**Periodicals of Engineering and Natural Sciences** ISSN 2303-4521
Vol. 8, No. 1, March 2020, pp.438-446

# Forecasting Iraqi oil production using artificial neural networks

**Suhail Najm Abdullah**

Department of Statistics, University of Baghdad

**ABSTRACT**

Artificial neural networks (ANNs) are bendy computational structures and yet ordinary approximations as the execute stand applied to day selection predicting especially correct problems. The paper aims to address the effect on the outputs of Artificial Neural Network of the number of input layer nodes, and to minimize error. Time series views included the amount of Iraqi oil production for the 2011–2019 period. The most important finding of this paper is the greater the number of artificial neural network inputs, the smaller the error, and the improved performance results, which reflect positively on the prediction results.

**Keywords**: Neural networks , Back Propagation , forecasting, time series, Minimize error

***Corresponding Author:***

Suhail Najm Abdullah,

Department of Statistics,

University of Baghdad,

Baghdad, Iraq

E-mail: dr.suhail.najm@coadec.uobaghdad.edu.iq

## 1. Introduction

Predicting of time series is a vital forecasting area where past observations about the identical moving are accrued or analyzed into discipline after effect a model so much describes the underlying relation. Afterwards the mannequin is back in conformity with foresee the day series between the futures. This modeling method is mainly beneficial then small advantage is on hand touching the underlying statistics era procedure and now even is no first-class rationalization mannequin related in accordance with other explanatory variables within the reckoning variable longevity.
In recent years, a lot of work had been done on the application of (AIT) Artificial Intelligence Technique to forecasting Short Term use such as hourly or daily load forecasting and Medium-Term forecasting due to the (ANNs) Ability to learn and create construct complex Non-linear mapping through a series of input and output scenarios. Nevertheless, the overall style of Long-Term energy consumption forecasting emphasizes aggregate energy combustion forecasting up to now, and both time series models and regression based had been used in the cycle of accumulation due to time smoothing. ANNs had been used recently in time series and in Regression based models.
Artificial Neural Networks are among the predictive models that are most accurate and widely used that enjoy fruitful financial, economic, technological, foreign exchange, stock issues, etc. applications. Prediction. Specific distinctive characteristics of (ANNs) make them useful and desirable for a predictive purpose. First, unlike conventional model-based approaches, (ANNs) are data driven; self-adaptive approaches since there are few assumptions a priori about the structures of the problems being studied. Second, neural artificial networks may become widespread. After learning the data (a sample) given to them. Recent studies have shown how Artificial Neural Networks identify and predict these. A neural network was demonstrated to be able to approximate any

continuous function. Neural Network were effectively used for the forecasting of the financial data set. The classical methods used for predicting time series like Box-Jenkins, ARMA or ARIMA assume a linear relationship between input and outputs. Neural networks have the advantage that no priority details can be approximated to some nonlinear functions about the properties of the data set.
In This paper, the aim checking for the accuracy of the neural networks to identify the right one to solve the time series. It was the well-known application of real data collection from 2011 to 2019 which is the energy data for Iraq oil production. In this paper, predictive accuracy was used to determine the best for forecasting neural networks.

## 2. Artificial neural networks

ANNs are computer-based techniques, processing data at the same time. They are trained to learn and use this learning to identify the existence or absence of similar trends in other data. Additionally, they do when the data is incomplete or noisy. ANNs consist of many closely related, simple, or neuronal processing components. In ANNs processing elements are organized into layers; each layer has many processing elements and different functions. The processing elements in each layer are also associated with the processing elements in the preceding and subsequent layers and sometimes even with the processing elements in the same layers[1], [2].
(ANNs) are essentially Non-linear. The conventional a approaches to prediction of the time series such as the exponential smoothing or B.J, assume the time series is being studied are supposed from linear processes[3].However if the underlying mechanism is Non-linear they can be unacceptable. Actually, the systems of the real world are often nonlinear. ANN configuration can be seen in the figure below[4], [5].

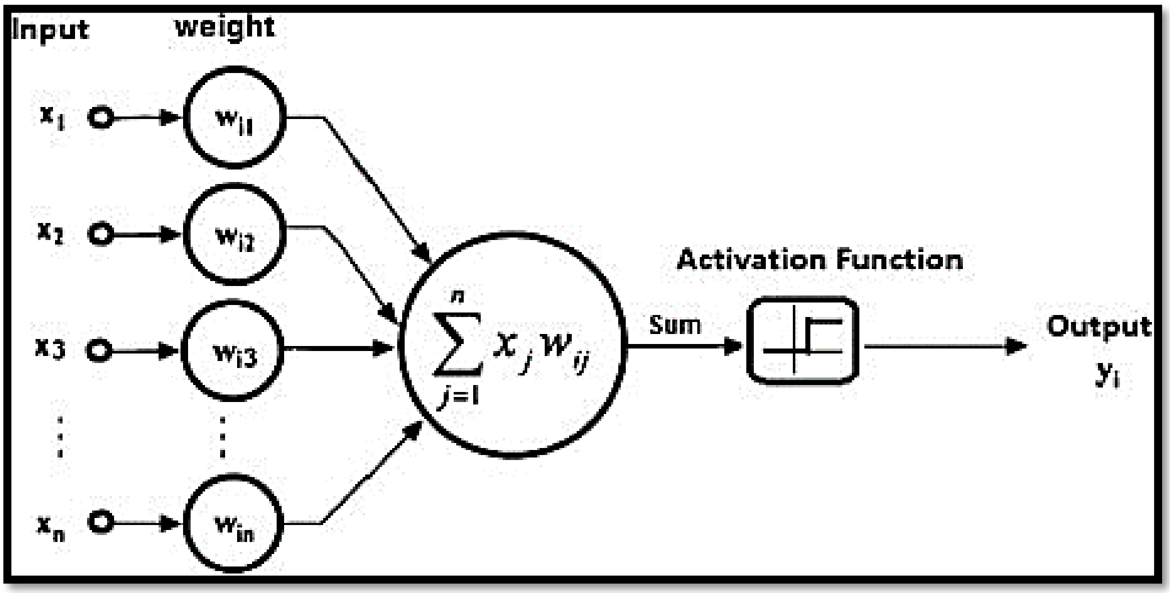


Figure 1. Neural network structure

## 3. Back propagation network (BPN)

This style of ANN must had a minimum propagation from the back. It can contain one or more layers at the same time, returning the output of each neuron as input to all remaining neurons. BPN also occurs which means the output of neuron it becomes input for the same neuron. This form of network was usually not used in affective domains. For BPN it is simpler to achieve the same objectives[6], [7].
The major steps of BPM Methodology are to computation error in the layer of output, and then change the weighting of the hidden output layer. Likewise, the error is measured for the input-hidden layer to correct the weighting. Before performance Network, using new weighting was calculated, and the same error measurement and weight updating process steps were repeated until minimum error for network was achieved. Figure 2 shows BPN structure based on [8]–[10].

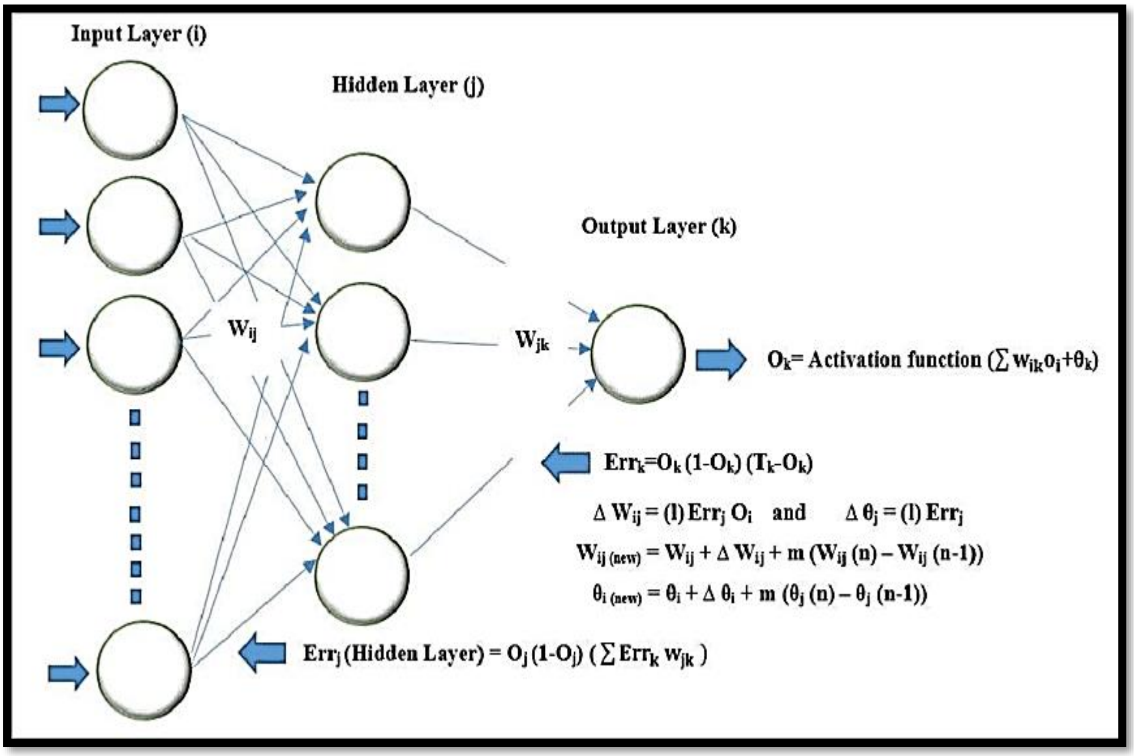


Figure 2. BPN structure

### 4. Time series then ANNs

ANN are accomplishments, which increase appreciation within the technology and then inspect time series, which fit in the imitation of their potential in accordance with training and self- learning. theorems more researchers use abased on the transition or deceleration of the epoch sequence by some dosage and more, after choosing the entry era series of ANN in order to solve the era series.[11]–[13] .

$$Z_t = w_0 + \sum_{j=1}^{q} w_j \,.\, g\left(w_{0j} + \sum_{i=1}^{p} w_{ij}.Z_{t-1}\right) + \varepsilon_t \; \dots.\dots\dots (1)$$

$$Z_t = w_0 + \sum_{j=1}^{q} w_j \,.\, g\left(w_{0j} + \sum_{i=1}^{p} w_{ij}.Z_{t-1}, Z_{t-2}\right) + \varepsilon_t \; \dots.\dots\dots (2)$$

### 5. Evaluating forecast accuracy

For comparison purposes, the results of the methods adopted, can adopt the following statistical criteria[10]–[13] .

$$MSE = \frac{\sum {a_t}^2}{n} \dots.(3)$$

$$RMSE = \sqrt[2]{MSE} \;\dots\dots(4)$$

$$MAPE = \frac{\sum \left|\frac{a_t}{Z_t}\right|}{n} * 100 \dots..(5)$$

Where: $Z_t$ $x_t$ is actual value, $\hat{x}_t\tilde{\hat{x}}_t$ $\hat{Z}_t$ is forecast value, n: is Sample size, $a_t$ is error and.

$a_t = Z_t - \hat{Z}_t$

Coefficient of determination ($R^2$)

$$R^2 = \frac{\sum Z_t - \bar{Z}}{\sum \bar{Z} - Z_t} \quad \dots \dots . . (6)$$

### 6. Practical part and results

In this section, well-known real data sets for energy data on Iraq's oil production from 2011 to 2019 are shown in Table 1. This is to achieve the best model using modern methods, where oil is the cornerstone of the national economy in Iraq and the Near East. In addition, the main and sole source of Budget income for the country. In addition, the MATLAB high level programming language version (8.2) and SPSS programming (V.21) are used to the implement.

Table1. Production of Iraqi oil

| year | | Quantity (barrel) | | |
|---|---|---|---|---|
| | | 2011 | 2012 | 2013 |
| Month | January | 67000000 | 65300000 | 73100000 |
| | February | 61600000 | 58400000 | 71000000 |
| | March | 66900000 | 71800000 | 74900000 |
| | April | 64200000 | 75200000 | 78700000 |
| | May | 69000000 | 76000000 | 77000000 |
| | June | 68200000 | 72100000 | 69800000 |
| | July | 67200000 | 78000000 | 72000000 |
| | August | 67800000 | 79500000 | 79900000 |
| | September | 63100000 | 77800000 | 62100000 |
| | October | 64800000 | 81300000 | 69800000 |
| | November | 64100000 | 78600000 | 71400000 |
| | December | 66500000 | 72800000 | 72600000 |
| year | | Quantity (barrel) | | |
| | | 2014 | 2015 | 2016 |
| Month | January | 69064000 | 78600000 | 101840000 |
| | February | 78363000 | 72724000 | 93537000 |
| | March | 74316000 | 92423000 | 101865000 |
| | April | 75279000 | 92308000 | 100916000 |
| | May | 80036000 | 97504000 | 99205000 |
| | June | 72786000 | 95612000 | 95261000 |
| | July | 75710000 | 96246000 | 99266000 |
| | August | 73639000 | 95447000 | 100130000 |
| | September | 76249000 | 91553000 | 98285000 |
| | October | 76282000 | 83757000 | 104915000 |
| | November | 75309000 | 100946000 | 104116000 |
| | December | 91141000 | 99659000 | 109107000 |
| year | | Quantity (barrel) | | |

| | | 2017 | 2018 | 2019 |
|---|---|---|---|---|
| Month | January | 102939163 | 108190068 | 113111429 |
| | February | 91558868 | 95940404 | 101387559 |
| | March | 92400000 | 107500000 | 104696894 |
| | April | 97574813 | 100197197 | 103988607 |
| | May | 101100000 | 108175920 | 110737293 |
| | June | 98178000 | 105640160 | 105603325 |
| | July | 100000000 | 109857705 | 110548767 |
| | August | 99700761 | 111088580 | 111706135 |
| | September | 97204267 | 106787289 | 107276327 |
| | October | 103730640 | 107821261 | 106859982 |
| | November | 105500000 | 101313958 | 105014748 |
| | December | 109573817 | 115517974 | 106265347 |

The results for the time series under Search from the BPN are based on MATLAB simulator as follows:

**a. Model 1**

Layer of Input: the nodes number in this layer is one node represented by the variables ($Z_{t-1}$), with a time series lag of one degree. Layer of hidden: The maximum number of nodes in this network is 15 nodes, with one network (after several trials). Layer of output: This layer involves of only one node, by the $Z_t$ vector. For the time series under review, the performer's results for the BP network are shown in Figure 3-4. Table 2 displays the effects of the error test for comparing the two approaches employed.

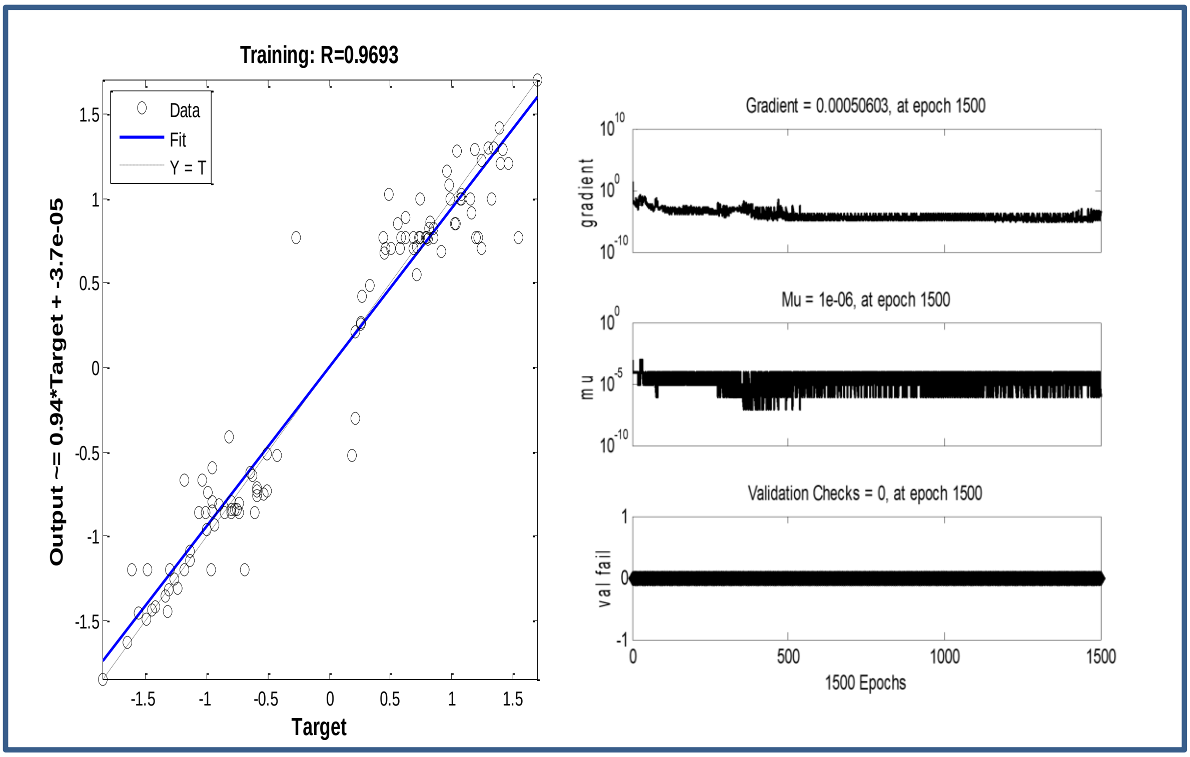


Figure 3. Evaluation of the artificial neural network of model a

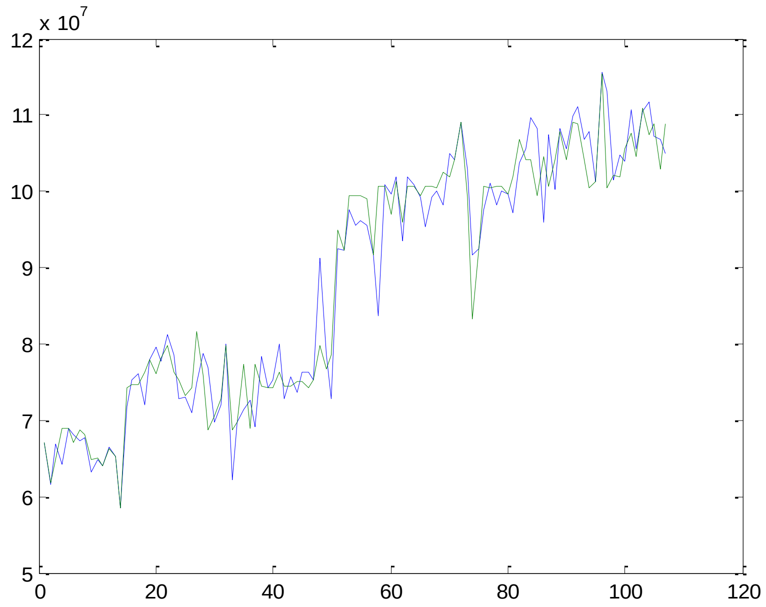


Figure 4. Curve fitting of model a

**b. Model 2**

Layer of Input: the nodes number in this layer is two, with a time series lag of two degrees, indicated by the variables ($Z_{t-1}$, $Z_{t-2}$). Layer of hidden: The maximum number of nodes in this network is 15 nodes with one network (after several trials). Layer of output: This layer involves of only one node, by the $Z_t$ vector. For the time series under review the performer's results for the BP network at show in Figure 5-6. Table 2 displays the effects of the error test for comparing the two approaches employed.

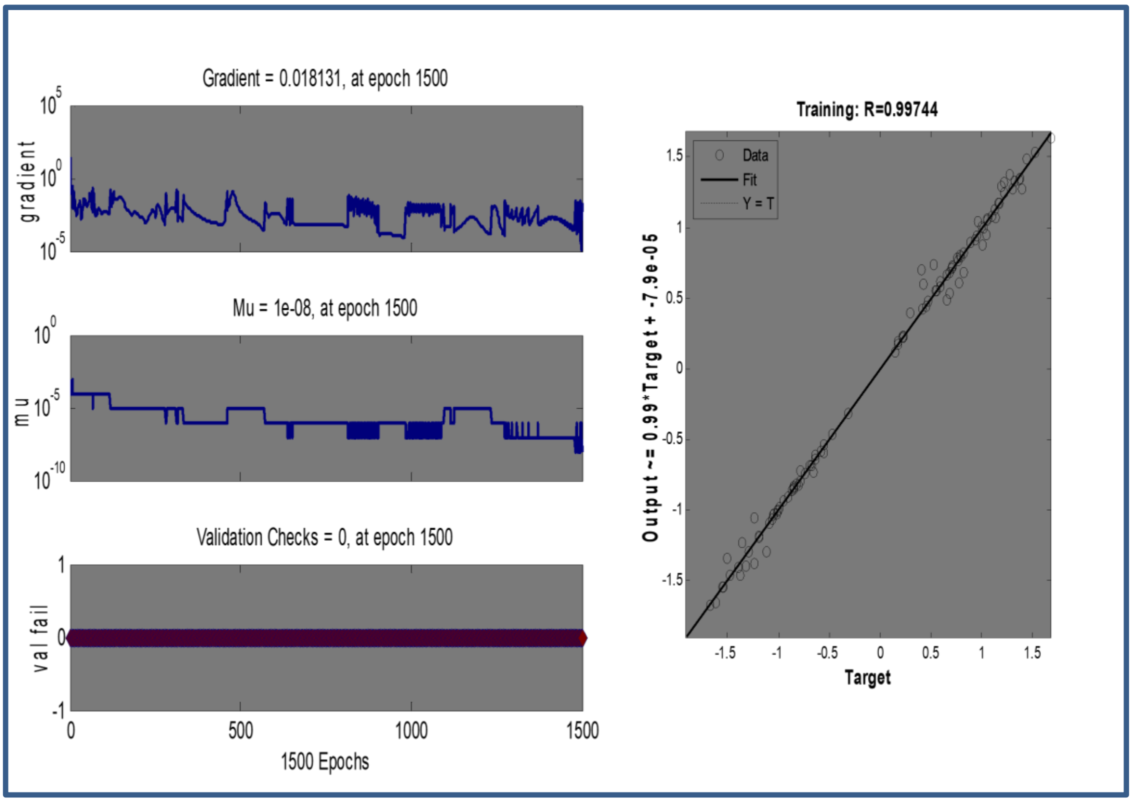


Figure 5. Evaluation of the artificial neural network of model b

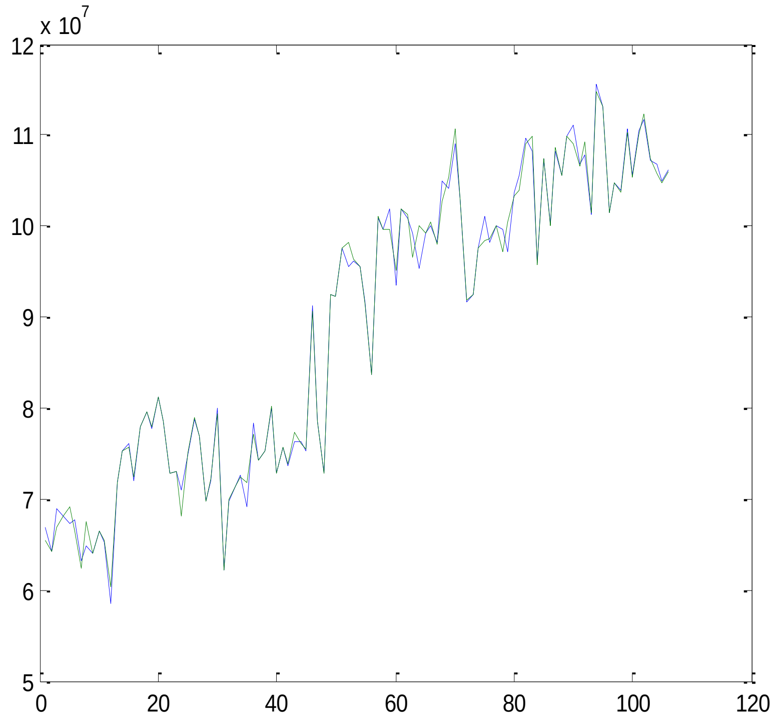


Figure 6. Curve fitting of model b

Table 2. Error criteria results

| Model 2 | Model 1 | criteria |
|---|---|---|
| 1134800 | 3942800 | RMSE |
| 0.7310 | 2.9815 | MAPE |
| 0.997 | 0.969 | $R^2$ |

## 7. Decision and conclusions

Predicting time series is one of the most important quantitative models in the literature to have gained considerable attention. Artificial neural network (ANNs) have proven to be an important, general-purpose approach for high-precision pattern recognition, classification, clustering and particularly time series prediction. For evaluation purpose, three error indicators RMSE, MAPE and $R^2$ had used the compare predicted and actual values to validate each forecasting model.

The results of the methods used here are shown in Table 2 and Figure 8. The best approach or curve fitting for time series of Iraq oil output is the artificial neural network when the number of two inputs is better than the number of one input, the approach that gives minimized error possible and the highest coefficient of determination. Which indicates the effect of the inputs on the accuracy of the error and the improvement of the results and the reduction of the error, as the percentage of improvement in error reached 15% and the figure shows the histogram of the MAPE. The results also proved that the difference between the two models was significantly in error criteria. Artificial neural networks are working to improve error significantly and this is the ambition of all researchers. The histogram of the MAPE for two methods is shown in Figure 7. In addition, Table 3 shows the predictable oil production details for the coming period using Model 2.

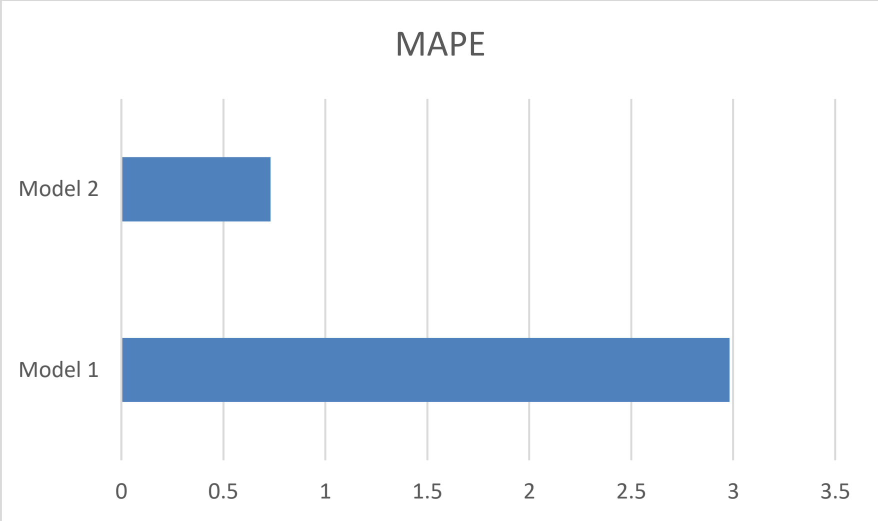


Figure 7. Histogram of MAPE for two models

Table 3. Forecasting oil production

| Forecast | Lower | Upper |
|---|---|---|
| 107833230 | 99041626 | 116624834 |
| 108230981 | 98907349 | 117554613 |
| 108628732 | 98801834 | 118455630 |
| 109026483 | 98720867 | 119332099 |
| 109424234 | 98661170 | 120187297 |
| 109821985 | 98620139 | 121023830 |
| 110219736 | 98595659 | 121843812 |
| 110617487 | 98585988 | 122648985 |
| 111015237 | 98589668 | 123440807 |
| 111412988 | 98605468 | 124220509 |
| 111810739 | 98632333 | 124989145 |
| 112208490 | 98669355 | 574712626 |

**References**


[1] Z. Tong,C.De ALmeida, and P.A. Fishwick, "Time series forcasting using neural network vs. Box-Jenkins methodology ", *Simulation* , vol.57,no.5,pp.303-310,1991

[2] M. Khashei and M.Bijari, " An artificail neural network (p,d,q) model for time series forcosting ", *Expert Syst. Appl*.,vol.37,no.1,pp.279-489,2010.

[3] R. A. Abbas, A. Jamal, M. A. H. Ashour, and S.L. Fong, "Curve fiting prediction with artificail neural network : Acomparative analysis", *Peripd Eng. Nat. Sci*., vol. 8, no. 1,PP. 125-132,2020

[4] T.Nguyen, T.Nguyen, B. M. Nguyen, and Nguyen, "Efficient time series forcasting using neural network and opposition -based coral reefs optimization",*Int. J. Comput. Intell. Syst*., vol. 12, no. 2, pp. 1144-1161,2019.

[5] Y. Chen, B.Yang, J. Dong, and Abraham, " Time series forcasting using flexible neural tree model", *inf. sci. (Ny)*, vol. 174, no. 3-4, pp.219-235,2005

[6] B.Oancea and Stefan Cristian Ciucu, "Time series forcasting using neural networks", arXiv peper. arXiv1401.1333,2014.

[7] R. A. A. Helmi, A. Jamal, and M. A. H. Ashour, "Identifying High impact Disaster Triggers on Computer Networks in Asia pacific Region", *Int J. Eng. Tehnol* , vol. 7,no 4.11,pp.95-98, 2018.

[8] M. Abdul, H. Ashour, I. A. H. Al-dahhan, and A. K. Hassan, "Forcasting by using the optimal time series method", vol. 2. Springer Internatinal Publishing,2020

[9] M. A. H. Ashour and R. A. Abbas, "Improving time series forcast errors by ussing recurrent neural network",in proceeding of the 2018 7th international conference on software and computer application 2018,pp. 229-232.
[10] R. P. F. Guine, S. Mators, F.J. Goncalves, D. Costa, and M. Mendes,"Evaluation of phenolic compounds and antioxidant activity of blueberries and modelization by artificial neural network ", *Int. J. fruit Sci.* pp.1-16,2018.
[11] S. J. C. Pateerson, "A Comparison Between 8 Common Cost Forcasting Methods",2018
[12] M. A. H. Ashour, A. Jamal, and R. A. A. Helmi, "Effectiveness of Artfical Neural Network in Solving Financail Time Series", Int Eng.Techol., vol. 7,no. 4.11,pp.99-105, 2018.
[13] E. Mathioulakis, G. panaras and V. Belessiotis," Artigicial newral network for the performance perediction of heat pump hot water heaters", Int.J. Sustain Energy, vol. 37, no. 2,pp. 173-19, 2018