Petroleum
Business
Review

Prediction of Natural Gas Price in European Gas Hubs Using Artificial Neural Network

Mansoureh Ram[a], Atefeh Taklif[b,*], and Ali Faridzad[c]

[a] *PhD Student, Department of Energy, Agriculture and Environmental Economics, Faculty of Economics, Allameh Tabatabai University, Tehran, Iran; Email: raam@nigc.ir*

[b] *Assistant Professor, Faculty of Economics, Allameh Tabatabai University, Tehran, Iran; Email: a.taklif@atu.ac.ir*

[c] *Associate Professor of Energy Economics, Faculty of Economics, Allameh Tabatabai University, Tehran, Iran; Email: ali.faridzad@atu.ac.ir*

ARTICLE INFO

Keywords:

Natural gas price prediction

Gas hub

Artificial neural network

JEL Classification: C53, C45

Received: April 1, 2019
Revised: May 15, 2019
Accepted: June 8, 2019

ABSTRACT

The liberalization of natural gas markets and the emergence of gas hubs in recent decades have shifted the natural gas trade from the regional to the global trade. The growth and maturity of these hubs have weakened the previously established relationship between the natural gas price and the prices of crude oil and petroleum products. Therefore, predicting the price of gas as a strategic commodity has become more important for different countries. Using the neural network method, this paper attempts to develop a model of the monthly prediction of natural gas price. Based on the time series data from 2012 to April 2019 as the input to the neural network, this model predicts the prices in five hubs and natural gas exchange centers in Europe. Based on the R^2 performance evaluation index of 98% of the neural network model fitted based on the aforementioned data series, the neural network model has acceptable performance in predicting the natural gas price. The results of this study show that using the artificial neural network (ANN) method, the gas prices in the European gas hubs, which are located in European countries, can be predicted with a high degree of accuracy.

1. Introduction

There are raised concerns over climate change and air quality in the world, limited growth and development of renewable energies, and failure to achieve low-carbon energy. These issues have caused that the natural gas with the lowest environmental pollution index and potentially easy and accessible use to play a significant role in the fuel consumption of countries in recent years (Su et al., 2019). Currently, the natural gas accounts for about 24% in the global fuel basket and 67% in the Iran's energy basket (BP, 2018).

Since natural gas is considered to be one of the important carriers of energy, it will also have a more important role in the future among the energy sources due to the environmental benefits. Therefore, it is very important to predict the price of natural gas. Predicting the natural gas prices as an essential and significant tool allows different stakeholders in the natural gas market to make better decisions when facing potential risks. On the other hand, predicting the price reduces the gap between supply and demand and optimizes the use of resources based on the accurate predictions (Su et al., 2019).

In many studies, the price of crude oil has been predicted, but since the natural gas price still follows the

* Corresponding Author

price of crude oil and petroleum products in some places, the prediction of natural gas price has less been studied compared to the oil price. In general, the price of natural gas in some markets including the Asian markets remains at a high level, and for solving the problem of high price of this energy carrier, developing the gas market and establishing the gas hubs[2] in East Asia are suggested. The prediction of natural gas price and use of prediction tools improve the efficiency of the developed markets. The accurate prediction of the natural gas price is essential to support effective investment and negotiation decisions on the import and export contracts (Jin and Kim, 2015). In addition, predicting the natural gas price not only provides an important guide on the effective implementation of energy policy making and planning, but is also very effective in the economic planning, energy investment, and environmental protection (Su et al., 2019).

Due to the importance of the oil and gas market, there have been several studies on the predictions in these markets where each study has examined and predicted a market aspect using a specific method. The studies on the monthly gas price prediction can be divided into two groups: the studies using econometric methods and the studies predicting the natural gas price using the machine learning tools such as neural network methods[3], genetic algorithm[4], support vector machine[5], etc. In either of the methods, the price of natural gas is predicted using either a single variable or several variables. The results of the above studies show that the accuracy of neural network prediction is significantly higher than that of the regression models in terms of error criteria. In other words, based on the research, the nonlinear and combined models, especially the neural networks, have a higher capability and accuracy than other models to predict the gas price because they include more factors in the modeling[6]. This paper aims to predict the gas price in the European hubs using the artificial neural network model. In the second section of this paper, the experimental studies of this field are reviewed. In the third section, the theoretical foundations are provided, and, in the fourth section, the application of neural network in the prediction of natural gas price, data analysis and description, performance evaluation features of prediction, the validation technique, and model parameters are investigated. Section 5 also presents the results of the neural network model, and Section 6 draws the conclusions.

2. Literature Review

Today, the future price or quantity is predicted using the computational and statistical sciences in different fields. Therefore, many methods have been considered for the analysis and prediction, among which the use of machine learning-based methods has attracted more attention in recent years. One of the studies conducted in this context is the research of Su et al. (2019) which predicts the price of natural gas in the Henry Hub market based on the 2001–2017 data by utilizing four models, namely artificial neural network (ANN), support vector machine (SVM), gradient boosting machine (GBM), and Gaussian process regression (GPR) models; the neural network model is better than the other three models based on the results. In another work, Jin and Kim (2015) predicted the natural gas price in Henry Hub using three methods, namely wavelet, time series, and neural networks, based on the 2000–2013 data, and showed that using the neural networks alone is the best method for the two-step prediction and combining the wavelet method with the ARIMA method is the best technique for the four-step prediction.

[2] Natural gas hubs tend to be at the heart of gas infrastructure networks such as pipelines and liquefied natural gas (LNG) terminals. The hub is used as a central pricing point for the network's natural gas. In some cases, a financial derivative contract is the priced-off gas delivered at this point as well. (Reuters)

[3] The artificial neural network is a computational model developed in 1943 by McCulloch and Pitts based on the mathematics and threshold logic algorithms (Su et al., 2019). A neural network (also called an artificial neural network) is an adaptive system that learns by using interconnected nodes or neurons in a layered structure that resembles a human brain. A neural network can learn from data, so it can be trained to recognize patterns, classify data, and forecast future events (mathworks.com).

[4] The genetic algorithm (GA) is a search technique used in computing to find exact or approximate solutions to search and optimization problems (Ganesh Bonde).

[5] The support vector machine (SVM) is machine learning algorithm that analyzes data for classification and regression analysis (techopedia.com).

[6] Price prediction can be performed through the economic and econometric methods such as time series methods, ARIMA, GARCH, machine learning methods (including genetic algorithm and support vector machine), and combined models.

In a paper reviewing the literature on natural gas prediction, Tamba et al. (2018) examined the prediction models presented for the production, consumption, demand, price and income elasticity, market liquidity, and price changes. Busse et al. (2012) estimated the day ahead spot price movement of natural gas in market area of NetConnect Germany[7] using the neural networks model with five factors, including temperature, exchange rate, and the price in three national balancing point (NBP), Net Connect Germany (NCG), and Dutch Title Transfer Facility (TTF) hubs based on the data from January 2010 to February 2011. Based on the results, the temperature has the greatest impact on the short-term gas price. Also, the prices predicted for the next four days have the highest impact.

In a research, Fabini predicted the gas prices in three NBP, ZEE and TTF hubs using the GARCH, EGARCH, and TGARCH models based on the 2008–2012 price data. The fitted models showed a positive correlation between the fluctuations of the three markets. Hosseinpoor (2016) predicted the natural gas price in the United States based on the 1997–2016 data in his master's thesis at the University of Oklahoma using the stochastic differential equations, ARIMA, and neural network and showed that the neural network model outperforms other models in predicting the prices.

However, in the field of studies conducted in Iran, the study by Mohammadi et al. (2017) entitled "Natural Gas Price Prediction Using Combined Wavelet Transform and Neural Network" based on the US gas market survey and the daily price data from 1997 to 2015 can be noted. In this study, the combined model of wavelet transform and neural network had a better prediction than the neural network model. Pourkazemi and Asadi (2009) used the artificial neural network model to dynamically predict the WTI crude oil price based on the 1998–2008 price data. In addition to the univariate neural network model, the oil storage data of OECD countries were added to the input to the network, and the results indicated the better prediction of the neural network model than the ARIMA model and the more accuracy of the bivariate model than the univariate model. In another study, Pourkazemi et al. (2005) predicted the urban gas demand using the fuzzy neural networks and ARIMA model based on the demand data from April 2003 to December 2004 and finally concluded that the neural network model provides the predictions with less error based on different criteria. Mirsoltani and Akhavan (2013) used the neural network model and fuzzy system to predict the energy price and consumption in the industry sector in Iran based on the 1996–2010 data and stated that, based on the consumption data of oil products and prices, the results of fuzzy logic have a smaller rate of error than the neural network model. Using a combination of neural network technique and rule-based expert systems, Abrishami and Varahrami (2011) predicted the gas price based on the price data during 2006–2010. Moshiri and Forootan (2004) compared the performance of the neural network model with the nonlinear GARCH models and ARIMA linear model for predicting the daily crude oil price during 1983–2003. They showed the superiority of the neural network model with a lower error rate. In a paper entitled "Prediction of Crude Oil Price Using Wavelet Smoothing and Artificial Neural Network", Behradmehr (2008) predicted the oil price based on the price data from January 2000 to September 2004. The results showed that the used combined model leads to the better performance in the prediction of oil price through the data denoising and smoothing.

In a study on the modeling and prediction of natural gas price using the artificial neural networks, Abbasi and Varnamkhasti (2014) sought to present a high-accuracy and acceptable price prediction model and indicated the higher accuracy of the fuzzy neural network model after examining the MLF networks with one and two hidden layers, RBF neural network, and neural-fuzzy network model. In a study aimed at predicting the natural gas consumption in the domestic and commercial sectors of Isfahan, Iran using the neural network model, traditional regression model, and ARIMA based on the consumption data from 2002 to 2011, Honari et al. (2016) found that the neural network models have less error than the other two methods. Abrishami et al. (2008) in another work modeled and predicted the gasoline price using the GMDH neural network from 1998 to 2007 and reported that the accuracy of the neural network predictions is significantly higher than the regression model, which indicates the high capability of the neural network model to model the complex processes and predict the nonlinear dynamic paths. Sotoudeh and Farshad (2012) in a paper entitled "Neural Network Application for US Gas Price Prediction" predicted the price using the variables of consumption, production,

[7] NetConnect Germany GmbH & Co. KG (NCG) together with Gaspool, is one of the two market area managers on the German natural gas market.

import and export of gas, crude oil reserves, and prices during 1949–2010 and showed that the model has an acceptable prediction.

As stated in the studies reviewed above, the prediction accuracy of the neural network model is significantly higher than the regression models in terms of error criteria. In other words, based on the aforementioned research, the nonlinear and combined models, especially the neural networks, have the capability and accuracy to predict the gas price better than the other models since they include more factors in the modeling.

Therefore, studies reviewed above show that there are publications regarding gas international price forecasting in European gas hubs. Therefore, the contribution of this paper is the forecasting of the gas international price in European hubs by utilizing the ANN as an accurate forecasting technique.

3. Theoretical Foundations

The natural gas market was established since the 1950s when the large Groningen field was discovered in Norway and then other fields in the Black Sea. In 1962, the Dutch gas exports to France, Belgium, and Germany began, and the natural gas pricing formula was based on the market value analysis where the market price of competing domestic fuels of each market, i.e. coal and gasoil at that time, formed the basis for the pricing. In addition, the gas price was determined according to the distance from the market and the storage capacity of each country.

Currently, Europe, Asia, and the US each have their own pricing mechanism. In Europe and in the areas where the free and integrated markets of natural gas have not formed or are emerging, the natural gas price is partly related to alternative energy prices, particularly fuel oil and thermal oil; in other sectors, it is determined by the market mechanism and the quantity of supply and demand. However, in recent years, the growth of the gas hubs in Europe has also influenced the pricing of gas in the international contracts and led to the gradual integration of national markets in the EU and the evolution of gas pricing in the long-term contracts. In other words, the gas market is emerging as a new market independent of the oil market, and for this reason, the relationship between the price of gas and the price of oil or petroleum products is weakening in this market. Therefore, as a result of the effects of market liberalization, this relationship is likely to change in the future and will evolve as what has happened in the US market in the past.

Certainly, there is no single mechanism for pricing the natural gas on a global scale. The differences in the gas pricing mechanisms in different markets are also due to the existence of gas-on-gas competition pricing indices, oil price index, monopoly, recursive and replacement methods, and market regulation.

The gas market can be divided into the three regional markets of the United States of America, Asia-Pacific, and Europe. Even in the last few years, the European market has been different in various segments, and with the development of the gas market on the continent and the emergence of gas hubs, the market spot pricing is actually emerging (Zajdler, 2012).

The review of the world gas market shows that there are currently two main pricing mechanisms for natural gas in the three regional markets. In the early 1960s in Norway and its borders, the pricing method based on the price of oil and its products was developed by gas producers. This method is the pricing based on the natural gas alternative fuels, which was also used in the Asia and is still practiced in some parts of the world. However, the second pricing method which was initially established in the United States and later transferred to the UK and other parts of Europe was based on the market or hub price where the gas price varies from market to market.

Over time, with dividing the British Gas Company into multiple sectors, separating the distribution from the commercial sector, and providing the access for both sectors to the transmission network, a virtual gas balancing location, which was later introduced as the British National Balancing Point (NBP), was established to conclude the spot contracts; thus, natural gas was traded like other commodities and similar to crude oil.

This pricing model was then applied to the British NBP in 1998 for the Dutch and Belgian markets, and lastly, the European gas producers in the British market took action considering the NBP hub gas price as a reference price for the Belgian Zeebrugge (ZEE) and Dutch Title Transfer Facility (TTF) hubs that quickly became the two important natural gas hubs in Europe. Currently, other local hubs across Europe determine the gas price based on the supply and demand.

Since the mid-2000s, the natural gas in Europe and the US has been traded based on the hub price rather than oil price index. Since 2001, the European Commission has implemented three waves of reform, most notably the

Third Energy Package in 2009, which aims to liberalize the European energy markets. In fact, Europe entered a period of reforms similar to Britain in the 1990s. One of the most important structural changes by Europe as an integral package was implemented in 2011 (Grandi, 2014), which has contributed to the development of the European energy and gas markets.

The development of gas infrastructure in the member states, including the pipelines, LNG terminals, and storage facilities, which has led to the flexibility of the gas supply chain, has made a significant change in the EU's gas market pattern, leading the gas market development to be subject to the competition laws and liberalization principles, which were adopted by the EU member states and are gradually being implemented.

The most progress has seen in the development of European gas hubs in the last decade, while few number of European countries had a traded gas hub in 10 years ago. Nowadays, new trading hubs were gradually formed and there are still 7 countries without a traded gas hub. However, It is possible that all European countries have a single gas market by 2025 (Patrick Heather, 2019).

The managing long term gas contracts, as oil-linked purchase prices rose significantly above hub-based prices, encountered increasing commercial difficulty in the late 2008 and reduced sales prices to customers. Therefore, Gas hub-based prices and oil-linked converged at the end of 2010, but this convergence was only temporary and was partly accounted for by very cold weather in Europe. In the context of a surplus of European gas supply over demand during this period (2010-2014) this was a commercially untenable position for European gas buyers (Jonathan Stern and Howard Rogers, 2011). The transition away from oil-linked and towards hub-based prices continued in the following years and became as a dominant price-setting mechanism (Jonathan Stern and Howard Rogers, 2014).

Hub-based pricing is supported to be able to better reflect the fundamental value of natural gas (Shi and Variam, 2016; Stern, 2014; Zhang et al., 2018a). In other words, the hub price of natural gas is determined by demand and supply in natural gas market (Zhang D. and et al., 2018b).

Stern (2014) and Zhang et al. (2018a) argue that oil and natural gas are not necessarily substitutes and have different underlying fundamentals. Their idea is consistent with the efficient market hypothesis (EMH) introduced by Fama (1965). The main argument is very simple: the price should reflect the fundamental value of underlying asset and in an efficient market it should respond to shocks accurately and quickly (Zhang D. and et al., 2018b). Supporting arguments claim that hub pricing can better reflect fundamentals in natural gas and thus create better efficiency. By contrast, the opinions against hub pricing and in favor of oil indexation advises that oil indexation is the best remedy for gas market failure (Zhang D. and et al., 2018b).

Obviously, the replacement of gas pricing mechanisms in different markets is done by exclusive methods and oil-based or recursive pricing with the pricing method based on the price index and gas-on-gas competition and, in other words, hub price index. Therefore, it is important to predict the gas prices in the hubs for the producers and also for the consumers. Given the enormous volume of Iran's gas reserves of about 34 trillion cubic meters (BP, 2018), the analysis of the price and pricing indices is very important. Therefore, it is necessary to predict the gas prices in the European hubs as a potential market for Iran.

4. Methodology

4.1. Application of Neural Networks for Prediction of Natural Gas Price

The neural network is a computational model developed in 1943 by McCulloch and Pitts based on the mathematics and threshold logic algorithms. The neural network is a framework to attract machine learning algorithms for the cooperative work and thus, it is not an algorithm. The neural network has become more important since the 1980s and has served as the focus of research in the field of artificial intelligence which has wide applications in the data processing, classification, performance approximation and numerical control (Su et al., 2019).

As one of the new approaches that has also been considered in the field of economics in Iran, artificial intelligence provides linear and nonlinear tools for the prediction. In other words, the artificial neural network, as an intelligent system, can detect linear and nonlinear relationships between inputs and outputs based on training data and identify their fundamental relationships and then generalize the detected relationships to other data. As such, by properly designing the neural network architecture and selecting the right training data, one can achieve a structure that can predict the time series (Pourkazemi and Asadi, 2009).

The neural network processes information in the same way as the neural network of the human brain by connecting different units and nodes called artificial neurons to create different complex networks. As it is shown in figure 1, each node contains an activation function to generate output based on one or more inputs. The output signal of one node can be transmitted by a weight connection to another node (Su et al., 2019).

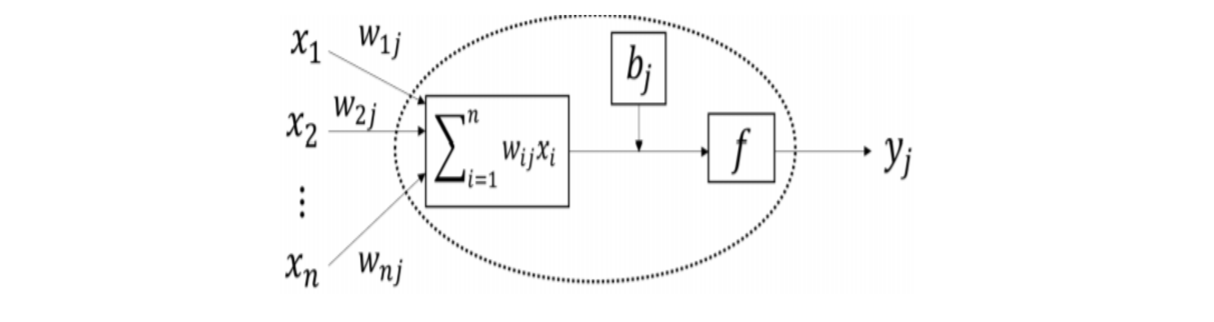


Figure 1. The structure of artificial neuron in ANN.

The relationship between the target variable (yj) and the independent variable (Xi) is obtained from Equation (1):

$$yj = f(\sum_{i=1}^{n} wijxi + tj) \quad (1)$$

Xi (i = 1, 2,… n) is the input data with the weights Wij. There are two basic units, namely summing and activating the input signals in the processing unit. Also, Yj is the unit output which is defined as Equation (2):

$$yj = f(\sum_{i=1}^{n} wijxi + tj) \quad (2)$$

The error backpropagation training algorithm is also used to train the networks. The algorithm was developed in 1986 by David Rumel hart and James McClelland. The error backpropagation is designated because the calculated error is returned from the output layer to the intermediate layer and finally to the input layer. The accuracy index of this algorithm is the mean square error. The activation function for the hidden and output layers is a standard sigmoid function, as shown in Equation (3):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

he output Yj is transmitted from the next layer as the input signal to the connected units. According to the designed model, all the units in the artificial neural network are interconnected in different layers. A simple example of a three-layer neural network is shown in Figure 2. In this network, the information flows through the input, hidden, and output layers where the input layer or the first layer contains the same number of units as the input vector. Then, there is the hidden layer with the desired number of units, and finally, the output layer is the weighted output of the hidden layer unit.

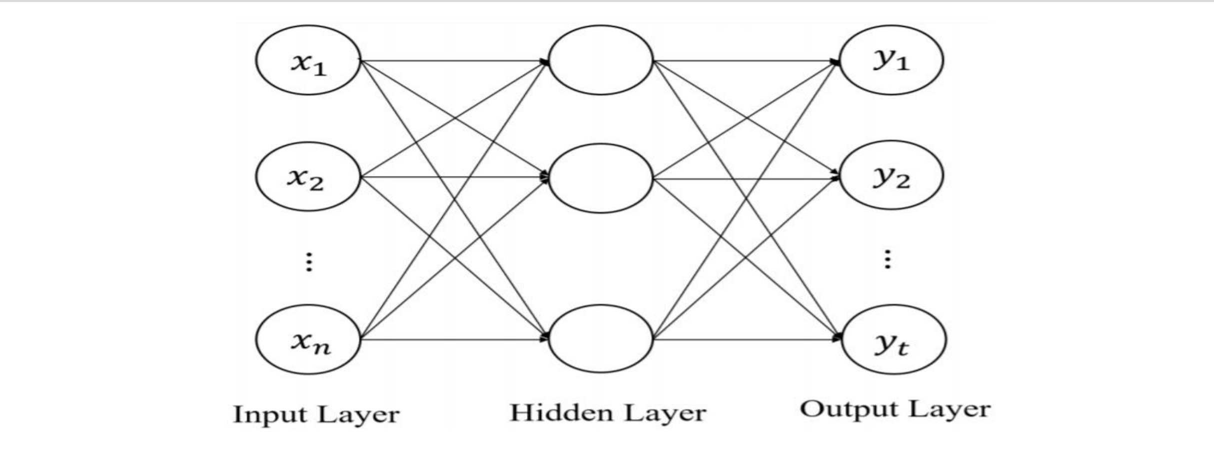


Figure 2. A simple structure of Artificial Neural network that comprise three layers.

The neural network intelligence is derived from the learning process, which allows the network to be capable of automatic adaptation, communication and memory for specific tasks. Although the gas price is primarily determined by the factors such as supply and demand, it is also influenced by the factors such as product inventory, stock market activities, exchange rate, and political issues (Abrishami and Varahrami, 2011).

Unlike traditional methods, neural networks are the self-adaptive and data-driven networks where there are little assumptions about the models used for different problems. The neural network provides better analyses when there are ambiguities in the independent variables than the linear regression analysis. In addition, the network structure is trained using part of the data and then tested using the rest of the data (Abrishami and Varahrami, 2011).

In addition, the neural networks are easy to generalize, so that after the learning using the sample data, they are able to correctly elicit and derive the unobserved parts of the population even with the possibility of disturbance in the sample data. In other words, the capabilities of these networks are the capability of training (adjust network weights using training data) and generalization (after training the network using the training data and adjusting the network weights, the network is able to accept an input and to provide an appropriate output) (Menhaj, 2005). Also, the biggest advantage of the neural networks is the ability to model complex nonlinear relationships.

4.2. Data Preparation and Description

Many factors such as crude oil price, fuel oil and gas oil prices, drilling activities, temperature, supply and demand of natural gas, storage, import of gas, etc. affect the price of natural gas. To predict the gas price as best as possible, these data should be taken into account (Su et al., 2019). However, it is somewhat impossible for the European Union (EU) to obtain the above data with respect to the gas transmission and distribution centers on the continent and to apply the data to the model. Therefore, the prediction of natural gas prices in the European hubs is based on the published historical price data as univariate.

Currently, the active gas hubs and transmission centers in Europe include the NBP (National Balancing Point), TTF (Title Transfer Facility), ZEE (Zeebrugge), NCG (Net Connect Germany), PSV (Punto di Scambio Virtuale), CEGH (Central European Gas Hub), GPL (Gaspool) and PECG (Points d'Échange de Gaz), and given the availability of the price data of the NBP, TTF, ZEE, NCG and PECG hubs, the price predictions were made for the five selected hubs.

The price predictions are based on the time series data from monthly hub prices available from 2012 to 2019. The comparison of gas prices in different hubs is shown in Figure 3.

Figure 3 shows that the natural gas prices do not have a regular upward or downward trend. In addition, the intensity of the fluctuations and their periods are not constant and identical. This indicates the existence of a nonlinear structure in the gas price data series.

During the study period, the average price in each hub, maximum price and minimum price were determined. Table 2 shows the results from the description and analysis of the natural gas spot price data in each of the hubs and the related variables including the maximum, minimum, median, mean and standard deviation.

Table 1. List of variables used in this study.

Variable	Data	Unit
Natural gas price	Gas price in NBP hub	$/thousand cubic meters
	Gas price in TTF hub	
	Gas price in NCG hub	
	Gas price in the ZEE hub	
	Gas price in PECG hub	

Source: Research findings

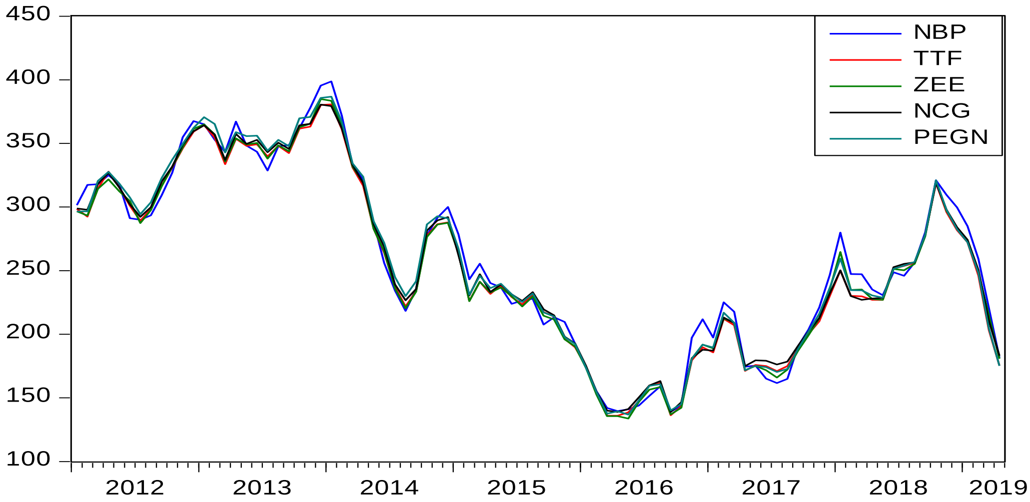


Figure 3. Historical trend of natural gas prices in European hubs.

Source: Data published by Platts and Argus

Table 2. Statistical summary of natural gas prices in selected European hubs during 2012-2019. ($/thousand cubic meters)

Hub	Max	Min	Median	Mean	SD
NBP	399	139	248	259	70.7
TTF	381	136	244	254	69.5
NCG	381	139	249	256	69.2
PEGC	387	137	247	258	71.2
ZEE	385	134	245	255	70.1

Source: Research findings

Table 3 shows the correlation between gas prices in different hubs. As shown in the table, there is a correlation between the gas prices in different hubs. The highest gas price correlation is seen in the two TTF and NCG hubs and the lowest gas price correlation in the two NBP and NCG hubs. The price correlations in different hubs are positive, meaning that increasing or decreasing the price in one hub leads to the increased or decreased prices in other hubs. The higher the correlation of prices in the two hubs, the greater the price changes between the two hubs.

Figure 4 shows the correlation between European hubs in a graphical and pairwise manner. As it can be seen, the gas price in all of the five hubs are nearly correlated and these prices move together in one direction.

Table 3. The correlation values between gas prices in different hubs.

Hub	NBP	TTF	ZEE	NCG	PEGC
NBP	1.0	-	-	-	-
TTF	0.9920	1.0	-	-	-
ZEE	0.9945	0.9991	1.0	-	-
NCG	0.9914	0.9996	0.9988	1.0	-
PEGC	0.9922	0.9993	0.9990	0.9992	1.0

Source: Research findings

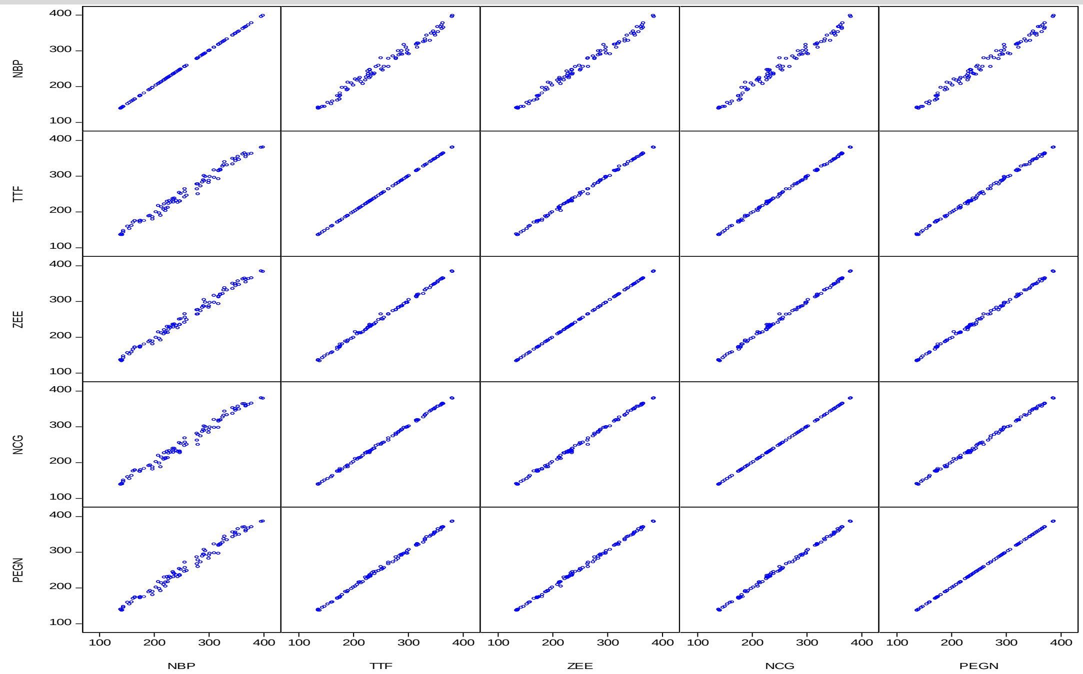


Figure 4. The charts of price highlight correlation of gas hubs in Europe.

Source: Research findings

4.3. Prediction Performance Evaluation Indicators

There are many evaluation criteria for measuring the performance of a prediction model. In fact, the criteria examine the validity of the prediction methods (difference between real and predicted value of dependent variable), some of which are given in this section for measuring the performance of the model. The R2 or R-square index is used to measure how the prediction model matches to actual data and obtained using Equation (4):

$$\mathrm{R-square(R2)} = \left[1 - \frac{\frac{1}{N}\sum_{t=1}^{N}(\tilde{y}t - yt)^2}{\mathrm{var(y)}}\right] \quad (4)$$

The values of yt and y ̃t represent the real values and predicted values at time t, respectively. N also represents the number of data used. R2 represents the conformity of the data with the estimated model and ranges 0-1, and the better the estimated model exhibits the changes, the closer the R2 to unity. In other words, in a model estimated with R2 greater than 0.8, the estimates can be sufficiently matched to the existing data.

Mean absolute error (MAE) is a performance evaluation tool for the prediction model which calculates the MAE based on Equation (5).

$$\mathrm{MAE} = \frac{1}{N} * \sum_{t=1}^{N} |\tilde{y}t - yt| \quad (5)$$

Mean square error (MSE) represents the mean squared deviation of the predicted value and the real value and is used to measure the amount of variations. The prediction model performs better when the MSE is lower. In contrast to MAE, the mean square error increases the prediction deviation. The mean square error is calculated based on Equation (6):

$$\mathrm{MSE} = \frac{1}{\mathrm{N}} * \sum_{\mathrm{t}=1}^{\mathrm{N}} (\tilde{\mathrm{y}}\mathrm{t} - \mathrm{yt})^2 \quad (6)$$

The root mean square error (RMSE) can be directly obtained by calculating the square root of the mean square error (MSE), which is sensitive to very large or very small error values and, as a result, is a good reflection on the prediction model and calculated based on Equation (7):

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} * \sum_{t=1}^{N} (\tilde{y}t - yt)^2} \quad (7)$$

The less the error model, the better the prediction capability. In other words, the lower the RMSE, the higher the capability of the prediction model.

Mean absolute percentage error (MAPE) is often used as a loss function, because it can intuitively describe the relative errors. It not only considers the deviation of the prediction value and real value, but also considers their ratio. To calculate the mean absolute percentage error, Equation (8) is used:

$$\mathrm{MAPE} = \frac{1}{\mathrm{N}} * \sum_{\mathrm{t}=1}^{\mathrm{N}} \left|\frac{\tilde{\mathrm{y}}\mathrm{t} - \mathrm{yt}}{\mathrm{yt}}\right| \quad (8)$$

4.4. Selection of Model Parameters

The machine learning methods have many parameters, some of which are the key parameters that should be carefully selected. The model complexity often depends on these parameters, which are called the model selection parameters. In a neural network model, a nonlinear auto-regression model with external input is selected (Su et al., 2019).

Selecting the number of hidden layers is very important, because the large number of the layers reduces the network performance. In theory, a neural network with two hidden layers is able to estimate any nonlinear function with an arbitrary degree of accuracy (Mohammadi et al., 2017).

4.5. Neural Network Validation and Verification

Cross-validation is an important technique in model selection to obtain the practical and stable models in the machine learning. In the cross-validation, the K validation is a common method to prevent excess connections in very complex models. Using the validation, K requires less data. Due to the MSE value of the validation samples, the training stops automatically when it is not possible to improve the generalization.

The training data is divided into K and its K-1 times is used for the model training and the rest is used for the validation. This training and validation operation is often performed by rotating the K folds. After performing this operation for K times, all errors are collected to calculate the final error. K is often considered equal to 10.

In the neural model used to predict the gas prices in the European hubs, the number of hidden neurons and number of layers are 10 and 2, respectively. In the training network, the selected training algorithm is the Levenberg-Marquardt algorithm which typically requires more memory and less time. In addition to building a neural network, a two-layer feedforward neural network with complete connection between the nodes is used, and the complete network information is shown in Table 4.

Table 4. The parameters of ANN that used in the designed model in this study.

Input layer nodes	Hidden layer nodes	Output layer nodes	Transfer function: hidden layer and output layer	Training and test data size
16	10	5	Tansig	20 and 80%

Source: Research findings

The feedforward backpropagation algorithm of artificial neural network was used in MATLAB R2017b software to design the prediction model.

5. Neural Network Simulation Results

Using the designed neural network model, applied data and prediction performance evaluation indicators, model validation technique, and model parameters, the prediction model performance indicators were extracted.

Figure 5 shows the model performance and validation indicators based on the results are as follows:

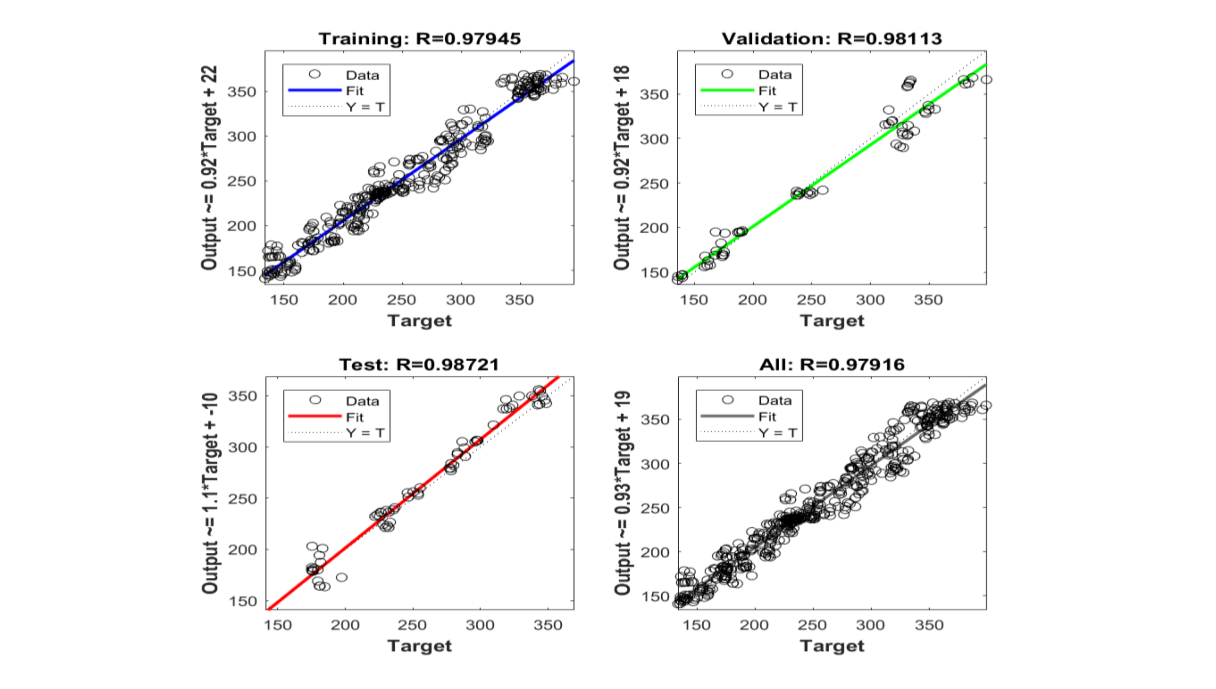
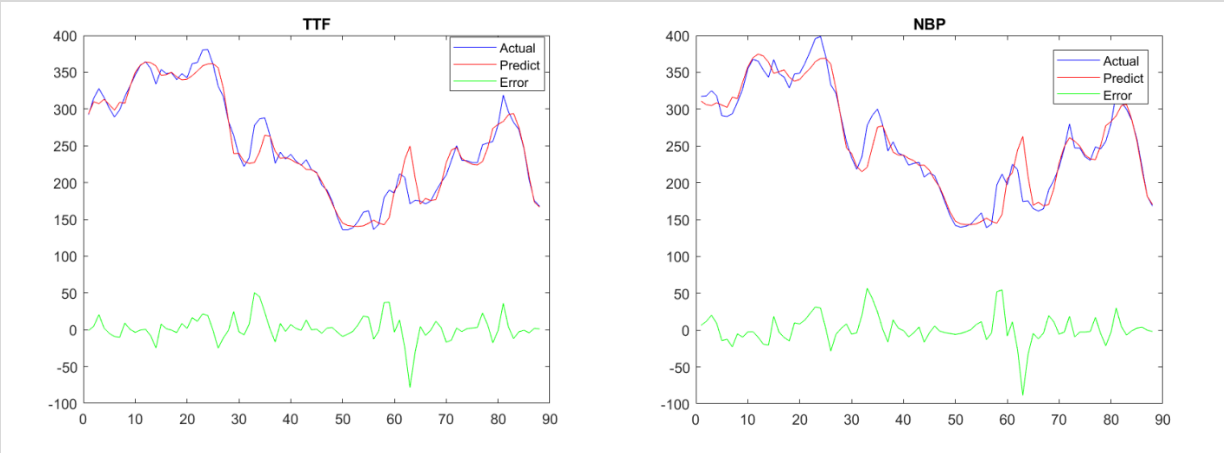


Figure 5. The performance and validation indicators validate the prediction model.

Source: Research findings

As shown in the figure 5, the R2 index is estimated for training data, test data, validation, and, generally, overall model. The index is about 98%, which indicates the very good performance of the prediction model through the use of the artificial neural network method.

Figure 6 shows the projected values of natural gas prices in the five European hubs based on 88 gas price data from the beginning of 2012 to April 2019.



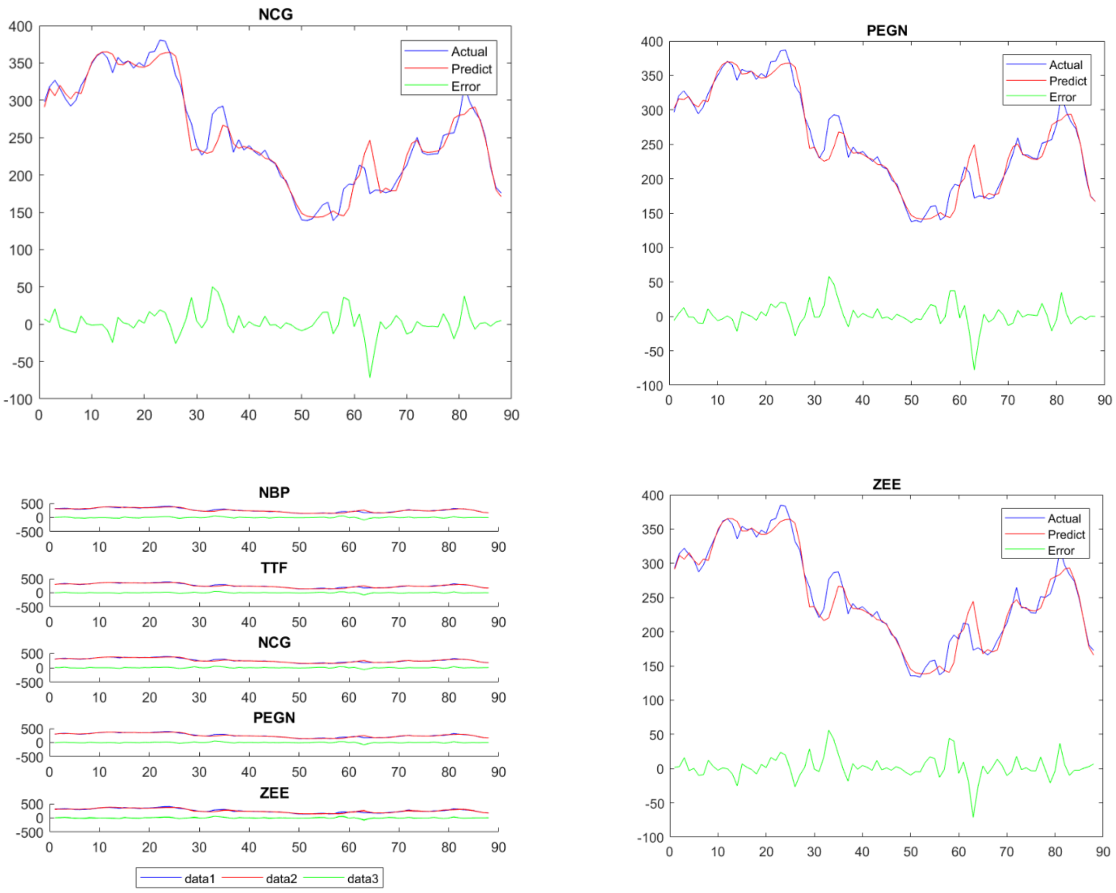


Figure 6. The comparison of prediction model and real prices in European hubs indicate an acceptable fitted model that cover the real values.

In the figure 6, the blue lines are the historical data the real values used to fit the prediction model, the red lines are the predictions based on the model, and the green lines are the model error rates fitted from the real values. As shown in the diagrams above for each of the gas hubs, the predicted values based on the fitted model were able to significantly cover the real values, and the R2 value of the model was about 98%, which shows a significant match. Based on the results of the neural network model, it can be stated that the neural network is well capable of self-learning, self-adaptation and self-organization which can analyze the patterns and rules of observed data and form complex nonlinear through the training and adapt to large-scale, multi-factor, incomplete and inaccurate data processing.

Based on the above results, the used ANN model has better prediction performance and accuracy (98%) than machine learning methods such as SVM (84%), GBM (80%) and GPR (83%) and also ANN (89%) that applied in Su et al. (2019) study. In the other study that conducted by Sebastian Busse et al. (2012), ANN model used for predicting the natural gas spot prices of the three major hubs (NBP, TTF, and NCG). The result revealed a forecasting accuracy of 64%.

Therefore, based on the historical data and using the ANN method in this study, the gas prices in the European gas hubs can be well predicted, where the predicted data can be widely used in the areas of investment, negotiating the import and export contracts, and, more generally, energy policy making and planning.

6. Conclusions

Due to the less pollution and easy access, natural gas has been able to make a significant contribution to the energy basket of countries in recent years. The importance of this fuel in recent years has caused the price prediction to be taken into account and various models have been introduced for the price prediction, and all these efforts have aimed to provide a prediction model of less error. Based on the conducted studies and according to the results of previous studies, the neural network model can well predict the price values . The purpose of this study was to predict the prices of natural gas in the European gas hubs based on the machine learning method using the neural network model.

As such, the monthly data on natural gas prices from the beginning of 2012 to May 2019 were used for the prediction. The prediction model inputs are the price data of the NBP, TTF, ZEE, NCG, PEGC hubs and the cross-validation method is used for the model learning. The performance evaluation factors of the prediction model also include the R2, MSE, RMSE and MAPE in the prediction method, where the R2 value obtained in the fitted neural network model is 98% which shows a very good performance in predicting the natural gas price.

To carry out further studies in this field, it is suggested to predict the gas prices in the European gas hubs using the multivariate models. Obviously, increasing the number of price-related variables such as gas storage, demand and supply, air temperature, etc. can improve the model output.

The prediction of natural gas prices in the European markets in the coming years can help to conclude the natural gas export contracts in Iran. Since new long-term contracts in the European markets are concluded based on the price index of gas sold in the hub, it is necessary to continuously examine the price trends in these markets, influencing variables, and also price prediction in the above markets to be included in the future contracts of Iran with the purchaser countries.

References

Abbasi, S. Varnamkhasti, J. G. 2014. Natural Gas Price Modeling and Prediction Using Artificial Neural Networks. Khorramshahr University of Marine Science and Technology.

ABRISHAMI, H. & VARAHRAMI, V. 2011. Different Methods for Gas Price Forecasting. Cuadernos de Economía, 34, 137-144.

Abrishami, H. Moeini, A. Mehrara, M. Ahrari, M. and Soleimani Kia, F. S. 2008. Modeling and Prediction of Gasoline Price Using GMDH Neural Network. Iranian Economic Research, 36, 37-58.

Behradmehr, N. 2008. Crude Oil Price Prediction Using Wavelet Smoothing and Artificial Neural Network. Journal of Energy Economics Studies, 18, 81-97.

BP. 2018. BP Statistical Review 2018 [Online]. [Accessed].

BUSSE, S., HELMHOLZ, P. & WEINMANN, M. 2012. Forecasting Day Ahead Spot Price Movements of Natural Gas, an Analysis of Potential Influence Factors on Basis of a NARX Neural Network. Multikonferenz Wirtschaftsinformatik 2012, Tagungsband der MKWI 2012, 1-12.

Dayong Zhang, et al, 2018. Is hub-based Pricing a Better Choice than Oil Indexation for Natural Gas? Evidence from a Multiple Bubble Test. Energy Economics.

Ganesh Bonde, Stock Price Prediction Using Genetic Algorithms and Evolution Strategies. Institute of Artificial Intelligence University Of Georgia Athens, GA-30601.

GRANDI, L. 2014. European Gas Markets: from Oil Pndexation Prices to Spot Prices. Energy Brains, Junio.

Honari, A., Yaghini, M. and Nadimi, M. 2016. Prediction Household and Commercial Gas Consumption over Five-Year Period in Isfahan Using Neural Networks. Production and Operations Management, 1, 247-262.

HOSSEINPOOR, S. S. 2016. Forecasting Natural Gas Prices in the United States Using Artificial Neural Networks. Master's Thesis, Oklahoma.

https://www.techopedia.com/definition/30364/support-vector-machine-svm

JIN, J. & KIM, J. 2015. Forecasting Natural Gas Prices Using Wavelets, Time Series, and Artificial Neural Networks. PloS One, 10, e0142064.

Jonathan Stern and Howard Rogers, 2011. The Transition to Hub-Based Gas Pricing in Continental Europe. The Oxford Institute for Energy studies.

Jonathan Stern and Howard Rogers, 2014.The Dynamics of a Liberalised

European Gas Market: Key determinants of Hub Prices, and Roles and

Risks of Major Players. The Oxford Institute for Energy Studies.

Kazemi, M. H. P., Pour, M. H., Afsar and Nahavandi. 2005. Comparative Study of ARIMA Linear and Nonlinear Methods of Fuzzy Neural Networks for

Prediction of Urban Gas Subscription Demand. Economic Research, 40.

Menhaj. 2005. Neural Network Basics. Amirkabir University of Technology.

MIRSOLTANI, S. M. & AKHAVAN, T. 2013. Forecasting Energy Price and Consumption for Iranian Industrial Sectors Using ANN and ANFIS. Iranian Journal of Economic Studies, 2, 73-102.

Mohammadi, T., Taklif, A. and Zamani, S. 2017. Prediction of Natural Gas Prices Using Combination of Wavelet Transform and Artificial Neural Network (Case Study: US Market). Iranian Economic Research, 22, 1-26.

Moshiri, S. and Forootan, F. 2014. Chaos Test and Prediction of Future Crude Oil Prices. Iranian Economic Research, 21, 67-90.

Patrick Heather, 2019. European Traded gas Hubs: a Decade of Change. The Oxford Institute for Energy studies.

Pourkazemi and Asadi, B. 2009. Dynamic Prediction of Crude Oil Price Using Artificial Neural Networks and Oil Storage of OECD Countries. Economic Research, 44.

SOTOUDEH, M. & FARSHAD, E. 2012. Application of Neural Network for Forecasting Gas Price in America. The Journal of Mathematics and Computer science, 4, 216-226.

SU, M., ZHANG, Z, ZHU, Y, ZHA, D & WEN, W. 2019. Data Driven Natural Gas Spot Price Prediction Models Using Machine Learning Methods. Energies, 12, 1680.

TAMBA, J. G., ESSIANE, S. N., SAPNKEN, E. F., KOFFI, F. D., NSOUANDÉLÉ, J. L., SOLDO, B. & NJOMO, D. 2018. Forecasting natural gas: a literature survey. International Journal of Energy Economics and Policy, 8, 216-249.

Xunpeng Shi, 2016. Gas and LNG Trading Hubs, Hub indexation and destination Flexibility in East Asia. Energy Policy.

ZAJDLER, R. 2012. The Future of Gas Pricing in long-Term Contracts in Central Eastern Europe. Global market Trends Versus Regional Particularities. Warsaw