recurrent neural network revealed that predictions of 1 day ahead, 3 days ahead and 5 days ahead for 50 days sample length, 100 days sample length, 200 days sample length and 400 days sample are characterize by sharp similarities. This shows that forecasting performance for 1 day ahead prediction, 3 days ahead prediction and 5 days ahead prediction for 50 days to 400 days sample length are equivalent in the root mean square error sense. Table 4 also shows that the RMSEs for 1 day ahead, 3 days ahead and 5 days ahead predictions for 800 days sample length indicates better forecasting performance than for 50 days to 400 days sample length in the root means square error sense as modeled by a two layered layer recurrent neural network using the log difference series of the NNPC data.

On the comparison of tables 3 and 4, one could deduce that forecasting performances of the log difference series of the NNPC data as modeled by a two layered layer recurrent neural network are better than forecasting performances of the difference series of the NNPC series as modeled by a two layered layer recurrent neural network. Also, comparing tables 1 and 3 shows that forecasting performance of the difference series of the NNPC data as modeled by a two layered recurrent neural network are better than forecasting performance of the difference series of the NNPC data as modeled by a quadratic regression modeled in the mean square error sense. This is because the outcomes of table 1 are larger than the outcomes of table 3. The outcomes of tables 2 and 4 reveals that forecasting performance of the log difference series of the NNPC data as modeled by a two layered layer recurrent neural network are better than forecasting performance of the log difference series of the NNPC data as modeled by a quadratic regression model. This is because the outcomes of table 2 are larger than the outcomes of table 4.

On the comparison of tables 1 and 2 which were modeled by a quadratic regression model for both difference as well as log difference series of the NNPC data, with tables 3 and 4 which were modeled by a two layered layer recurrent neural network for both difference and log difference series of the NNPC series, indicates that for both difference and log difference series, forecasting performance resulting from the series modeled by a two layered neural network are better than forecasting performance of the series modeled by a quadratic regression model in the mean square error sense. This resulting analogy shows that a two layered layer recurrent neural network gives better forecasting performance than a quadratic regression model for both difference and log difference series of the NNPC data in the mean square error sense.

IV. Conclusions

This study have shown that a two layered layer recurrent neural network model have better forecasting performances greater than a quadratic regression model with regards to the difference series and log difference series of the daily crude oil production series of the NNPC. We could deduce from the study that root mean square errors (RMSEs) and mean absolute errors (MAEs) are almost identical for short data lengths such as 50 days to 200 days data samples as well as with longer data lengths such as 400 days to 800 days data sample length. Speaking firmly on the autocorrelation pattern of the difference series and log difference series of the daily crude oil production of the NNPC series, one could deduce that the daily crude oil production series of the NNPC series indicates a significant autocorrelation and partial autocorrelation. However, we might ruminate that forecasting performance of the daily crude oil production series of the NNPC are at its best with log difference series, irrespective of the forecasting method. There have been factors affecting the daily crude oil production of the NNPC such as the Niger Delta militancy unrest which has led to the vandalism of oil pipelines, kidnapping of oil workers, oil theft as well as depriving oil workers access to the oil fields where they work. Future investigation on the daily crude oil production series of the NNPC might consider nonlinear mathematical models of the neural network category with hyperbolic activation functions as well as sine and cosine activation functions to eliminate the irregular factors that affects the daily crude oil production series of the NNPC, such that, a proper spring board will be established upon which models will be developed for forecast.

Acknowledgement

This research is partly financed by Fundamental Research Grant Scheme (203/PMATHS/6711368) from the Ministry of Higher Education, Malaysia and Universiti Sains Malaysia.

References

[1] M. Baghebo, T. O. Atima, The impact of petroleum on economic growth in Nigeria, Global Business and Economics Research Journal 2 (2013) 102–115.
[2] J. Neter, M. H. Kutner, C. J. Nachtsheim, W. Wasserman, Applied linear statistical models, volume 4, Irwin Chicago, 1996.
[3] R. W. Farebrother, Non-linear curve fitting and the true method of least squares, Journal of the Royal Statistical Society: Series D (The Statistician) 47 (1998) 137–147.
[4] R. Tobia, An introduction to least squares regression, SAS Institution Inc (1999).
[5] L. M. Healy, Logistic regression: An overview, Eastern Michigan University, College of Technology. In Proceeding of COT 7 (2006) 30–37.
[6] J. L. Elman, Finding structure in time, Cognitive science 14 (1990) 179–211.
[7] A. C. Tsoi, A. Back, Discrete time recurrent neural network architectures: A unifying review, Neurocomputing 15 (1997) 183–223.
[8] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proceedings of the national academy of sciences 79 (1982) 2554–2558.
[9] M. I. Jordan, Attractor dynamics and parallellism in a connectionist sequential machine (1986).
[10] A. F. Atiya, A. G. Parlos, New results on recurrent network training: unifying the algorithms and accelerating convergence, Neural Networks, IEEE Transactions on 11 (2000) 697–709.
[11] Z. Zhang, Z. Tang, C. Vairappan, A novel learning method for elman neural network using local search, Neural Information Processing–Letters and Reviews 11 (2007) 181–188.
[12] Q. Liu, Z. Guo, J. Wang, A one-layer recurrent neural network for constrained pseudoconvex optimization and its application for dynamic portfolio optimization, Neural Networks 26 (2012) 99–109.
[13] I. O. Aksu, R. Coban, Identification of disk drive systems using the multifeedback-layer neural network and the particle swarm optimization algorithm, in: Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2013 International Conference on, IEEE, pp. 230–234.
[14] A. B. D. Becker, Decomposition methods for large scale stochastic and robust optimization problems, Ph.D. thesis, Massachusetts Institute of Technology, 2011.