

Forecasting Natural Gas Prices Using Nonlinear Autoregressive Neural Network

Abdelkader Sahed

Department of Economics, The University Center in Maghnia, Algeria
E-mail: sahed14@yahoo.fr

Mohammed Mekidiche, Hacem Kahoui

Department of Economics, The University Center in Maghnia, Algeria

Received: 22 June 2020; Accepted: 15 July 2020; Published: 08 October 2020

Abstract: When forecasting time series, It was found that simple linear time series models usually leave facets of economic and financial unknown in the forecasting time series due to linearity behavior, which remains the focus of empirical and applied study. The study suggested the Nonlinear Autoregressive Neural Network model and a comparison was made using the ARIMA model for forecasting natural gas prices, as obtained from the analysis, NAR models were better than the completed ARIMA model, measured against three performance indicators. The decision criterion for the selection of the best suited model depends on MSE, RMSE and R2. From the results of the criterion it has found that both the models are providing almost closed results but NAR is the best suited model for the forecasting of natural gas prices.

Index Terms: Forecasting, ARIMA, NAR, Natural Gas Price.

1. Introduction

1.1. Background Information

Natural gas plays a very important role in economic life, it has increased in recent years the use of natural gas, which continues to grow rapidly for several decades [1].

In general, natural gas has become a clean and effective fossil fuel, which makes it an energy source. The percentage of natural gas demand is high in many countries [2]. In 2017 on the global natural gas demand has risen annually by about 3.3%.

Because of the uneven distribution of natural gas supplier in the world. This shows the need to transport gas via pipelines, which are safer and less expensive over long distances [3].

The production of natural gas, sold by Algeria in 2018, is estimated at 97 billion cubic metres, making it the world's tenth largest producer, according to the managing director of the Sonatrach group, reported the daily El khabar.

According to information from the company "Sonatrach", Algeria is currently the second supplier of natural gas to Europe after Russia, Algeria exported more than 20 billion cubic meters of natural gas, while exports amounted to Europe 16 billion cubic meters, and on the other hand Algeria produced about 130 Billion cubic meters of gas during 2019. More than 51 billion cubic meters were exported abroad.

In addition, Algeria is linked to Europe through three pipes that cross the Mediterranean, the first passes through Tunisia to the Italian island of Sicily, the second through the Moroccan territories to Spain, and the third through Almeria in southern Spain.

The importance of predicting natural gas prices in the short term is to work to balance the supply and demand naturally and reduce the risk of high fuel prices. Where the demand for natural gas in the winter is greater than the quantity supplied, and thus the underground natural gas reserves play an effective role in meeting the growing demand. In addition, natural gas has become the preferred energy source because it is cleaner and cheaper than most other gases. Moreover, the understanding of the dynamics of change in natural gas prices will help decision-makers to plan energy [4].

The objective of this work is to identify the ARIMA models, identify NAR models and determine the optimal model for forecasting natural gas price.

1.2. Statement of the problem

In order to explain the above gaps, the purpose of this study is to address the following main research question:

What is the ideal model among ARIMA models and NAR models for forecasting natural gas prices?

To answer this question, This research paper is divided as follows: The second section provides previous studies in this field, and the third section provides the tools and methods used and is represented in both the arima method and the NAR method, the fourth section explains the results of forecasting natural gas prices and comparison between these methods.

2. Literature review

Previous studies include many studies on the subject of forecasting natural gas prices using a non-linear autonomic nervous network (ANN), in this section of the study the most important of which will be presented:

Hosseinipoor, S. (2016) studied the subject to Forecasting Natural Gas Prices in the United States Using Artificial Neural Networks, In this paper it was suggested, three different categories of models are examined, namely stochastic differential equations, ARIMA and autoregressive neural networks. The results indicate that the NAR neural network adapts better to the given data than the other proposed models. The three-layer NAR model with 6 hidden neurons was found to be the best performing model for forecasting natural gas prices in the United States [5].

Hosseinipoor, S., et al. (2016) investigated the subject about Application of ARIMA and GARCH Models in Forecasting the Natural Gas Prices, In this paper, it was suggested a combination of ARIMA (5,1,9) and GARCH (1,1) models is a univariate time series model. Their model, ARIMA (5,1,9) / GARCH (1,1), demonstrates that the price of natural gas is likely to rise slightly in the future, which is not important and cannot exceed \$ 3.2 / MMBtu even in case of optimistic scenario. From their findings, the forecast graph indicates that in 2016 the price of natural gas tends to stay above \$ 1.5 / MMBtu. In other words, with 95% confidence, the price of natural gas will fluctuate between 1.5 and 3.2 dollars per million British thermal units [6].

Siddiqui, A. W. (2019) studied the subject of Predicting Natural Gas Spot Prices Using Artificial Neural Network. In this paper, they suggest using the ARNN model to forecast gas prices. The model was measured against the linear ARIMA model. ARNN performed well and showed a 33.4% improvement in performance compared to the ARIMA model. Also, compare their results with the typical neural network structure suggested in a previous article. The model was tested using the same data set, i.e. Henry Hebe Daily Spot Prices. The reported improvement in their work with MSE as a measure of error is about 58%. Obviously, such an improvement is important in terms of making immediate gas purchasing decisions [7].

Čeperić, E., et al. (2017) investigated the subject about Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. Among the results of short-term forecasting of instant natural gas prices for Henry Hub, Their findings are based on the performance of classic time series models and machine learning methods, specifically; neural networks (NN) and seasonal adjusted strategic regression support machines (SSA-SVR). It gives good results in short term forecasting of natural gas prices [8].

Azadeh, A., et al. (2012) studied the subject of A hybrid neuro-fuzzy approach for improvement of natural gas price forecasting in fuzzy and noisy environments: domestic and industrial sectors. their studies aimed to advance a hybrid mathematical approach to the hybrid neural network to improve the estimation of natural gas prices. It consists of the artificial neural network (ANN), the obscure linear regression (FLR), and the traditional regression (CR). The preferred FLR, ANN, and CR models are chosen by the average absolute error percentage (MAPE). Then the smart approach to this study is applied to estimate the prices of natural gas in both the domestic and industrial sectors in Iran. Given the lowest MAPE, ANN and FLR production, the best models for domestic and industrial sectors are respectively identified [9].

From the above researchs, ARIMA or NAR have shown that the natural gas price can be predicted. The authors tried to distinguish alternative models, but the authors used ARIMA to forecast in most articles.

The work aims to pick the appropriate forecast model for natural gas prices.

3. Methods and Materials

3.1 ARIMA Model

The ARIMA model was used as a control group to compare the performance of forecasts. The ARIMA model is widely used for research and forecasting and is often called the Box-Jenkins model or technique. It is considered the most successful social science prediction tool and is commonly utilized in time series. ARIMA is important for the projection of time series, as it implies that there is no knowledge of any underlying pattern or relationship as is the case for certain other approaches. ARIMA depends primarily on historical series values and previous prediction error situations. In comparison to short-run forecasting, though, ARIMA models are much more stable and effective than complicated systemic models [10]. The linear ARIMA models are the most common, the ARIMA models having an AR component called autoregressive and an MA component called moving average. Where ARIMA takes the following equation (1):

$$Y_t = \phi_{t-1}Y_{t-1} + \phi_{t-2}Y_{t-2} + \dots + \phi_{t-p}Y_{t-p} + C + \varepsilon_t - \theta_{t-1}\varepsilon_{t-1} - \theta_{t-2}\varepsilon_{t-2} - \dots - \theta_{t-q}\varepsilon_{t-q} \quad (1)$$

The parameters of forms p and q are determined by the autocorrelation function (ACF), the partial autocorrelation function (PACF) and the Akaike criterion information (AIC) [11].

After the p and q values have been determined by comparing several models, the maximum likelihood method is used in the estimation step.

According to the estimate of several models of the diagnostic step is to validate the model in the prediction process. For this purpose the series residues are examined and subjected to several tests to ensure the model does not include to problems arising from the error autocorrelation and the Heteroscedasticity problem.

Therefore, we diagnose the model, there are two cases. First, if the model is best, we immediately go to the last step, which is forecast. If the model is not good, then we do the feedback process by changing the p and q parameters, and then we make sure that the model is good until we reach the best model, see Fig.1.

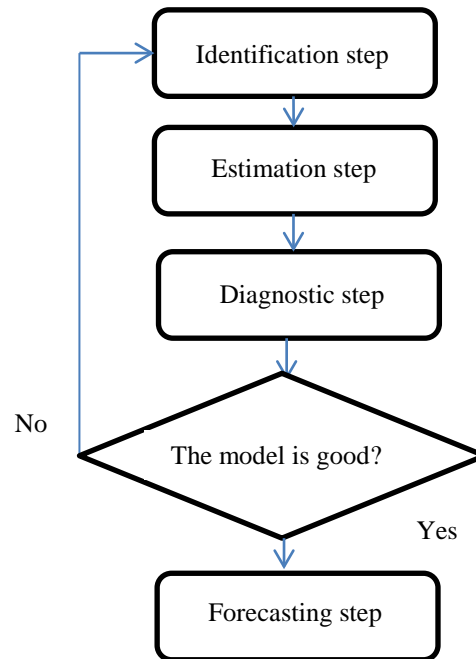


Fig. 1. ARIMA Processes

3.2 Nonlinear Autoregressive Neural Network

Neural networks are an artificial intelligence system that simulates the human brain using weights to communicate between neuronal units for the transmission of information.

Among the neural network types, there is a nonlinear autoregressive neural network, it has a two-layer network with a feedforward network, in addition to a sigmoid transfer function in the hidden layer and a linear transfer function in the output.

Nonlinear Autoregressive Neural Network has the ability to model nonlinear patterns, due to its flexibility convergence function, is a powerful method for predicting time series [12]. The main reason in their capability of correctly leveraging this non-linear information is in the time series for the building an accurate forecast model.

In predicting Nonlinear Autoregressive, future values depend on the previous values for the same variable under study (see Fig.2.), as it can be expressed mathematically as follows [13].

$$y(t) = f(y(t-1), \dots, y(t-p)) \quad (2)$$

Where:

P : number of delays.

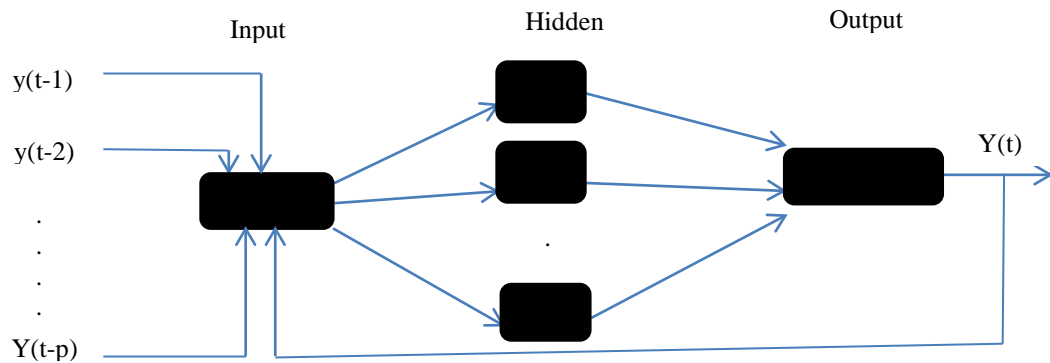


Fig. 2. Non-linear autoregressive (NAR) network

However, there are three methods of training were examined in this study of evaluating the performance of non-linear autoregressive model of neural networks.

And besides, not all algorithms are able to evaluate target output. In this study, attempts were made to implement three training algorithms such as Levenberg-Marquardt, Bayesian regularization and Scaled Conjugate Gradient. The detailed explanation related to the three training algorithms, is described as follows:

- The Levenberg-Marquardt algorithm (LM) is one of the common training methods used by the NAR method. This algorithm is characterized by a fast training process, also known as the least squares method. This algorithm is designed to work with a loss function, which takes the form of a sum of squared errors. The algorithm works without calculating the exact matrix of Hessain. Instead, it works with a gradient vector and a Jacobian matrix, the Hanassin calculation is rounded off in this algorithm and shown in equation (3), and the gradient is calculated in equation (4) [14].

$$H = J^T J \tag{3}$$

$$g = J^T e \tag{4}$$

Where:

J : Jacobian Matrix.

e : vector of network error.

It method is used by LM in the Newton-like update mentioned in Equation (5).

$$X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T e \tag{5}$$

Where:

X_{k+1} : is new weight.

X_k : is current weight.

- Bayesian regularization (BR) algorithm is a method based on reducing the composition from the square of errors, weights and biases, and this method continues in training until you obtain the optimum combination of weights and biases.

- Scaled Conjugate Gradient (SCG) can train any network as long as its weight, net input, and transfer functions have derivative functions is recommended as it uses gradient calculations which are more memory efficient than the Jacobian calculations the other two algorithms use [15].

3.3 Performance Evaluation

In order to evaluate the performance of the methods used in this study, each of the was used Coefficient of Determination (R2) is expressed as the ratio of the explained variance to the total variance, Mean Squared Error (MSE) of the model is the mean of the squared prediction errors for a test set and Root Mean Squared Error (RMSE) is just the square root of the mean square error, It is shown in the following table [16].

Table 1. validation criteria

Code	Definition	Equation
(R ²)	Coefficient of Determination	$R^2 = \frac{SSE}{SST}$
(MSE)	Mean Squared Error	$MSE = \frac{\sum(Y_t - \hat{Y}_t)^2}{N}$
(RMSE)	Root Mean Squared Error	$RMSE = \sqrt{\frac{\sum(Y_t - \hat{Y}_t)^2}{N}}$

Where:

Y : is the values of the output

\hat{Y} : is the predicted values

N : is the number of observations

4. Results and discussion

4.1 Data

In this study, we have dropped the theoretical aspect of the applied aspect, since the famous time series for Natural Gas Prices (NGP) was studied from January 2000 to December 2019, where the data was collected on the following website (<https://www.eia.gov/dnav/ng/hist/rngwhhdm.htm>).

Fig.3. shows the plot of the natural gas price series, where there are Greater volatility from periodic and this explains oil price volatility, where the average value was 4.57(Dollars per Million Btu) and the maximum value was 13.42 and the minimum value was 1.730 and a standard deviation It is equal to 2.17, and through these data it is clear that the data are heterogeneous, that is to say, they have a wide dispersion, and through Jarque Bera's value, which was 148.04 with a probability 0.00, we accept the alternative hypothesis that the data do not follow a normal distribution (see Table 2).

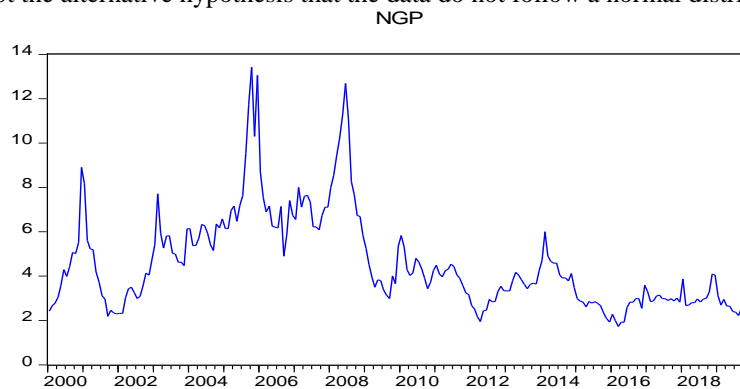


Fig. 3. Plot for a natural gas price series

Table 2. The descriptive statistics for the NGP

	NGP		NGP
Mean	4.574208	Kurtosis	5.487881
Median	3.980000	Jarque-Bera	148.0433
Maximum	13.42000	Probability	0.000000
Minimum	1.730000	Sum	1097.810
Std. Dev.	2.178603	Sum Sq. Dev.	1134.369
Skewness	1.467547	Observations	240

4.2 ARIMA Model

In ARIMA models, it is necessary to make the time series stationary, the fig.4, represents the autocorrelation function, Note that there is an autocorrelation in the time series. In addition to the Augmented Dickey - Fuller test was used, we find in table (3) that the probability value is greater than 5%, meaning that the time series is not stationary. so we use the difference method to make it stationary, in table (4) we find that the probability value is less than 5%, meaning that the difference series is stationary.

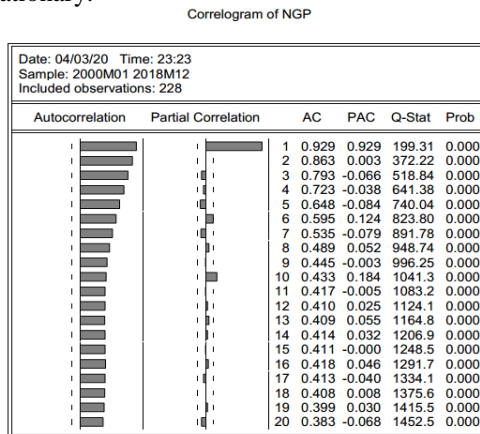


Fig. 4. Correlogram of NGP

Table 3. Augmented Dickey-Fuller of NGP

Null Hypothesis: NGP has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 0 (Automatic - based on SIC, maxlag=14)		
		t-Statistic Prob.*
Augmented Dickey-Fuller test statistic		-3.391094 0.0551
Test critical values:	1% level	-3.998997
	5% level	-3.429745
	10% level	-3.138397
*MacKinnon (1996) one-sided p-values.		

Table 4. Augmented Dickey-Fuller of D(NGP)

Null Hypothesis: D(NGP) has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 0 (Automatic - based on SIC, maxlag=14)		
		t-Statistic Prob.*
Augmented Dickey-Fuller test statistic		-15.49935 0.0000
Test critical values:	1% level	-3.999180
	5% level	-3.429834
	10% level	-3.138449
*MacKinnon (1996) one-sided p-values.		

For the identification of the model, we used autocorrelation function and partial correlation function of the differences series for a to calculate a sample of delays, the models are shown in the table (5) were identified, the ARIMA (5, 1.9) model was the best that based on the Akaike information criterion.

Table 5. Akaike information criterion

Model	AIC
ARIMA(5,1,5)	2.41
ARIMA(9,1,9)	2.36
ARIMA(9,1,5)	2.36
ARIMA(5,1,9)	2.35

After diagnosing the model, the maximum Likelihood method was used to estimate the model, shown in the table (6), where we note that all parameters are significantly different from zero.

Table 6. Estimate the ARIMA (5, 1.9) model

Dependent Variable: D(NGP)				
Method: ARMA Maximum Likelihood (OPG - BHHH)				
Date: 04/03/20 Time: 23:30				
Sample: 2000M02 2018M12				
Included observations: 227				
Convergence achieved after 13 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(5)	-0.158108	0.063979	-2.471256	0.0142
MA(9)	-0.253094	0.072345	-3.498436	0.0006
SIGMASQ	0.601439	0.031329	19.19730	0.0000
R-squared	0.080608	Mean dependent var		0.007137
Adjusted R-squared	0.072399	S.D. dependent var		0.810595
S.E. of regression	0.780701	Akaike info criterion		2.359060
Sum squared resid	136.5266	Schwarz criterion		2.404324
Log likelihood	-264.7534	Hannan-Quinn criter.		2.377325
Durbin-Watson stat	2.059458			

4.3 NAR Model

The forecasting using the non-linear autoregressive neural network are the following steps:

- Determine the series of Target, In this study we relied on the series of natural gas price, which contains 240 data.
- randomly divide to 70% for training, 15% for validation, and 15% for testing.
- In order to build the feedforward network, with a structure, it will import the transfer function in the hidden layer and a linear transfer function in the output layer (see Fig.5.).
- We tested a number of hidden neurons, out of a group of numbers, where number 11 was identified.
- We tested a number of delays, from a group of numbers (1,2,..., 12), where the number 5 was determined.
- In a training algorithm step, the Levenberg-Marquardt algorithm was chosen, which relied on minimizing the Mean Squared Error.

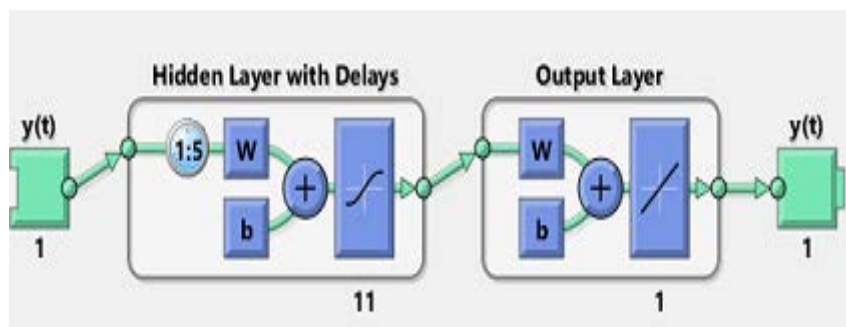


Fig. 5. NAR Architecture

- Whereas Best Validation performance is 1.22 at epoch 4 (see Fig.6.)

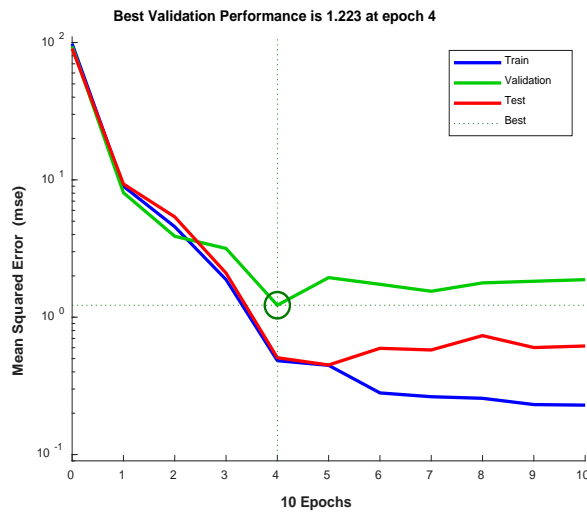


Fig. 6. Best Validation performance

- We have plotted and calculated the coefficient of determination For training, validation, testing and all samples. For comparing the Target values with the Output values. It is clear from Figure (7) that the coefficient of determination is close to one, this means that the NAR (5,11,1) is the best for modeling and forecasting the series of natural gas price.

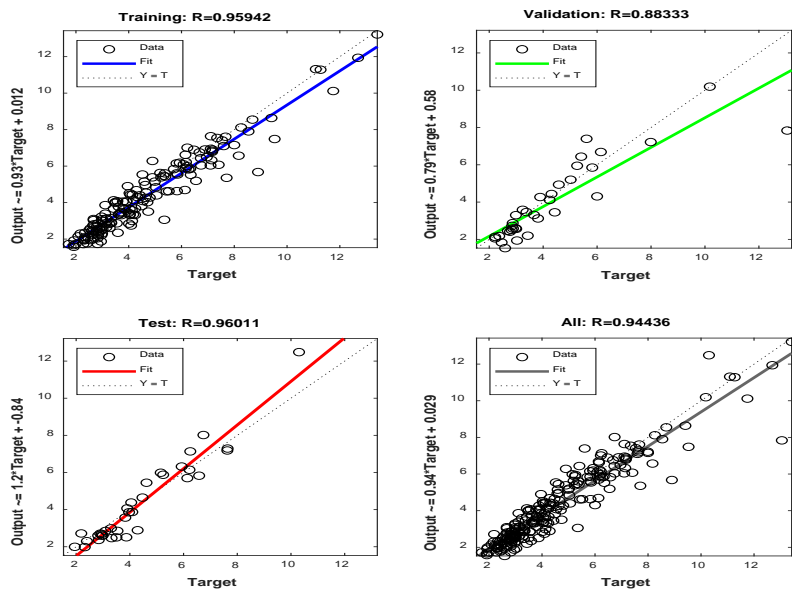


Fig. 7. the coefficient of determination

4.4 The comparison of models

To compare ARIMA model with NAR in the forecasting of natural gas prices for the 12 months of the year 2019, we calculated both MSE, RMSE and R2 shown in Table 7. Where the smallest value for RMSE and the highest value for R2 was for NAR model, therefore we concluded that NAR model is better than ARIMA model in forecasting natural gas prices, which is consistent with most of the preceding studies.

Table 7. A comparison of models

DATE	NGP	ARIMA(5,1,9)	NAR(5,11,1)
Jan-19	3.11	3.990472106	3.39161205
Feb-19	2.69	3.122059625	2.81272996
Mar-19	2.95	2.53967684	2.68738956
Apr-19	2.65	2.914333315	2.74894623
May-19	2.64	2.624779338	2.24579869
Jun-19	2.4	2.780934359	2.29983266
Jul-19	2.37	2.328767632	2.17730829
Aug-19	2.22	2.191319474	2.05430654
Sep-19	2.56	2.287636314	1.94693666
Oct-19	2.33	2.781444796	2.00535525
Nov-19	2.65	2.475641712	1.9914896
Dec-19	2.22	2.552252316	2.13095802
MSE		0.147232165	0.11049504
RMSE		0.383708437	0.33240794
R ²		0.42	0.61

5. Conclusion

In reality, study considers the natural gas as a critical commodity especially for a developing economy. The research therefore focuses on the forecast of natural gas prices by determining the best way to predict the future while taking into consideration its own lagging values. That study identified the more than decadal data on the natural gas prices, i.e. from 2000.

It is shown that the Autoregressive Integrated Moving Average ARIMA model is unable to fully capture the behavior of a nonlinear time series. This raises the need to develop a Nonlinear Autoregressive Neural Network model to adequately predict nonlinear time series, which can have NAR as a starting point in predicting the natural gas price.

This study applied both the Nonlinear Autoregressive Neural Network model and ARIMA model for forecasting natural gas prices, as obtained from the results analysis, NAR models were better than the ARIMA model, measured against three performance indicators MSE, RMSE and R². The results show that the NAR model was the best forecasting of natural gas prices.

References

- [1] Al-Fattah, S. M., and R. A. (2000). Startzman. Forecasting world natural gas supply. *Journal of petroleum technology* 52.05: 62-72.
- [2] Li, J., Dong, X., Shangguan, J., & Hook, M. (2011). Forecasting the growth of China's natural gas consumption. *Energy*, 36(3), 1380-1385.
- [3] Szoplik, J. (2015). Forecasting of natural gas consumption with artificial neural networks. *Energy*, 85, 208-220.
- [4] Kaboudan, M. A., & Liu, Q. W. (2004). Forecasting quarterly US demand for natural gas. *Information Technology for Economics and Management*, 2(1).
- [5] Hosseinipoor, S. (2016). Forecasting Natural Gas Prices in the United States Using Artificial Neural Networks. MASTER OF SCIENCE IN NATURAL GAS ENGINEERING AND MANAGEMENT, UNIVERSITY OF OKLAHOMA.
- [6] Hosseinipoor, S., Hajirezaie, S., & Nejati, J. (2016). Application of ARIMA and GARCH Models in Forecasting the Natural Gas Prices. *Student J. OKLAHOMA Univ.* PP.1-16.
- [7] Siddiqui, A. W. (2019, May). Predicting Natural Gas Spot Prices Using Artificial Neural Network. In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS) IEEE. PP. 1-6.
- [8] Čeperić, E., Žiković, S., & Čeperić, V. (2017). Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. *Energy*, 140, 893-900.
- [9] Azadeh, A., Sheikhalishahi, M., & Shahmiri, S. (2012, March). A hybrid neuro-fuzzy approach for improvement of natural gas price forecasting in vague and noisy environments: domestic and industrial sectors. In *Proceedings of the International Conference on Trends in Industrial and Mechanical Engineering*, Dubai, UAE (pp. 24-25).
- [10] Adebisi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014, 1-7.

- [11] Zhang, Y., Yang, H., Cui, H., & Chen, Q. (2019). Comparison of the Ability of ARIMA, WNN and SVM Models for Drought Forecasting in the Sanjiang Plain, China. *Natural Resources Research*, 1-18.
- [12] Wang, K. W., Deng, C., Li, J. P., Zhang, Y. Y., Li, X. Y., & Wu, M. C. (2017). Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiology & Infection*, 145(6), 1118-1129.
- [13] Beale, M. H., Hagan, M. T., & Demuth, H. B. (2016). *Neural network toolbox™ getting started guide*. Natick: The MathWorks.
- [14] Ruiz, L. G. B., Cuéllar, M. P., Calvo-Flores, M. D., & Jiménez, M. D. C. P. (2016). An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. *Energies*, 9(9), 684.
- [15] Markova, M. (2019, October). Foreign exchange rate forecasting by artificial neural networks. In *AIP Conference Proceedings* (Vol. 2164, No. 1, p. 060010). AIP Publishing LLC. 1-15.
- [16] Le, T. T., Pham, B. T., Ly, H. B., Shirzadi, A., & Le, L. M. (2020). Development of 48-hour Precipitation Forecasting Model using Nonlinear Autoregressive Neural Network. In *CIGOS 2019, Innovation for Sustainable Infrastructure* (pp. 1191-1196). Springer, Singapore.

Authors' Profiles



Dr. **Abdelkader Sahed** works at the University Center in Maghnia as a Lecturer, Department of Economic Sciences, Faculty of Economics, Management and Commerce. He has Masters and PhD both obtained at the University of Tlemcen. His research area includes Statistics and Probability Theories and econometrics. The author can be contacted at, sahed14@yahoo.fr.



Dr. **Mohammed Mékidiche** is currently Professor in the faculty of economics and commerce, University of Tlemcen, Maghnia Center, Algeria, where he teaches Statistics and Operations Research. He received the MS degree in production and operations Management at Faculty of Economics and commerce, University of Tlemcen, Algeria in 2005. He is currently a PhD holder in the field of production and operations Management at University of Tlemcen in 2012. His research project is optimization in production planning, fuzzy optimization and its application in production planning and scheduling. The author can be contacted at, mkidiche@yahoo.fr.



Hacem Kahoui is currently preparing PHD in the management and economics from the University Center in Maghnia, His research interests include decision making, goal programming, fuzzy sets, decision making, operation and production management, and Econometrics. The author can be contacted at, hacenkahwi@gmail.com

How to cite this paper: Abdelkader Sahed, Mohammed Mékidiche, Hacem Kahoui. " Forecasting Natural Gas Prices Using Nonlinear Autoregressive Neural Network ", *International Journal of Mathematical Sciences and Computing (IJMSC)*, Vol.6, No.5, pp.37-46, 2020. DOI: 10.5815/ijMSC.2020.05.04