

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

FORECASTING NATURAL GAS PRICES IN THE UNITED STATES USING
ARTIFICIAL NEURAL NETWORKS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN NATURAL GAS ENGINEERING AND
MANAGEMENT

By

SEYEDSAEID HOSSEINIPOOR

Norman, Oklahoma

2016

FORECASTING NATURAL GAS PRICES IN THE UNITED STATES USING
ARTIFICIAL NEURAL NETWORKS

A THESIS APPROVED FOR THE
MEWBOURNE SCHOOL OF PETROLEUM AND GEOLOGICAL ENGINEERING

BY

Dr. Suresh C. Sharma, Chair

Dr. Zhen Zhu

Dr. Xingru Wu

© Copyright by SEYEDSAEID HOSSEINIPOOR 2016
All Rights Reserved.

I dedicate this thesis to my beloved wife, Nooshin Nassr, who has been a constant source of support and love during the challenges of graduate school and life. I am truly thankful for having you in my life. You taught me how to love and how to devote.

This work is also dedicated to my parents, Mahmoud and Farideh, who have always loved me unconditionally, and also my brothers Sajjad and Vahid, my sister Matin, and my niece Hana, who are in my heart anytime and anywhere.

Acknowledgements

I would first like to thank my academic advisor Dr. Sharma. The door to Prof. Sharma's office was always open whenever I ran into a trouble spot or had a question.

I would like to thank my thesis advisor, Dr. Zhu, for his support, patience, and encouragement throughout my work.

I would also like to show my gratitude to Dr. Wu for comments that greatly improved the manuscript.

I would also like to thank Sassan Hajirezaie, Amirhossein Kamali, and Nooshin Nassr as the second readers of this thesis, and I am gratefully indebted to them for their very valuable comments on this thesis.

I would also like to acknowledge EIA and NOAA for their online free data.

Finally, I must express my very profound gratitude to Dean Stice, Dean Paul, and Dr. Abousleyman and special thanks to Dr. Pournik, for their gratis support. I have never been able to complete my graduate study without their help.

Table of Contents

Acknowledgements	iv
Table of Contents	v
List of Tables	viii
List of Figures.....	ix
Abstract.....	xiv
Chapter 1: Introduction	1
1.1. What is Natural Gas.....	2
1.2. Forms of Natural Gas	4
1.3. Regulations on Natural Gas Market	8
1.4. Natural Gas Market	10
1.5. Chapters and Contexts	12
Chapter 2: Literature review	15
Chapter 3: Methodology.....	20
3.1. Basic Definitions	20
3.1.1. Random Variables	21
3.1.2. Stochastic Process	21
3.1.3. Monte Carlo Simulation	22
3.1.4. Wiener Process	22
3.1.5. Time Series.....	23
3.1.6. White Noise	23
3.1.7. Stationary Time Series	23
3.1.8. Autocorrelation and Partial Autocorrelation	24

3.2.	Stochastic Differential Equation (SDE) Models	24
3.2.1.	Geometric Wiener Process	25
3.2.2.	Constant Elasticity of Variance (CEV) models	25
3.2.3.	Linear Drift-Rate Models	25
3.3.	Time Series Models	25
3.3.1.	AR Model	26
3.3.2.	MA Model	26
3.3.3.	ARMA Model.....	27
3.3.4.	ARIMA Model	28
3.3.5.	ARCH Model.....	28
3.3.6.	GARCH Model.....	29
3.4.	Artificial Neural Networks (ANN) Models.....	29
3.4.1.	Models of a Neuron	33
3.4.2.	Types of Activation Function.....	34
3.4.3.	The Architecture of Neural Networks	36
3.4.4.	Learning Process	39
3.5.	Factors and Variables	43
3.5.1.	Natural Gas Production and Consumption	44
3.5.2.	Pipeline Capacity	44
3.5.3.	Storage of Natural Gas	45
3.5.4.	Substitutions	46
3.5.5.	Weather Effect.....	47
Chapter 4:	Results and Discussions	49

4.1.	Stochastics Differential Equation Models	50
4.1.1.	Geometric Brownian motion (GMB) Model	50
4.1.2.	Constant Elasticity of Variance (CEV) models	55
4.1.3.	Linear Drift-Rate Models	58
4.1.4.	Summary of SDE Models	62
4.2.	Time Series Models	62
4.2.1.	ARIMA/GARCH Model	63
4.3.	Neural Network Models	72
4.3.1.	Univariate Neural Network Autoregression (NAR) Model	72
4.3.2.	Multivariate Neural Network Autoregression (NARX) Model	79
4.3.3.	Sensitivity Analysis	85
4.4.	Result Summary	92
4.5.	Natural Gas Price Forecaster Package	93
Chapter 5:	Conclusions	97
References	99
Appendix A:	Nomenclature	107
Appendix B:	Software Package	110

List of Tables

Table 1. Summary of SDE models forecasting	62
Table 2. Result summary for stationary tests	65
Table 3. Akaike Information Criteria Results for Different Lags in ARIMA Model	67
Table 4. Bayesian Information Criteria Results for Different Lags in ARIMA Model .	67
Table 5. ARIMA/GARCH model's parameters.....	70
Table 6. The measured mean error in sensitivity analysis.....	91
Table 7. Fit results of models.	92
Table 8. Predicted natural gas price for different models.	93

List of Figures

Figure 1. The relationship between reservoir depth and temperature to the likely proportion of hydrocarbon produced (Chandra, 2006).....	3
Figure 2. US natural gas pipeline network (EIA, 2009).....	4
Figure 3. Production volume versus distance to market framework for gas technologies (Wood, Mokhatab, & Economides, 2008).....	6
Figure 4. World energy consumption 2013, Source: (BP, 2014)	7
Figure 5. Henry Hub Natural Gas Spot Price, (Source: EIA)	11
Figure 6. Natural gas supply and disposition in the United States, 2014 (EIA, 2014)...	12
Figure 7. Schematic of artificial neural network	32
Figure 8. Nonlinear model of a neuron (Haykin, 1999).....	34
Figure 9. Illustration of different activation functions.	35
Figure 10. Single layer neural network.	37
Figure 11. Multilayer neural network.....	38
Figure 12. Recurrent neural network.....	39
Figure 13. Normality test on gas price time series.	49
Figure 14. GBM model back test results (constant parameters)	51
Figure 15. The obtained solutions for 1, 6, 12, and 24 steps ahead by GBM model (constant parameters).....	52
Figure 16. Gas price prediction based on GBM model with constant statistical parameters (constant parameters).....	52
Figure 17. GBM model back test results (variable parameters).....	53

Figure 18. The obtained solutions for 1, 6, 12, and 24 steps ahead by GBM model (variable parameters)	54
Figure 19. Gas price prediction based on GBM model with constant statistical parameters (variable parameters)	54
Figure 20. CEV model back test results (constant parameters).....	55
Figure 21. The obtained solutions for 1, 6, 12, and 24 steps ahead by CEV model (constant parameters)	56
Figure 22. Gas price prediction based on CEV model with constant statistical parameters (constant parameters).....	56
Figure 23. CEV model back test results (variable parameters)	57
Figure 24. The obtained solutions for 1, 6, 12, and 24 steps ahead by CEV model (variable parameters)	57
Figure 25. Gas price prediction based on CEV model with constant statistical parameters (variable parameters)	58
Figure 26. LDR model back test results (constant parameters).....	59
Figure 27. The obtained solutions for 1, 6, 12, and 24 steps ahead by LDR model (constant parameters)	59
Figure 28. Gas price prediction based on LDR model with constant statistical parameters (constant parameters).....	60
Figure 29. LDR model back test results (variable parameters)	60
Figure 30. The obtained solutions for 1, 6, 12, and 24 steps ahead by LDR model (variable parameters)	61

Figure 31. Gas price prediction based on LDR model with constant statistical parameters (variable parameters)	61
Figure 32. Autocorrelation and Partial Autocorrelations functions for Henry Hub Spot Price	64
Figure 33. Natural Gas Spot Price (a) and First Difference of Price (b) Historical Data	64
Figure 34. Standard Deviation of Differenced Time Series	65
Figure 35. Autocorrelation and Partial Autocorrelations functions for First Difference of Spot Price.....	66
Figure 36. Variance and Mean of Time Series during Time	68
Figure 37. ACF and PACF of Squared First Difference of HH Spot Price.....	69
Figure 38. (a) Model Back Test Results, (b) Residual Value for Model.....	71
Figure 39. Model Forecasting for 12 Month in Future.....	71
Figure 40. Autocorrelation of error for 1-layer 1-lag NAR model.....	73
Figure 41. Autocorrelation of error for 1-layer 9-lag NAR model.....	73
Figure 42. Autocorrelation of error for 1-layer 10-lag NAR model.....	74
Figure 43. Performance of neural networks vs. number of neurons in the hidden layers for NAR model.	74
Figure 44. Back testing of the NAR model with 10 lags and 1 neuron in the hidden layer	75
Figure 45. Regression results and R squared for 10-1-1 NAR model.....	76
Figure 46. Performance progress for 10-1-1 NAR model.	77
Figure 47. Regression results and R squared for 10-6-1 NAR model.....	77
Figure 48. Performance progress for 10-6-1 NAR model.....	78

Figure 49. Back testing of the NAR model with 10 lags and 6 neurons the hidden layer	78
Figure 50. Autocorrelation for 1 neuron 1-lag NARX.....	79
Figure 51. Autocorrelation of error for 1-neuron 9-lag NARX model.....	80
Figure 52. Autocorrelation of error for 1-layer 9-lag NARX model.....	80
Figure 53. Performance of neural networks vs. no. of neurons in the hidden layer in NARX model.....	81
Figure 54. Regression results and R squared for 21-1-1 NARX model.....	82
Figure 55. Performance progress for 21-1-1 NARX model.....	82
Figure 56. Back testing of the NARX model with 9 lags and 1 neuron in the hidden layer.	83
Figure 57. Regression results and R squared for 21-4-1 NARX model.....	83
Figure 58. Performance progress for 21-4-1 NARX model.....	84
Figure 59. Back testing of the NARX model with 9 lags and 4 neurons in the hidden layer.	84
Figure 60. Sensitivity analysis for West Texas oil price.....	85
Figure 61. Sensitivity analysis for total production of natural gas.....	86
Figure 62. Sensitivity analysis for natural gas storage capacity.....	86
Figure 63. Sensitivity analysis for natural gas storage volume.....	87
Figure 64. Sensitivity analysis for natural gas storage injection.....	87
Figure 65. Sensitivity analysis for natural gas storage withdrawal.....	88
Figure 66. Sensitivity analysis for total consumption of natural gas.....	88
Figure 67. Sensitivity analysis for cooling degrees day Temperature.....	89

Figure 68. Sensitivity analysis for heating degrees day temperature.	89
Figure 69. Sensitivity analysis for extreme maximum temperature at New Orleans, LA.	90
Figure 70. Sensitivity analysis for extreme minimum temperature at New Orleans, LA.	90
Figure 71. Sensitivity analysis for mean temperature at New Orleans, LA.	91
Figure 72. Software package's main menu.	93
Figure 73. SDE model selection window.	94
Figure 74. ARIMA / GARCH model selection window	94
Figure 75. Neural networks model selection window	95
Figure 76. Output setting window for SDE model.	96
Figure 77. Output setting window for ARIMA model.	96
Figure 78. Output setting window for ANN.	96

Abstract

Prediction of the natural gas price is imperative to producers, suppliers, traders, market makers, and bankers involved in the natural gas exploration, production transportation, and trading. Additionally, consumers are also highly affected by the changes in the price of oil and gas products. Several attempts have been made to model the energy commodity prices over the past few decades. Stochastic differential equation, linear and nonlinear regression, auto regression, and neural networks are the main techniques that have been implemented.

In this thesis, three different categories of models are examined which are, stochastic differential equations, ARIMA, and autoregressive neural networks. The results indicate that, the NAR neural network provides a better fit to the given data as compared to the other proposed models. The three-layer NAR model with 6 hidden neurons was found to have the best performance in terms of one month ahead price prediction.

The accuracy of the NARX model with 6 neurons was found to be higher than that of the other models. Although, this model provides a reasonable fit to the given data, it fails to capture the price spikes effectively. The sensitivity analysis shows that CDD/HDD temperatures, extreme minimum temperature, and WTI oil prices have an insignificant effect on the results. On the other hand, total consumption, total production, and mean temperature of weather impact the results significantly.

Chapter 1: Introduction

Prediction of the natural gas price is imperative to producers, suppliers, traders, market makers, and bankers involved in the natural gas exploration, production, transportation and trading as well as consumers involved in the utilization of the natural gas (Mishra, 2012). The price of energy commodity, including natural gas price, is dramatically volatile that encounters the parties to a high risk and uncertain situations. More accurate forecasting helps them to select an appropriate strategy in order to reduce the uncertainty by means of hedging the risks. This research attempts to find the best way to model and forecast the natural gas prices among different approaches, and choose the optimum practice.

The natural gas price has also considerable effect on the evaluation of gas reserves (Caldwell & Heather, 1996), which is a major part of Gas Company's assets. Fluctuation in gas prices disturb the company's value and increase the investment risk which result in reduction of the company's stock price. Determination of the factors which control the natural gas market and price, as well as spread of information between the all parties in the market, stabilize the firm value and reduce the volatility.

Comparing to the other fossil fuels like coal or petroleum products, natural gas is less pollutant and more environment friendly. Since natural gas is the cleanest and the most abundant fossil fuel compared to other fossil resources in the world, by improving transportation technology and decreasing the handling costs, it is becoming the most popular source of energy globally (Conti, et al., 2015).

1.1. What is Natural Gas

Natural gas, also called marsh or swamp gas in older texts and more recently Landfill gas, is a combustible mixture of gaseous hydrocarbons consisting of methane as a major component and heavier hydrocarbons, including ethane, propane, butanes, and very small amount of pentanes or heavier components (Speight, 2007). The natural gas mixture also includes the unsaturated hydrocarbons such as aromatics, olefins, naphthenic, mercaptanes, and more complex hydrocarbons. The inorganic components are available in the natural gas mixture. The most important inorganic materials are nitrogen, carbon dioxides, hydrogen sulfide, and water that should be removed from the stream before injecting it into a pipeline (GPSA, 2013).

There are three main theories about the process of oil and gas formation underground; Thermogenic, biogenic, and abiogenic processes. In the most accepted theory, thermogenic process, the dead body of the organic species such as plants and animals buried into the deep waters by mud, soil, sands, rocks, and deposits. Over million years, the accumulation of the organic and inorganic materials made a heavy layer of these deposits with high pressure and temperature, which decomposed the complex organic components into the simpler one. Natural gas often discovered in the deeper reservoir as a single phase resource and it exists as an associated gas in shallower formation adjacent to the oil. The nature of organic debris also affects the product of this transformation (Bahadori, 2014). Figure 1 illustrates the relationship between the depth of the formation and the possibility of existence of the formed hydrocarbon.

Even though natural gas was discovered in the 17th century, storage and transportation were the main issues. It was used as a lightener in the streets of Baltimore,

MD in 1816. To serve all the continent, natural gas pipeline grid has been developed by pipeline operators until all the country was in the coverage of the natural gas supply.

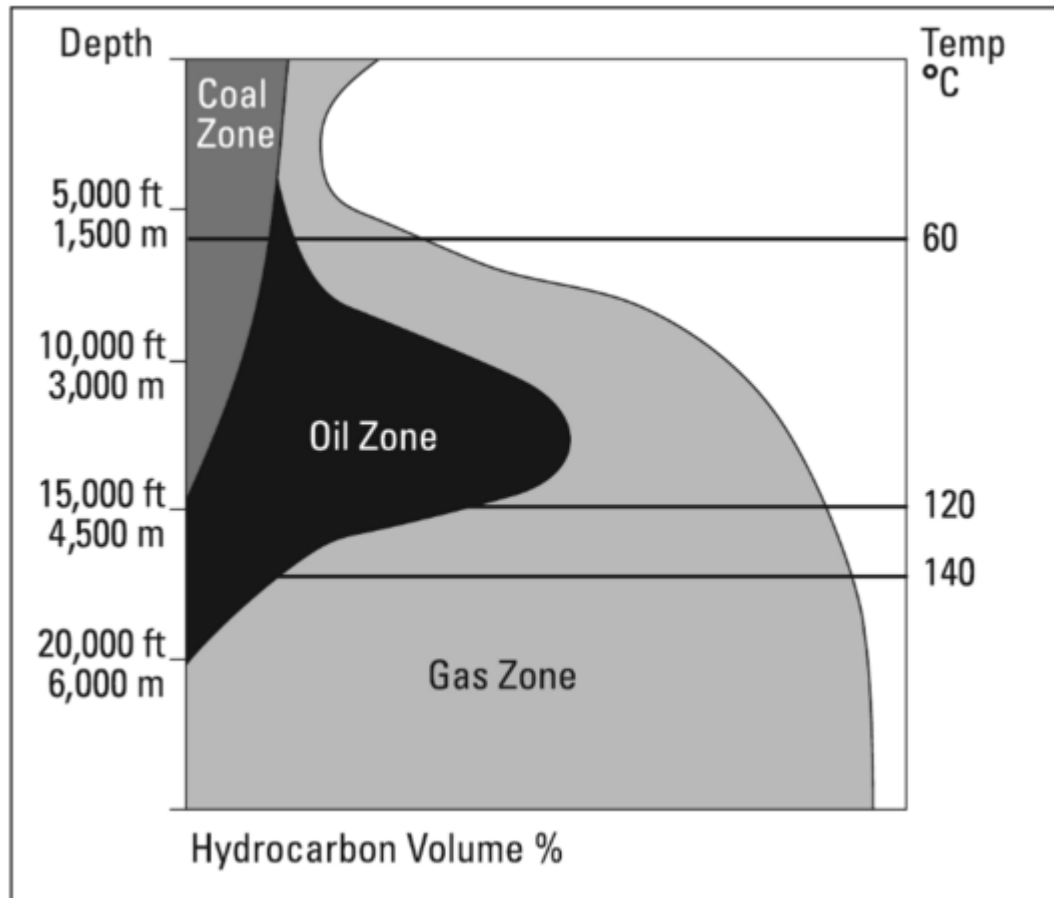


Figure 1. The relationship between reservoir depth and temperature to the likely proportion of hydrocarbon produced (Chandra, 2006)

Now, the most part of the consuming natural gas in the United States is produced in the Rockies, Gulf of Mexico (inshore and offshore), Western Canada, and Midcontinent including Texas, Oklahoma, Nebraska, and Arkansas. The major part of the natural gas is transferred by the interstate pipeline network which connect the producers

and consumers all around the country. Figure 2 shows the current national network of natural gas pipeline in the United States.

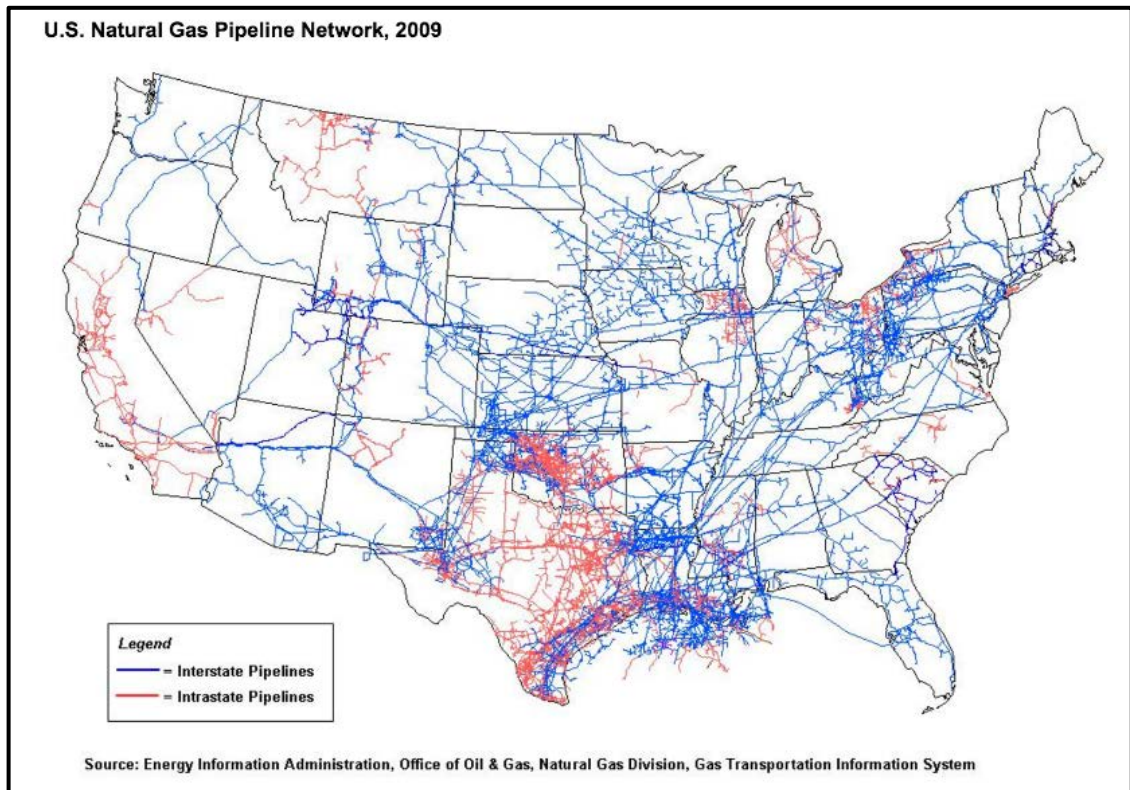


Figure 2. US natural gas pipeline network (EIA, 2009)

1.2. Forms of Natural Gas

For several decades, natural gas was vented as a dangerous byproduct of the oil wells. Most of the petroleum fluids in the reservoir condition contain various amounts of dissolved gas, which is liberated in the stock tanks. Since dissolved natural gas causes several problems in oil transportation, it should be separated in order to reduce the transportation cost.

Since storage and transferring of natural gas in gaseous phase was not feasible for a long time, it used to be consumed locally near to the production fields. By development of pipelines and new transportation technologies, such as LNG and CNG, natural gas can be conveniently delivered to the markets in the remote areas across the world (Mishra, 2012).

Natural gas is stored in different forms of underground storage, which are not available in every part of the country and often are located close to the production areas. Depleted reservoirs, aquifers, and salt caverns are the underground places that could be used as reliable storage for natural gas (Wang & Economides, 2013).

A very famous form of natural gas has been emerging since the mid-70s is Liquefied Natural Gas (LNG), which is a liquid form of the natural gas that cooled by cryogenic technology to approximately -260°F (-162°C). LNG technology implements complex facilities and equipment in production, storage, and transportation sides which are expensive and hard to operate (Mokhatab, Mak, Valappil, & Wood, 2013). The good news is that LNG technology is improving in ease of handling and lowering the cost dramatically by time. The share of LNG in the transportation of natural gas is small but it is increasing.

Compressed Natural Gas (CNG) is the other form of natural gas that is being applied to handle and transport the natural gas as a carrier of energy. CNG is the form of natural gas, which is compressed to high pressure around 4000 psi. It needs to be carried by special container and ships (Mokhatab & Poe, 2012).

Hydrate is another form of natural gas, which is an ice-like crystal formed by existence of liquid water in high pressure and low temperature conditions. Natural gas

hydrate is abundant on the ocean's bed. Collecting of hydrate and extraction of gas from the solid hydrate is not economic yet, but it is believed to be an alternative to LNG or pipelines in the future (Mokhatab & Poe, 2012).

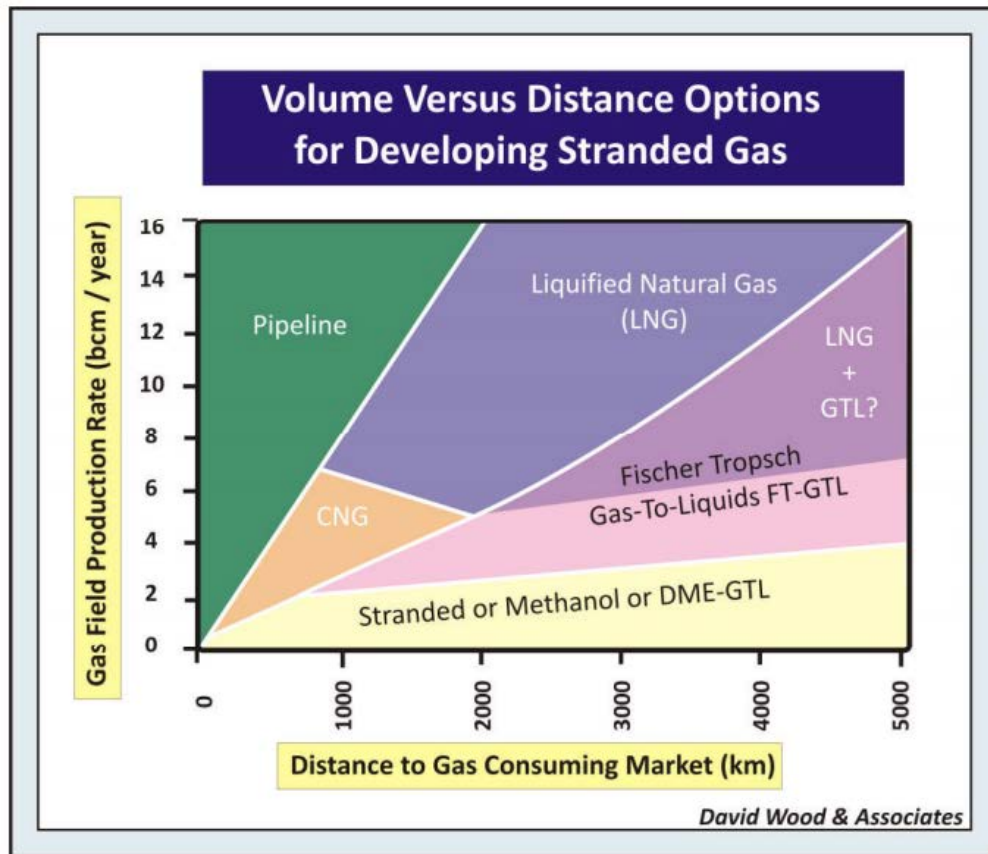


Figure 3. Production volume versus distance to market framework for gas technologies (Wood, Mokhatab, & Economides, 2008)

Gas to Liquid (GTL) is a process in which natural gas in the presence of a catalyst converts into a heavier hydrocarbon mixture such as methanol and ammonia with lower volume and almost the same content of energy (Anderson, Kölbl, & Rálek, 1984). This

liquid form could be transferred easier than the gaseous form of the natural gas. This technology is not new, and was introduced seventy years ago (Schulz, 1999).

Figure 3 shows a rule of thumb for application of different technologies and various forms of the natural gas in transportation based on the destination distance and the volume of the transferring gas. The pipeline is the best choice for transmission of huge amounts of natural gas in short distances. For long distance, especially oversea transportation with large amount of gas, LNG is a good option.

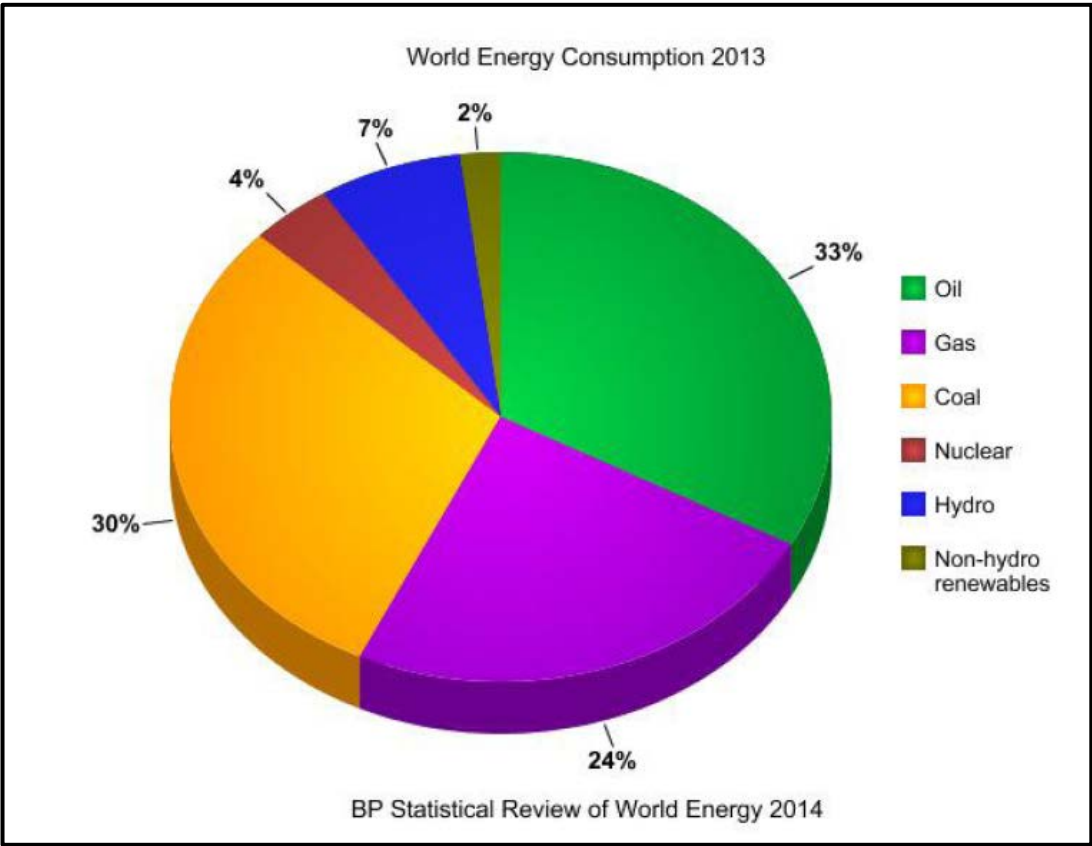


Figure 4. World energy consumption 2013, Source: (BP, 2014)

The production of shale gas is dramatically increasing. By increasing the shale gas production in the U.S., oil import will decline and enough gas will be available to

replace the coal in power generation industry (Panella, Barcellona, & D'ecclesia, 2012). New technologies and explorations enabled us to produce more natural gas from tight and unconventional reservoirs. Unconventional reservoirs such as shale have a considerable share of reserve of natural gas in the North America. Hydraulic fracturing in combination of horizontal drilling was greatly developed in the 70s and 80s to produce hydrocarbon from tight formations. Moreover, the rapid change in the economy of China needs a large amount of energy (U.S. EIA | IEO, 2013). In addition, the global energy demand is increasing as well as the portion of the demand for natural gas in the total energy consumption (BP, 2014).

1.3. Regulations on Natural Gas Market

The history of natural gas regulation goes back to the history of this industry. In mid-1800s, natural gas was produced from coal mines and consumed locally in municipal areas. In order to reduce the cost of operations, the production and distribution of product were performed by a single company. Therefore; the regulation on price was set by government to prevent a monopoly. By expanding the pipelines in the 1900s and beginning of interstate transportation, local governments invented a new interstate natural gas market and determined the rate that could be charged by distributors.

In 1938, Federal Government intervened for the first time in the interstate natural gas market and passed the Natural Gas Act (NGA) to give jurisdiction to Federal Power Commission (FPC) over regulation of the interstate natural gas sales to control every feature from wellhead to burner tip (IEA, 1995). The wellhead price was regulated in 1940 in terms of production cost. After 1960 the geographical area became the base for

rate regulation. Since this regulation was not successful, FPC has determined national price ceiling for interstate sales in 1974. The determined price was below the market value of the commodity that made producers reluctant to explore and drill new reserves. The ceiling price was set for wellhead but interstate sales were almost free which caused some problems with consuming states (MacAvoy & Pindyck, 1974).

The regulation of natural gas has dramatically changed during 30 years past. In November of 1978, Natural Gas Policy Act (NGPA) as part of the National Energy Act (NEA), has passed (Richardson & Nordhaus, 1995) by Congress to reach the following objectives:

1. Creating a single national natural gas market
2. Equalizing supply with demand
3. Allowing market forces to establish the wellhead price of natural gas

New regulatory body, the Federal Energy Regulatory Commission (FERC), has been created to control the interstate natural gas market as well as deregulation of natural gas prices at the wellhead which resulted a significant amount of drilling rigs in early 80s. The main objective of this act was balancing the market. Power Plant and Industrial Fuel Use Act (FUA) at 1978 which, restricted the usage of natural gas in power plant to substitute other fuels such as coal and oil products on one hand, and the increment in production on the other hand leads a huge over-supply of natural gas. On the supply side, they encouraged the producers to explore and develop new fields, and on the demand side, they encouraged the consumers to consume other fuels. These actions together led the market into having a large excessive amount of gas supply (Texas, 1976).

Natural Gas Well Head Decontrol Act of 1989 (NGWDA) deregulated the first sales of natural gas determined by FERC (Huston, 2015). It was the beginning of deregulation in the natural gas market, which finalized by Order 636 in 1993 (Reiter & Economides, 1999).

Finally, under the current regulations, only pipelines and Local Distribution Companies (LDCs) are directly regulated by FERC. Producers and marketers are controlled by authorization other than FERC while the price is determined by supply and demand law in a competitive market. Pipeline operators are obligated to accept natural gas into their infrastructures if it meets the technical specifications.

1.4. Natural Gas Market

Deregulation of price in the gas market, made a competitive market in which producers, consumers, and market makers could find the commodity at a fair price. This liberation was also a good platform for derivatives market. In the early 1990s, natural gas financial markets started to operate when the regulation in the U.S. was adjusted in order to liberate the market, as well as national grid for natural gas transportation was expanded. All of these evolutions prepared a reliable infrastructure for natural gas trading and availability all around the country. New financial markets for energy commodity, including financial derivatives like future contracts and options, facilitates the investors and traders to hedge their investment. In April, 1990 New York Mercantile Exchange (NYMEX) was started to trade natural gas futures contracts with Henry Hub in Louisiana as a delivery location.

The price of energy in general and the price of natural gas in particular, are very volatile. For the investors who plan to build million-dollar power plants, the incoming cash flow involves them with a high degree of uncertainty because the fuel price flocculation brings instability in revenues. The price oscillation in short time denotes the risk of this commodity which makes the investors worried about the fuel price as a major part of the operation costs. Figure 5 demonstrates the natural gas price changes since 1997 which, changes widely during the time.

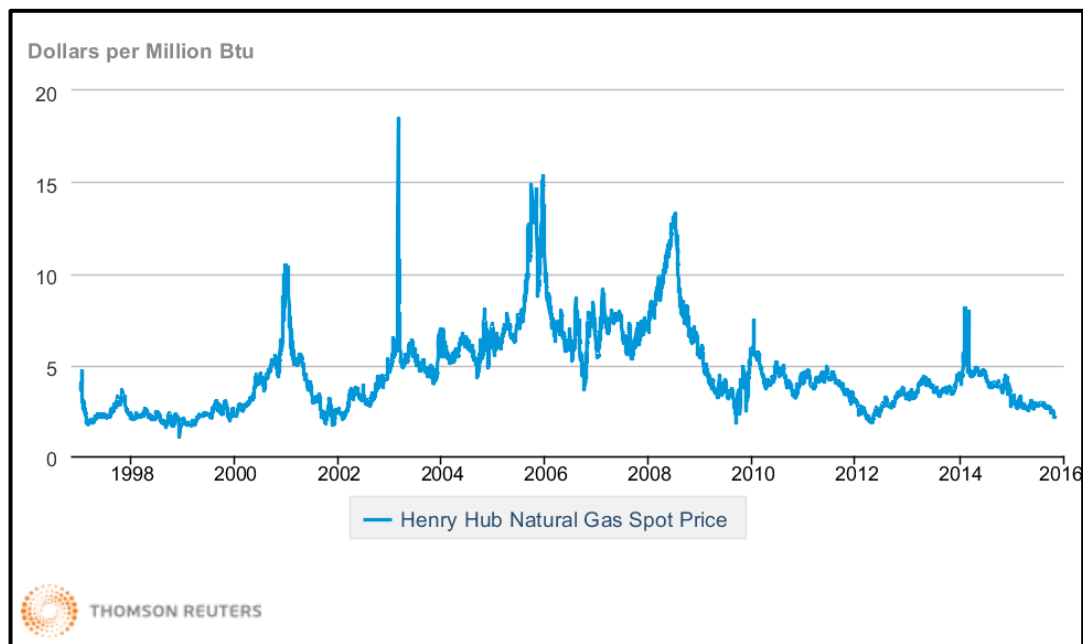


Figure 5. Henry Hub Natural Gas Spot Price, (Source: EIA)

Natural gas markets were considered as domestically isolated markets in the past, but by expanding the pipelines around the world and developing LNG technology, it could be transferred easier than before (Mishra, 2012). It fills the price gap between

different regions and converges the price globally. Figure 6 illustrates the flow of natural gas in the United States' economy.

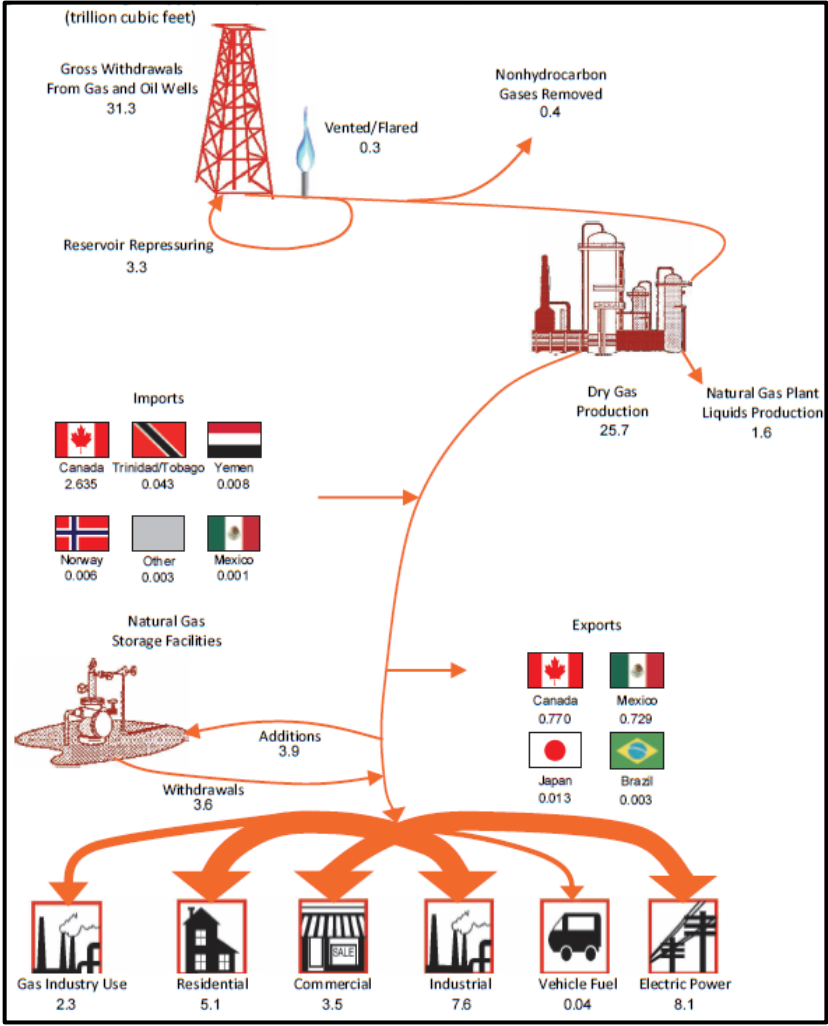


Figure 6. Natural gas supply and disposition in the United States, 2014 (EIA, 2014)

1.5. Chapters and Contexts

This research is conducted to examine various models to describe, fit, and predict the natural gas price in the United States market based on the historical spot price at Henry Hub, and other related parameters such as meteorological data, economic data etc. The

objective is to find the simplest and the best fitted model to predict the natural gas price, which is very important to investors, producers, and consumers to hedge and manage the risks associated with their economic and financial activities. This thesis contains 5 main chapters. A software package developed by MATLAB® is attached to this booklet.

Chapter 2: This chapter presents a survey of the most relevant studies reported in the literature about the prediction of the energy prices. The studies are critically reviewed and cross-compared in order to determine a) the most influential factors for the price of energy in general and natural gas in particular, and b) the methods that have successfully been used for the prediction of the energy price. The survey highlights the potential and capabilities of ANNs for similar applications and justifies their use for the gas price prediction.

Chapter 3: In this chapter three categories of models are presented; Stochastic differential equation models, time series models, and neural networks models. The first part defines some basic concepts which are used in our models. Then models are defined one by one. Finally, the neural network model is described in more details as a main model for this research. In this chapter the neurons and linkage between them is explained and some description about the architecture, activation functions, and learning algorithm have been given.

Chapter 4: The methods, variables, sources, data gathering, tools applied, and method validation are described in details. In this chapter, one of the most important part of this chapter is validation and evaluation of the applied methods. A part of data is kept for testing the solved neural network. The results for models which are presented in this chapter, are comparable because of using the same data set.

Chapter 5: A summary of the results is presented and recommendations for further studies are provided.

Chapter 2: Literature review

Hotelling (1931) could be recognized as the first researcher who introduced a model to describe the behavior of exhaustible resources, such as natural gas. This model is known as the Hotelling's rule; which states that the producers of a non-renewable commodity (e.g. Natural gas) tend to sell their commodities when the benefits from the sale are more than the benefits of keeping it. In other words, the extraction cost and the present value of the commodity in one side, and the cost of the storage and future value of the commodity in other side, which are related by interest rate, will determine the proficiency of keeping or selling the commodity. It is assumed that the markets are efficient and the owners are motivated by the profit. This rule does not consider the emergence of new technologies and resources that may be explored and discovered in the future. Pindyck (1978) has optimized Hotelling's model for oil and gas case by taking the oil and gas reserve increment by exploration into account. An MIT Energy Lab report (MacAvoy & Pindyck, 1974) revealed the problem of natural gas shortages with econometrics models; which noted that the ceiling price was set by the Federal Power Commission and did not represent the price of the supply and demand equilibrium.

Linear models considers a linear correlation between independent and dependent variables. They assume that the same stimulus in a system will result the same response. Observations are used to forecast the future outcomes in definite conditions. The most linear models have poor performance in prediction due to the complex behavior of the system which, could be recognized as nonlinearity (Agbon & Araque, 2003).

In the past, the linear models were very common to use predicting the oil and gas prices. A pessimistic case, most likely case, and an optimistic case were reported by a

sensitivity analysis (Powers & Stevenson, 1987). Although the relationship between natural gas and oil price is not a linear correlation some simple and linear correlation have been developed which are useful as a rule of thumb (Brown & Yücel, 2008). Some models were established based on price elasticity, economic growth effect, and energy intensity (Roberts, 1989). Price elasticity is the ratio of the percentage change in quantity demanded and the percentage change in the commodity price.

Bopp (1990) and Hsieh (1990) applied econometrics model, and Pilipovic (2007) implemented time series models to predict natural gas prices. Reiter and Economides (1999) applied both multivariate econometrics and neural network models to forecast the natural gas price in the short term. Inikori, et al. (2001) investigated the effect of oil and natural gas prices on the oil and gas industry by establishing a linear regression model, which forecasted said prices. They considered price elasticity and supply demand balances as input variables.

Two forecasting models have been developed by Nogales, et al. (2002) to predict the daily price of natural gas. They have used the time series analysis approach to establish dynamic regression and transfer function for Spain and California Market, which resulted in the average errors of 5% and 3% respectively. Agbon and Araque (2003) applied chaos time series analysis and fuzzy neural network model with a nonlinear model to predict the oil and gas prices. Ogwo, et al. (2007) developed an equitable pricing model to predict the natural gas price. Mishra (2012) modeled the natural gas price with time series as well as a nonparametric approach to forecast the price. Hu and Trafalis (2011) developed a new kernel for a neural network model (vector support machine) to predict the natural gas price. Panella, et al. (2012) suggested a new

approach to implement the neural network to model a nonlinear regression time series, associated with energy commodity prices. Yi & Wang (2013) addressed the effect of oil price on the international price of natural gas using wavelet based Boltzmann cooperative neural network. Ekweanua, et al (2014) found a robust correlation between the natural gas price and production, import, and export amount by neural network model.

To predict natural gas price in long term or short term, several approaches and models have been proposed and applied by different researchers. Each method and model has its own advantages and disadvantages. The multivariate models which take several variables into account are more accurate than the univariate models. However the external variables often need to be predicted themselves. For example, a predictor model for natural gas price which is a function of oil price is struggling with the same uncertainty in oil price. It is not easy to predict the variable in more than one time step in the future.

Crude oil price has a stochastic nature that most of the time is normally distributed and sometimes it behaves like a nonlinear manner (Caldwell & Heather, 1996). Crude oil price is not just a function of supply and demand, but depends on more variables such as changes in technology, culture, ultimate resource base, consumption patterns and population growth (Skov, 1995). This complexity shows the non-linear nature of the crude oil price. The oil supply has a dominant role in oil price modeling. Small changes in supply made a large change in the oil price illustrating a strong relationship between the supply and price elasticity, which may not be linear (Dougherty, 1987).

The energy commodity prices behave a stochastic process (Schwartz, 1997). Several attempts have been performed to capture the stochastic nature of the energy price movement, by single-factor models (Lucia & Schwartz., 2002) and (Barlow M. T., 2002).

Agbon and Araque examined the spot oil and gas price and concluded that are likely chaotic and fractal (Agbon & Araque, 2003). Barlow et al (2004) examined three different multi-factor stochastic models describing by stochastic differential equations. They used Kalman Filters to calibrate the Models' factors. They concluded that the log spot price mean-reverting to generalized Wiener process and Pilipovic model, are applicable to model electricity spot prices (Barlow, Gusev, & Lai, 2004). A research on Canadian oil and gas price claims that the long term prices tend to revert back to the long term average, while short term prices were not predictable (Morgan, Mikalson, & Herchen, 2012).

High spikes and oscillations on price trends assert that the economic system is very complex and is affected by a large number of variables. The complexity is due to initial conditions, the system paths, and the final conditions which lead us to the non-linear models (Agbon & Araque, 2003). Polynomial, and Gaussian function curve fittings are a traditional method to establish a nonlinear model. Even though these solutions may not model the spike prices but offer a reasonable explanation for the most conditions (Agbon & Araque, 2003).

The neural network has been implemented in several areas of science (Qian & Sejnowski, 1988), social science (Agarwal, Saferpour, & H., 2014), engineering (Hoskins & Himmelblau, 1988), and finance (Trippi & Turban, 1992). The application of neural network in petroleum engineering is increased in different subjects such as production (Al-Fattah & Startzman), well testing (Al-Kaabi & Lee, 1993), and phase behavior (Habiballah, Startzman, & Barrufet, 1996). Reiter and Economides (1999) claim that short term prediction of gas price is feasible by using lagged variables and neural network

models. Hu (2011) built a support vector regression to model the non-linear behavior of gas price. She took advantage of neural networks to solve the models.

A study has been conducted to evaluate different models including linear regression, time series models, and neural network based models, with and without explanatory variable, to predict the electricity price in the short term. They found that neural network based models are showing better results with an accuracy of 0.5 to 47%, but there were not enough evidence to prove it (Yıldırım, Bayrak, & Weber, 2014).

Chapter 3: Methodology

All the data, related to prices, consumptions, reserves, storages, productions, pipelines etc., are collected from the U.S. Energy Information Administration's website (EIA, US Energy Information Administration, 2016) which are accessible in Microsoft Excel® format online for free. Climate data including cooling/heating degree days, extreme maximum/minimum temperature, mean temperature, and mean maximum/minimum temperature for New Orleans, LA, were downloaded from the National Centers for Environmental Information department of the National Oceanic and Atmospheric Administration (NOAA, 2016). All these data were combined together in a single Microsoft Excel® file as an input data file.

The data file was loaded into MATLAB® environment which is the main platform to analyze the data in this research. The Econometric Toolbox™, Financial Toolbox™, and Neural Network Toolbox™ are the most important toolboxes that were used to simulate and analyze the given data. Codes and routines have been developed for analysis, which are available as a software package attached to this thesis. A simple graphical user interface (GUI) has been developed to utilize the written code easily.

In this chapter a brief explanation is proposed about the key concepts and definitions which are used in the implemented models. After basic definitions, models are defined in details.

3.1. Basic Definitions

The objective of this investigation is to find a model, fit, and predict the natural gas spot price at the Henry Hub. Natural gas price is a random variable which follows a

stochastic process with a random trend. Initially, few important definitions which are useful and applicable in the modeling process are presented. Then the brief statement of the models has been introduced.

3.1.1. *Random Variables*

A random variable is a measurable function of X from the probability space Ω into the set of real numbers \mathbb{R} known as the state space. Three modifications are needed to make this definition more precise (Gallager, 2013):

- (i) The mapping $X(\omega)$ must have the property that $\{\omega \in \Omega : X(\omega) \leq x\}$ is an event for each $x \in \mathbb{R}$.
- (ii) Every finite set of random variables X_1, \dots, X_n has the property that for each $x_1 \in \mathbb{R}, \dots, x_n \in \mathbb{R}$, the set $\{\omega \in \Omega : X_1(\omega) \leq x_1, \dots, X_n(\omega) \leq x_n\}$ is an event.
- (iii) X might be undefined or infinite for a subset of Ω that has 0 probability. In other words, the probability of events $\{X = \pm\infty\}$ is zero.

3.1.2. *Stochastic Process*

In probability theory, a stochastic process is a family of random variables X from the probability space Ω into \mathbb{R} indexed by time set $t \in \mathbb{T}$, which could be denoted by $\{X(t, \omega) : t \in \mathbb{T}, \omega \in \Omega\}$. We simplify the notation to $\{X(t) : t \in \mathbb{T}\}$ or $\{X_t : t \in \mathbb{T}\}$ when the time variable t is continuous ($\mathbb{T} = \mathbb{R}$) or is a discrete variable ($\mathbb{T} = \mathbb{N}$) respectively (Prabhu, 2007). In a stochastic or random process, we always, encountered to some

uncertainty: even if the initial condition of the process is fixed, there are several directions, named a sample path, in which the process may proceed.

3.1.3. Monte Carlo Simulation

Monte Carlo methods are a wide range of computational algorithms that rely on the repeated random sampling or output of a random-base function to obtain the best numerical results. In a stochastic process or time series simulation, the functions generate different results in different iterations. They are often used in physical, mathematical and statistical problems and are most useful when it is difficult or impossible to use the other mathematical methods. These models contain a white noise term in their definition. Different random variables lead us to the different values and possible paths. For these types of modeling it is possible to run the model for several times, which are generated from various and independent white noises. Every run results its specific solution. Monte Carlo simulation is a useful tool to analyze these outputs statistically (Glasserman, 2003).

3.1.4. Wiener Process

Weiner process or Brownian motion W_t is a continuous-time stochastic process over time t , which is used to model the integral of a white noise Gaussian process. The process is characterized by the following properties:

- (i) $W_0 = 0$ with probability of 1.
- (ii) W_t is normally distributed for all $t \geq 0$ with mean 0 and variance t .
- (iii) W_t has independent increment: dW_t for any time is independent.

- (iv) W_t has continuous paths, meaning that any realization of W_t is a continuous function of t .

3.1.5. Time Series

The time series is a stochastic process of random variable X indexed by time. X_t is a notation used for discrete parameter process and $X(t)$ is a notation for continuous parameter process. In this research, we are dealing with the discrete parameter process of time series.

3.1.6. White Noise

A time series is called discrete white noise if

- (i) The X_t 's are identically distributed
- (ii) $\gamma(t_1, t_2) = 0$ when $t_1 \neq t_2$
- (iii) $\gamma(t, t) = \sigma^2$, where $0 < \sigma^2 < \infty$

Where auto-covariance function γ defined as;

$$\gamma(t_1, t_2) = E\{[X(t_1) - \mu(t_1)][X(t_2) - \mu(t_2)]\} \quad \text{Eq. 1}$$

3.1.7. Stationary Time Series

A time series is stationary if its statistical parameters such as mean and variance remain constant for all the time steps. The Engle test is one of the standard procedure to determine stationarity of a time series.

3.1.8. Autocorrelation and Partial Autocorrelation

Autocorrelation function is a correlation of time series with itself. This function measures the correlation between the variable X_t and its lag X_{t-k} which is a real value between -1 and 1 where -1 implies a complete negative correlation and 1 is a complete positive correlation while 0 indicates no correlation between the variables. In a similar way, partial autocorrelation is a correlation between X_t and X_{t-k} when X_{t-k} comes into the picture and improves the correlation to X_t . On the other hands, partial autocorrelation of a variable in k order is the amount of correlation between the variable and its kth lag that was not explained by the correlations at all lower orders lags.

3.2. Stochastic Differential Equation (SDE) Models

Stochastic Differential Equation (SDE) is a differential equation with one or more stochastic variable. It could be defined based on a Wiener process in general form:

$$dS_t = F(t, S_t)dt + G(t, S_t)dW_t \quad \text{Eq. 2}$$

where F is drift function, and G is diffusion function. Changes on these functions with different definitions result different models as described below. If S_t was defined as the rate of return and X_t as price, we could rewrite the Eq. 2 in form of:

$$\frac{dX_t}{X_t} = F(t, S_t)dt + G(t, S_t)dW_t \quad \text{Eq. 3}$$

3.2.1. Geometric Wiener Process

Geometric Brownian motion or exponential Brownian motion is a continuous-time stochastic process which is satisfying the following stochastic differential equation:

$$dX_t = \mu(t)X_t dt + \sigma(t)X_t dW_t \quad \text{Eq. 4}$$

where μ is drift parameter, σ is volatility parameter, and W_t is a Wiener process or Brownian motion.

3.2.2. Constant Elasticity of Variance (CEV) models

CEV models are a special case of drift-rated models which defined as:

$$dX_t = \mu(t)X_t dt + D(t, X_t^\alpha)\sigma(t)dW_t \quad \text{Eq. 5}$$

3.2.3. Linear Drift-Rate Models

This model is described based on following SDE:

$$dX_t = (A(t) + B(t)X_t)dt + D(t, X_t^\alpha)\sigma(t)dW_t \quad \text{Eq. 6}$$

3.3. Time Series Models

Time series models represent different stochastic processes and experience many varieties such as the autoregressive (AR) models, the integrated (I) models, and the moving average (MA) models, which are dependent linearly on their previous data points. Combinations of these models produce new models as autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA). These classes also could be merged with vector-valued data as multivariate time-series models may

called as VAR for vector auto-regression. An extra set of references of these models is available for use where the observed variable is affected by exogenous variables where the acronyms are terminated with an "X" for exogenous term.

Among other types of non-linear time series models, there are some models to denote the changes of variance over time. These models represent autoregressive conditional heteroscedasticity (ARCH) and the group of a wide variety of models (GARCH, etc.). Here changes in variability are related to, or predicted by, recent past values of the observed series.

3.3.1. AR Model

AR (r) or Autoregressive model of order r is a model which is relating a time series variable into the r past data

$$\begin{aligned}
 X_t - \phi_0 - \phi_1 X_{t-1} - \dots - \phi_r X_{t-r} &= a_t \\
 X_t - \phi_0 - \phi_1 B^1 X_t - \dots - \phi_r B^r X_t &= a_t
 \end{aligned}
 \tag{Eq. 7}$$

$$\phi(B)X_t = \phi_0 + a_t$$

where a_t is white noise and μ is mean of X_t . Another expression for AR (r) model is:

$$X_t = \phi_0 + \sum_{i=1}^r \phi_i X_{t-i} + a_t
 \tag{Eq. 8}$$

3.3.2. MA Model

Moving Average model of order m, MA (m), is generating realization from

$$\begin{aligned}
X_t - \mu &= a_t - \theta_1 a_{t-1} - \dots - \theta_m a_{t-m} \\
X_t - \mu &= a_t - \theta_1 B^1 a_t - \dots - \theta_m B^m a_t \\
X_t - \mu &= \boldsymbol{\theta}(B) a_t
\end{aligned}
\tag{Eq. 9}$$

Another expression for MA (m) model is

$$X_t = \mu - \sum_{j=1}^m \theta_j a_{t-j}
\tag{Eq. 10}$$

3.3.3. ARMA Model

Autoregressive Moving Average process, ARMA (r, m), is a combination of Moving Average MA (m) and Autoregressive AR (r) processes. Suppose $\{X_t; t = \pm 1, \pm 2, \dots\}$ is a causal, stationary, and invertible process. Therefore, it satisfies the following equation:

$$\begin{aligned}
X_t - \mu - \phi_1 (X_{t-1} - \mu) - \dots - \phi_r (X_{t-r} - \mu) \\
&= a_t - \theta_1 a_{t-1} - \dots - \theta_m a_{t-m} \\
X_t - \mu - \phi_1 B^1 (X_t - \mu) - \dots - \phi_r B^r (X_t - \mu) \\
&= a_t - \theta_1 B^1 a_t - \dots - \theta_m B^m a_t \\
\boldsymbol{\phi}(B)(X_t - \mu) &= \boldsymbol{\theta}(B) a_t
\end{aligned}
\tag{Eq. 11}$$

Another expression for ARMA (r, m) model is

$$X_t - \mu - \sum_{i=1}^r \phi_i X_{t-i} = a_t - \sum_{j=1}^m \theta_j a_{t-j}
\tag{Eq. 12}$$

3.3.4. ARIMA Model

The Autoregressive Integrated Moving Average, ARIMA (r, d, m), process of orders r, d, and m is a process, X_t , whose differences $(1 - B)^d X_t$ satisfy an ARMA (r, m) model that is a stationary model in which d is a non-negative integer. We use the following notation:

$$\phi(B)(1 - B)^d X_t = \theta(B)a_t \quad \text{Eq. 13}$$

where all the roots of $\phi(z) = 0$ and $\theta(z) = 0$ are outside of the unit circle, and $\phi(z)$ and $\theta(z)$ have no common factors. Parameter d in the ARIMA model represents the dth difference of X_t to find a stationary time series. Assume that X_t is not stationary, then we can reproduce a new time series with differencing the original time series such as $X_t - X_{t-1}$ for d = 1 and $(X_t - X_{t-1}) - (X_{t-1} - X_{t-2})$ for d = 2 and so on until a stationary time series is obtained.

ARIMA (0,1,0) is the famous random walk model as follows:

$$X_t = \mu + X_{t-1} \quad \text{Eq. 14}$$

3.3.5. ARCH Model

To address the conditional volatility behavior, (Engle, 1982) introduced the Autoregressive Conditional Heteroscedasticity (ARCH) model. ARCH (q) is defined based on an ARMA model in which the term a_t is a function of conditional variance. Let $\sigma_{t|t-1}^2$ be the conditional variance of X_t , and suppose

$$\begin{aligned} a_t &= \sigma_{t|t-1} \varepsilon_t \\ \sigma_{t|t-1}^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_q a_{t-q}^2 \end{aligned} \quad \text{Eq. 15}$$

In this equation, $0 \leq \alpha_1 < 1$ and ε_t 's are independent, identically distributed, zero mean and unit variance random variables that are independent of a_{t-k} , $k = 1, 2, \dots$. We can formulate the ARCH (q) model as follows:

$$\begin{cases} a_t = \sigma_{t|t-1} \cdot \varepsilon_t \\ \sigma_{t|t-1}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i a_{t-i}^2 \end{cases} \quad \text{Eq. 16}$$

3.3.6. GARCH Model

GARCH (p, q) is a generalized ARCH (p, q) model introduced by (Bollerslev, 1986) and (Taylor, 2007) which includes the lagged terms of the conditional variances. This model is defined as

$$\sigma_{t|t-1}^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_q a_{t-q}^2 + \beta_1 \sigma_{t-1|t-2}^2 + \dots + \beta_p \sigma_{t-p|t-p-1}^2 \quad \text{Eq. 17}$$

The compact form of GARCH (p, q) is

$$\begin{cases} a_t = \sigma_{t|t-1} \cdot \varepsilon_t \\ \sigma_{t|t-1}^2 = \alpha_0 + \sum_{j=1}^p \beta_j \sigma_{t-j|t-j-1}^2 + \sum_{i=1}^q \alpha_i a_{t-i}^2 \end{cases} \quad \text{Eq. 18}$$

3.4. Artificial Neural Networks (ANN) Models

Natural gas price prediction is important as it helps to have a better picture of the market in the future and enables us to monitor the factors that might affect the price movement. There are several technics to model and predict the natural gas price trend. Since the natural gas price change is random, the stochastic process could explain the nature of this oscillation. In addition, the periodic fluctuations of the natural gas price

introduce some other variables which can describe these vacillations. Another method that could explain this random trend is the time series techniques. The price could be modeled based on various auto-regression methods. Artificial neural network came into the picture when computers became more popular. Using fast computers enables us to run and train a neural network with high speed and accuracy.

We could take advantage of neural networks because of their properties, which help us to build a strong and robust model. The **ability to learn** and **generalization** are two important features of neural networks. Neural networks are able to solve complex problems by dividing them into simple tasks. An artificial neural networks consist of activation function that could be linear or nonlinear. By choosing the various nonlinear functions and different arrangement of the neurons and layers, nonlinear response to the stimuli would be captured. **Nonlinearity** handling is one of the most useful features of the neural networks which helps us to model complex functions (Haykin, 1999).

In neural network design, the networks learn to decrease the difference between the response and actual answers. There is an **input-output mapping** between the inputs and responses based on model-free estimation. After training the neural networks and defining the synaptic weight parameters, there is a single response for each specific inputs (Haykin, 1999).

Adaptivity is an important feature of the neural networks. The trained neural networks are able to modify the synaptic weight for new examples. They could be trained for new situation and reevaluate the responses comparing to the real answers. This feature is very useful for the problems that deal with online and dynamic data (Haykin, 1999).

In machine learning and data analysis, an Artificial Neural Network (ANN) is a family of models and data structures that simulate the operation of the human brain and nervous system. Artificial neural network is a branch of artificial intelligence emerged since the 1940s. (Hecht-Nielsen, 1989) proposed a good definition of an artificial neural network: *"... a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs."*

ANNs are generally described as a set of interconnected neurons grouped in layers which transfer signals between each other. The connections have numerical weights that can be adjusted based on feedback or comparison to a reference data set that make it adaptable to inputs and capable of learning.

Layers are key concepts in neural network architecture which, are made of a set of interconnected node named neurons containing an output function named activation function. The linkage between neurons are called synapsis which contained numerical weights. Independent variables as input data enter into the network via first layer called input layer, and dependent variables as results are obtained from the last layer named output layer. As shown in Figure 7, there are one or more layers, between input and output layers, as hidden layers. Each node of a layer is connected to the all nodes of the next layer. Each input to a layer is multiplied by a weight, then will be combined together with reference to a threshold and activation function and use them to determine the layer's outputs. The output could be a final result or an input to the next layer. The weights are initialized by random numbers and then trained by a large number of available data to get an accurate response.

Different learning algorithms are designed to train neural networks, which modify the weights of the synapsis to achieve the best fit to the inputs. Backpropagation is one the most famous types of neural networks, which performs a gradient descent within the solution's vector space towards a global minimum along the fastest vector of the error surface.

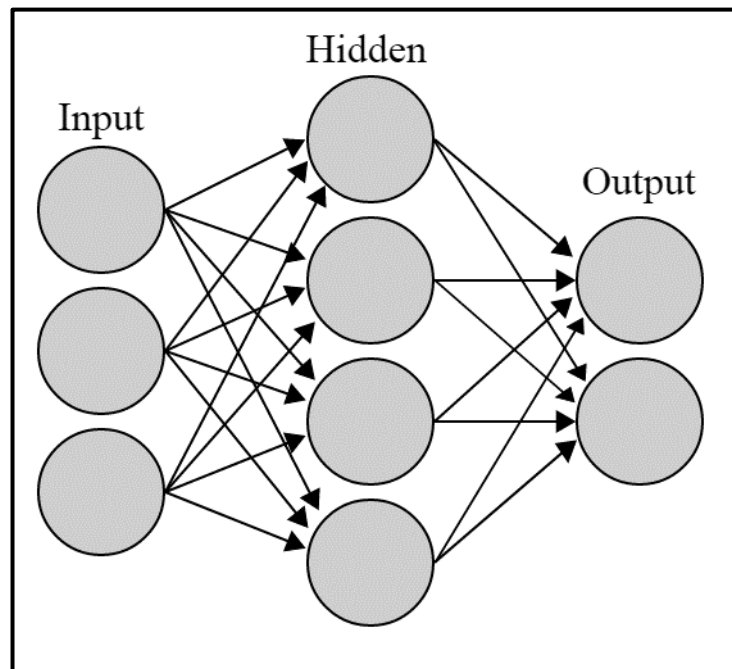


Figure 7. Schematic of artificial neural network

Neural networks are often applied to solve the problems with unknown correlation or cumbersome non-linear models which cannot be defined easily. As the exact solution or even the form of the solution is vague, neural networks require a large number of runs to determine the best solution. A neural network after training with defined weights is ready to use as a powerful analytical tool for other data, plugging in inputs and readily get the results at the output layer.

3.4.1. Models of a Neuron

Neuron is the simplest unit of neural networks acts as an information-processing unit, which forms a complex neural network to process the input data. A neuron consists of more basic elements like synapses, an adder, and an activation function demonstrated in Figure 8. According to the figure, a set of synapses or connecting links, transfers the input signal x_i to the neuron k which determined by the synaptic weight w_{ki} . In other words, w_{ki} is a synaptic weight that could take a negative or positive value multiplies the input signal i and sends the result into the neuron k . An adder adds up all the weighted input signals and bias in the neuron. The summation of the weighted signals and bias go to the activation function, which is a linear or nonlinear function. Activation function keeps the output values into a limited ranges. The output could be an input for another neuron in the next layer.

In the mathematical term the Figure 8 is written as:

$$v_k = \sum_{i=0}^m w_{ki} x_i \quad \text{Eq. 19}$$

and

$$y_k = \varphi(v_k) \quad \text{Eq. 20}$$

where x_i 's are the input signals. To make it easier for showing the formula as well as a matrix algebra operation, a hypothetical signal $x_0 = 1$ is considered to take the bias term into account. The weigh for bias is defined as $w_{k0} = b_k$.

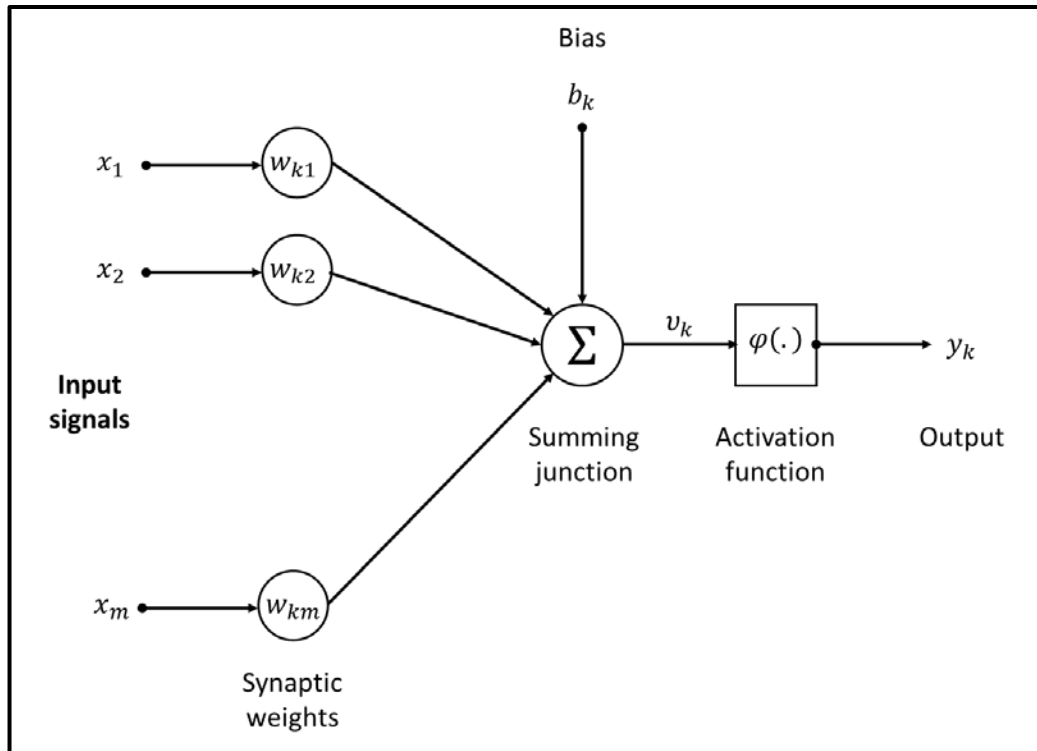


Figure 8. Nonlinear model of a neuron (Haykin, 1999).

3.4.2. Types of Activation Function

The activation function $\varphi(v_k)$, translates the induced local field v_k into a limited value. This function is defined in several ways, here we introduce three basic types of the activation function.

3.4.2.1. Threshold Function

The threshold function is defined as:

$$\varphi(v_k) = \begin{cases} 1 & v_k \geq 0 \\ 0 & v_k < 0 \end{cases} \quad \text{Eq. 21}$$

The output of this function is equal to unit for the positive inputs and is zero for negative inputs. The threshold or reference value for the neuron is bias coefficient. The Figure 9 shows the illustration of this function.

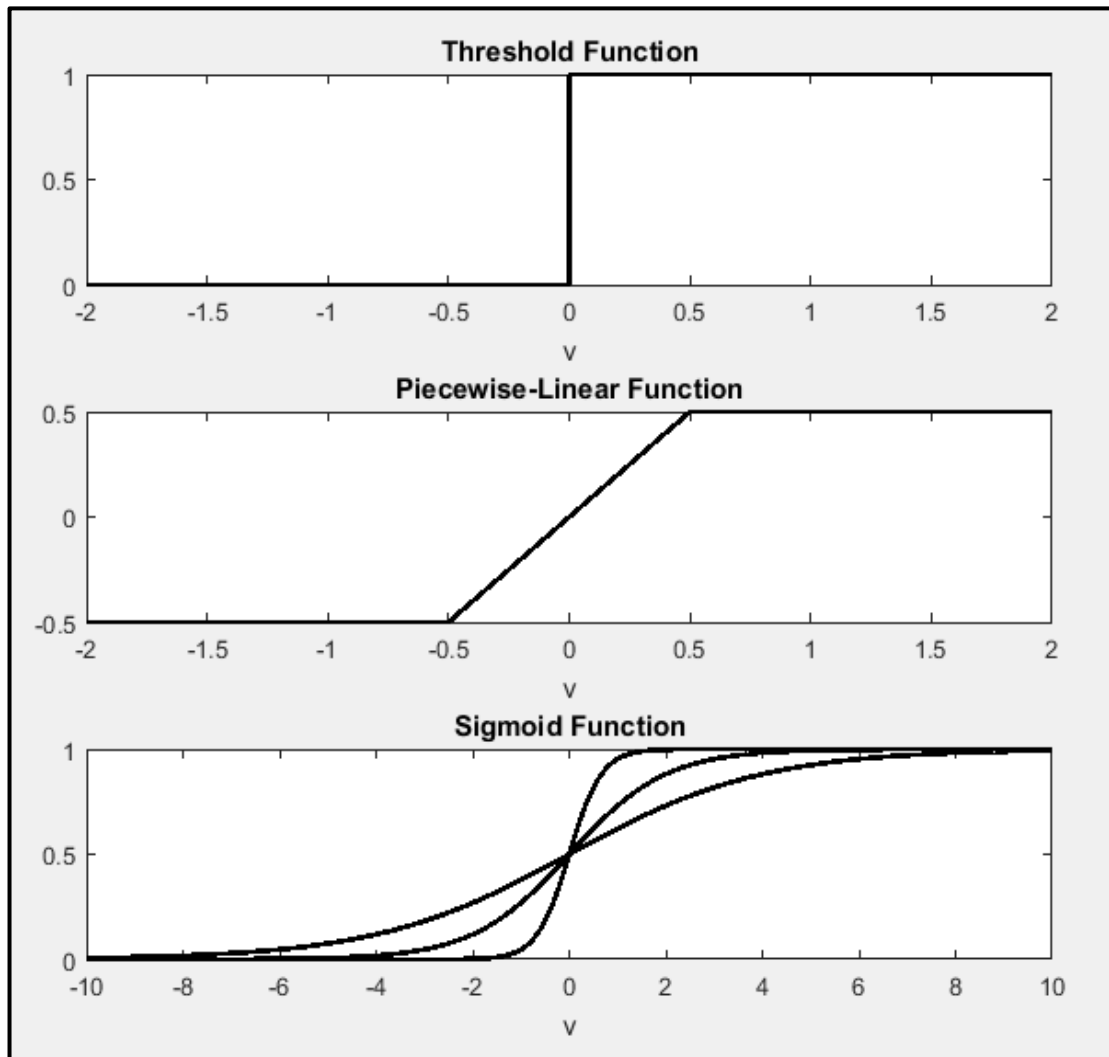


Figure 9. Illustration of different activation functions.

3.4.2.2. Piecewise-Linear Function

The piecewise-linear function is a combination of linear function and threshold functions. The results for input values between -0.5 and 0.5 is a linear output. The piecewise-linear function is defined as:

$$\varphi(v_k) = \begin{cases} 1 & v_k \geq 0.5 \\ v_k & 0.5 > v_k \geq -0.5 \\ 0 & v_k \leq -0.5 \end{cases} \quad \text{Eq. 22}$$

As seen in Figure 9, the middle part of the output behave as a linear function. The out bonds have constant values.

3.4.2.3. Sigmoid Function

The sigmoid function is very common to use in neural networks. It is defined as:

$$\varphi(v_k) = \frac{1}{1 + \exp(-av_k)} \quad \text{Eq. 23}$$

where a is a slope parameter that may have different values. The sigmoid, similar to the threshold and piecewise-linear functions has a value between 0 and 1 but the most important characteristics of the sigmoid function is its differentiability. The Figure 9 shows the illustration of this function with different a 's.

3.4.3. The Architecture of Neural Networks

Neural networks are arranged in different ways from simplest one as a single-layer network to multi-layer and recurrent networks. Single-layer network as showed in Figure 10, does not have any hidden layer. The output layer acts as computational layer.

The output from this layer is the final output and is considered as the response of the neural network. This network is a feedforward network.

Multilayer neural network as shown in Figure 11, has one or more hidden layers. The figure illustrates a 5-4-3-1 multilayer neural network in which the computational neurons are located in the hidden layer. This network has 5 nodes as input which feed into the second layer or first hidden layer as input values. The outputs from the first hidden layers go to the second hidden layer as input. The last layer is the output layer which has a single output as a final result.

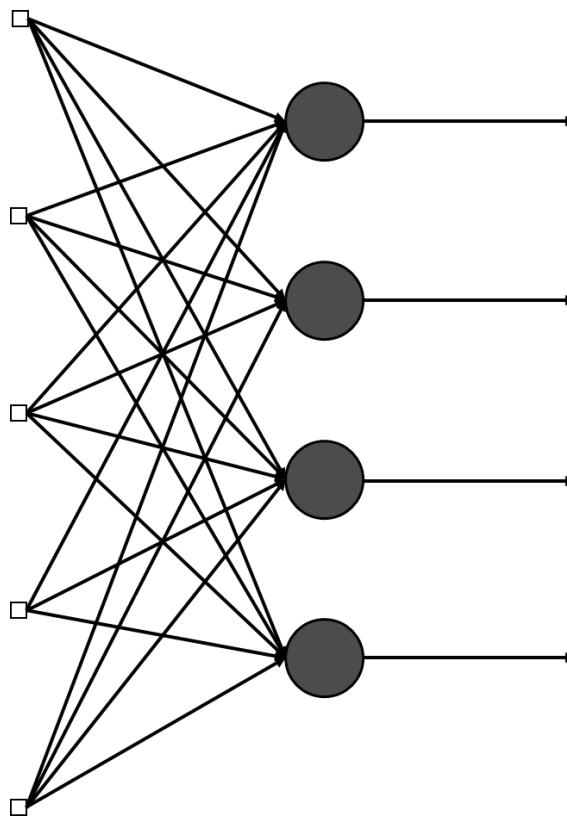


Figure 10. Single layer neural network.

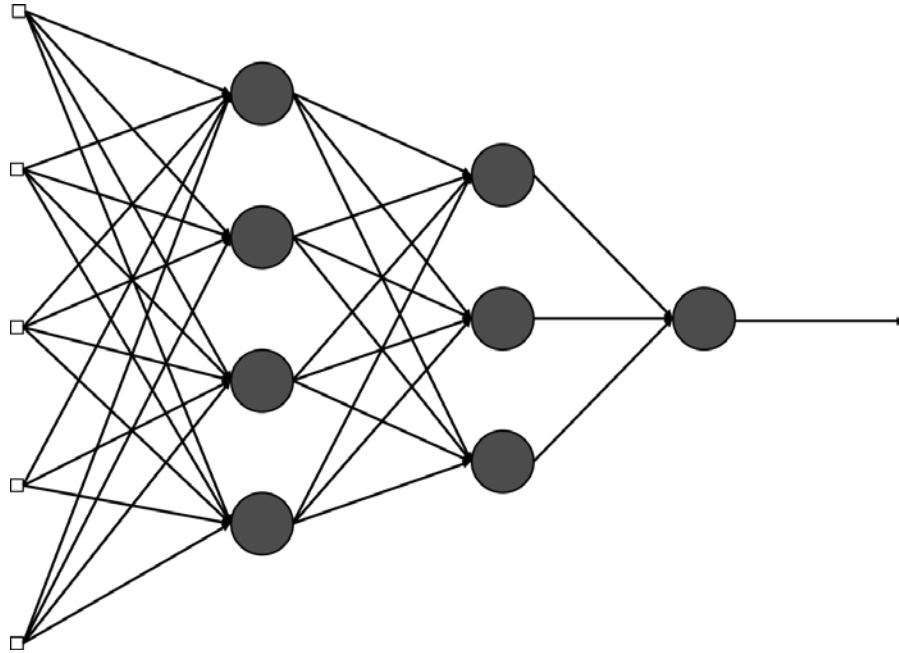


Figure 11. Multilayer neural network.

The other type of neural network arrangement is recurrent neural network, which has a feedback from the result affecting the input layer with delayed values. As seen in Figure 12, the delayed values could come from a hidden layer as final or intermittent results, which are stored in the memory by a delay operator. This feedback improves the results.

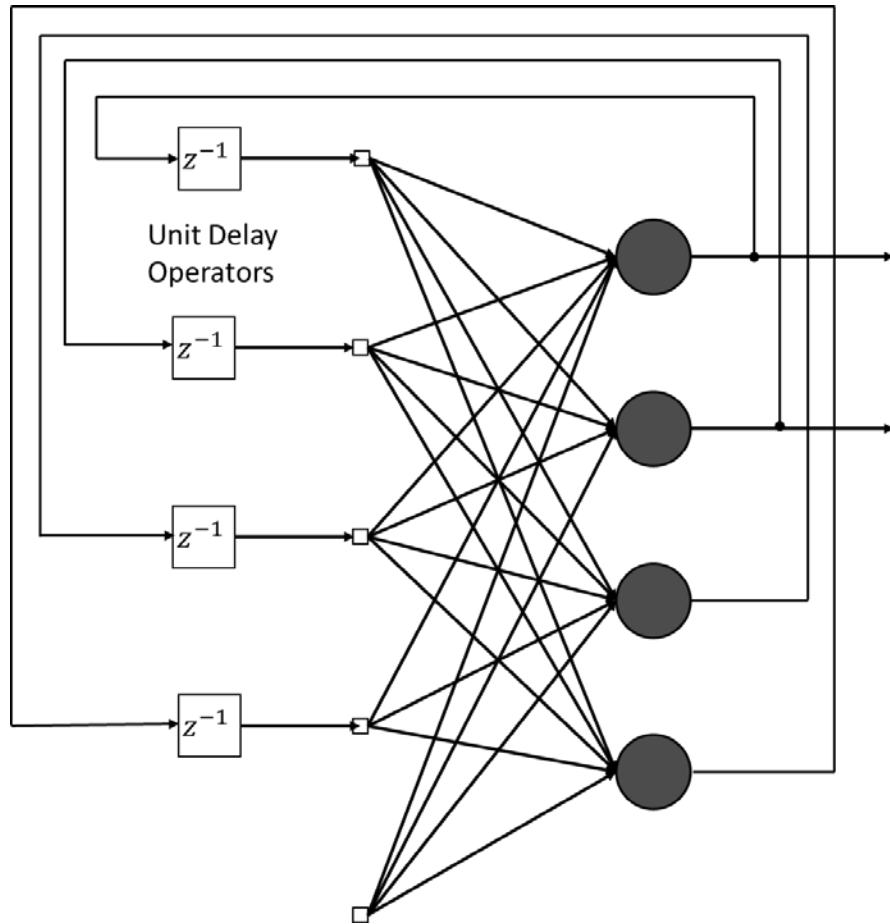


Figure 12. Recurrent neural network.

3.4.4. Learning Process

The ability to learn of the neural network is a very important feature, which help it to improve by new examples. Neural networks are stimulated by environment as the input signals which, results in the respond based on the current situation. Comparing the response and the environment, neural networks adjust the synaptic weights to gain a better response. This process is learning process classified in different categories as describe in following sections.

3.4.4.1. Error-Correction Learning

At this type of learning, the response of the neural network is compared to the desired answer that is given and the error function is calculated. The objective is to minimize the cost function which is calculated from the produced error. Let take $y_k(n)$ as a response of neuron k, $d_k(n)$ as the desire response of that neuron, and $e_k(n)$ as an error. n demonstrates the time step of the process.

$$e_k(n) = d_k(n) - y_k(n) \quad \text{Eq. 24}$$

Mean square error (MSE) is defiend as

$$MSE(n) = \frac{1}{2} \sum e_k^2(n) \quad \text{Eq. 25}$$

To minimize the MSE, we define $\Delta w_{ki}(n)$, the adjustment to the synaptic weights as:

$$\Delta w_{ki}(n) = \eta e_k(n) x_i(n) \quad \text{Eq. 26}$$

where η is a positive value named learning rate. The synaptic weight in the next time steps would be determined by:

$$w_{ki}(n + 1) = w_{ki}(n) + \Delta w_{ki}(n) \quad \text{Eq. 27}$$

3.4.4.2. Memory-Based Learning

In this method, all or part of the previous experiences are kept in a large memory of correctly classified examples: $\{(\mathbf{X}_i, d_i)\}_{i=1}^N$, where \mathbf{X}_i is the input signal vector, and d_i is the corresponding desired response. The memory based algorithm consists of two parts (Haykin, 1999):

- Criterion used for defining the local neighborhood of the vector, and,

- Learning rule applied to the training in the local neighborhood.

There are variety of algorithm in which these parts with neighborhoods are defined. One of the most famous memory based learning processes is the nearest rule, where the local neighborhood is defined as the training example that lies in the immediate neighborhood of the test vector.

3.4.4.3. Hebbian Learning

Hebbian learning rule is the oldest and most famous learning rule proposed by (Hebb, 2005) as “*when an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A’s efficiently as one of the cells firing B, is increased*”. The other name for this learning type can be correlational learning.

Based on Hebb’s statement we may have two following parts:

- If two neurons linked by a synapse are activated simultaneously, then the synapse is strengthen selectively.
- If two neurons linked by a synapses are activated asynchronously, then the synapse is weakened or eliminated selectively.

This form of synapse is called Hebbian synapses. There are four mechanisms that distinguish a Hebbian synapse: time dependent mechanism, local mechanism, interactive mechanism, and correlational mechanism.

The simplest way to show the Hebbian learning in mathematical form is as follows:

$$\Delta w = x_i y \quad \text{Eq. 28}$$

This rule is totally based on feed forward, unsupervised learning. It means that if the cross product of output and input is positive, the weight is increased, otherwise it is decreased (Haykin, 1999).

3.4.4.4. *Competitive Learning*

The competitive learning process is designed based on Hebbian learning method, unless the output neurons compete with each other and only one neuron would be activated instead of the several neurons. This learning method has three basic rules (Rumelhart & Zipser, 1985):

- All the neurons are set similarly except the weights which are distributed randomly that result in the neurons respond to differently to the same input patterns.
- The strength of each neuron is limited.
- There is a mechanism that allows only one neuron or one in a group to get activated at the same time. The neuron that wins the competition is called the winner-takes-all neuron.

3.4.4.5. *Boltzmann Learning*

Boltzmann learning is a stochastic learning method. In Boltzmann network, the neurons have a recurrent structure and they respond in a binary form. A main concept in this learning process is energy function which is defined as:

$$E = -\frac{1}{2} \sum_i \sum_j w_{ij} x_i x_j \quad \text{for } i \neq j \quad \text{Eq. 29}$$

where x_i is state of the i th neuron, w_{ij} is the synaptic weight between neuron i and j , and $i \neq j$ states that the neuron does not have self-feedback. The neurons are chosen randomly.

The neurons are separated into two visible and hidden classes. Visible neurons have interface with the environment while the hidden neurons operate independently (Haykin, 1999).

3.5. Factors and Variables

In a competitive market (e.g. the U.S. energy market), the price of commodity will vary until it reaches the economic equilibrium point which is a point where the quantity demanded will be equal to the quantity supplied at the current price. According to this theory, excess supply results the price decline and excess demand yields increment of the price. The equilibrium state remains same until the market situation is not changed. If some condition changed in the current market, for example an unforeseen cold winter occurs, the level of demand will change. This change defines a new point as an economic equilibrium state and prices move to the new point. The new equilibrium price is the price in which supply and demand are balanced in the same quantity transaction.

The models for prediction the price are classified into two different categories as univariate and multivariate models. Univariate models just attempt to simulate the phenomena by the target variables itself and its lags. In this models, other parameters that may affect the behavior of the dependent variables are not considered assuming all the

causes are reflected in the prices. The second class, multivariate models, enters the different parameters, which are called exogenous variables. These variables may affect the dependent variable behavior based on the economic, social, or political situations. The problem with these models is to require forecasting these parameters in the future in order to predict the target variable. At this section, we review some exogenous variables, which were considered as dominant variables respect to the gas price movement in the literature.

Some factors may also affect the supply and demand of the natural gas such as macroeconomic parameters, natural gas production and consumption, storage and inventories, weather conditions, pipeline capacity, and fuel substitutions. Political issues, military interface, and economic instability as an important factor in the energy market, particularly in the oil market are not considered in the previous researches (Chiroma, Abdulkareem, Sameem, Abubakar, Adamu, & Mohammed, 2013) because evaluation and quantizing of these variables is not an easy task.

3.5.1. Natural Gas Production and Consumption

Production and consumption of natural gas are the variables that are directly tied to the supply and demand. The amount of the production and consumption as well as their difference have significant effect on gas prices. The gap between supply and demand acts as a driving force for the price movement.

3.5.2. Pipeline Capacity

Since most of the transportation of natural gas is performed by pipeline In the U.S., the capacity of the pipeline is restricting the natural gas transportation. When

demand of the natural gas rises rapidly and reaches the pipeline capacity limits, more production and storage withdrawal cannot compensate the market demands. Therefore the maximum capacity of the pipeline and its available capacity would be important factors in the market adjustment (Avalos, 2012).

3.5.3. *Storage of Natural Gas*

Geman and Ohana (2009) asserted there is a correlation between the inventory and natural gas price when inventory is in below its long run average. Storing of the natural gas is not easy. Natural gas mainly stores in underground reservoirs. This type of storage is not available anywhere, we need it. Natural gas production is almost constant during the year, but the consumption varies in different seasons. To damp the seasonal fluctuation of natural gas demand, it stores in the underground storage and restores when the demand is higher than production. Song et. al. (2013) showed that natural gas price respond to the storage report shock. When the storage at a certain time is less than expected value at the similar time, the market is shocked about the storage amount. If the expectation was much higher than the existent stored gas, the price rises.

The inventory itself is important, but not include enough information about the supply or demand of the commodity. The market expectation of the sacristy of the commodity is more important than the absolute value of current storage. Busse et al. (2012) found that the level of storage does not have a significant effect on the natural gas price movement.

3.5.4. *Substitutions*

Several years ago, before 1990, burner tips were designed in such a way that could switch the fuel between natural gas and oil products easily. This feature made the refined petroleum products as a perfect substitution for natural gas. Recently the situation became different and designs changed. It is not so easy to switch the fuel as before (Brown & Yücel, 2008).

There are various researches with different result about the relationship between the refined petroleum products and natural gas. Villar and Joutz (2006) show a strong relationship between the natural gas and the oil, in the other hands, Bachmeier and Griffin (2006) claim there is a relationship between these fuels but it is weak. Brown and Yücel (2008) found that the relationship is significant which is conditioned by the other factors like weather, inventories, and shut in production. Part of the price for natural gas and oil comes from the cost of exploration, extraction and taxes which are almost same for both commodities.

Recently world powers had a contraction to reduce the CO₂ emission in the world. Natural gas as clean fuel, which has a lower emission respect to other fossil fuels like coal and petroleum residuals is a suitable replacement. In order to reduce the CO₂ emission and air pollution, the government policy is to encourage the consumers to move on more environmental friendly fuels like green energies and natural gas.

An important question here is that if the oil price is an exogenous variable or endogenous. Villar and Joutz (2006) state that it is an exogenous variable. This question is debatable and needs more investigation. As a rule of thumb, most researchers agree with Villar and Joutz.

3.5.5. *Weather Effect*

Energy consumption is different in various weather conditions and temperatures. For instance, residents consume more energy to heat up their home, when the weather temperature is too low. Also in high temperature condition, lots of energy is required to cool down the places we live. Therefore, in high and low temperatures, energy demand increase dramatically, respect to the moderate temperatures (Considine, 2000). This rise in consumption affects the demand side of the market (Bower & Bower, 1985). Since this demand is predictable in severe conditions of summer and winter, suppliers would be prepared for the high demand. The market flocculation is damped by storing the natural gas in summer and restoring in winter. This seasonality effect changes the U.S. natural gas market demand up to 50% (Mu, 2007). The residential and commercial sector consumption is more sensitive to the weather changes respect to the industrial section (Elkhafif, 1996). Considine (2000) also states that warmer climate condition slightly reduces the natural gas consumption and energy demand.

The weather is more imperative when a sudden change occurs in climate. A big change that was not predicted before, a change with a low probability for that season, may lead the price level to a dramatic high spike. We call this incident as weather shock. Mu (2007) shows that weather shock has an effect on natural gas market in both spot and future trading.

Heating/Cooling degree days are indicators of household energy consumption for space heating/cooling. Heating/Cooling degree days are defined as the amount of temperature below/above the base temperature. For an average outdoor temperature of 65

degrees Fahrenheit, most buildings require to heat/cool to maintain a 70 degree temperature inside (ASHRAE, 1993). Therefore, we assume 65 °F as the base temperature in our calculations.

The following linear functions are cooling and heating degree days in term of temperature of current day.

$$\begin{aligned} CDD &= \begin{cases} 0 & T \geq 65^{\circ}\text{F} \\ 65 - T & T < 65^{\circ}\text{F} \end{cases} \\ HDD &= \begin{cases} 0 & T \leq 65^{\circ}\text{F} \\ T - 65 & T > 65^{\circ}\text{F} \end{cases} \end{aligned} \qquad \text{Eq. 30}$$

Chapter 4: Results and Discussions

The first step in statistical analysis is to investigate the time series distribution. The Jarque-Bera normality test shows that the price time series is not normally distributed but it follows a log-normal distribution (Jarque & Bera, 1987). Figure 13 illustrates the distribution of the logarithmic price with respect to the reference red line in the normal plot. Therefore, the log price or the return of the natural gas price is normally distributed and follows the Gaussian distribution. Now, it is appropriate to assume normal distribution for first the difference of variable and/or the logarithm of the gas price.

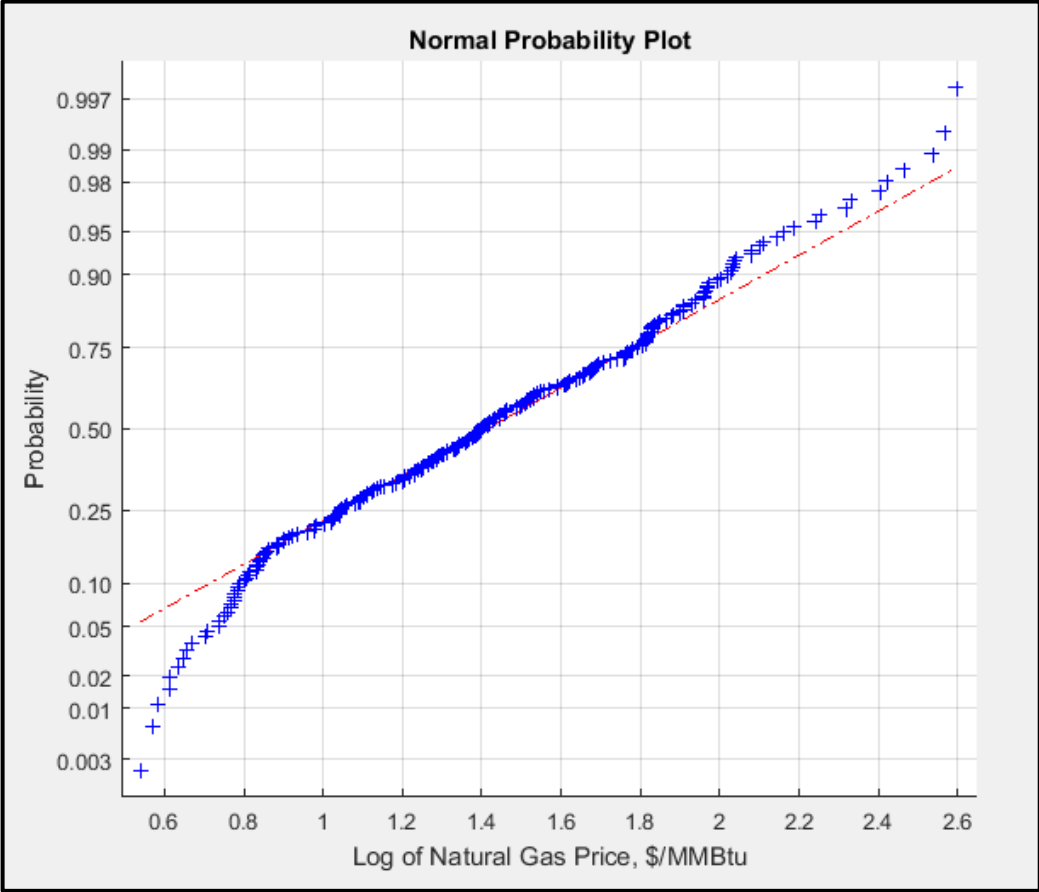


Figure 13. Normality test on gas price time series.

4.1. Stochastic Differential Equation Models

In this section, some SDE models have been examined to fit and predict the natural gas prices. The models are univariate single factor model which are analyzing the gas price behavior based on its historical information.

4.1.1. Geometric Brownian motion (GBM) Model

This model, as mentioned before, is described by the following SDE:

$$dX_t = \mu(t)X_t dt + \sigma(t)X_t dW_t \quad \text{Eq. 31}$$

The drift and diffusion terms could be considered as constant values or variables varying with time. Figure 14 demonstrates the solution with the constant parameters. The model was run for the existing time series since the beginning. The solid line is the historical price while the dotted red line is the one step ahead forecast based on the GBM model. As mentioned before, the stochastic differential equation model solutions contain random variables. Each realization has its own path. Therefore, the model was run for 1000 times and the mean value of these solutions was considered as the forecasted value.

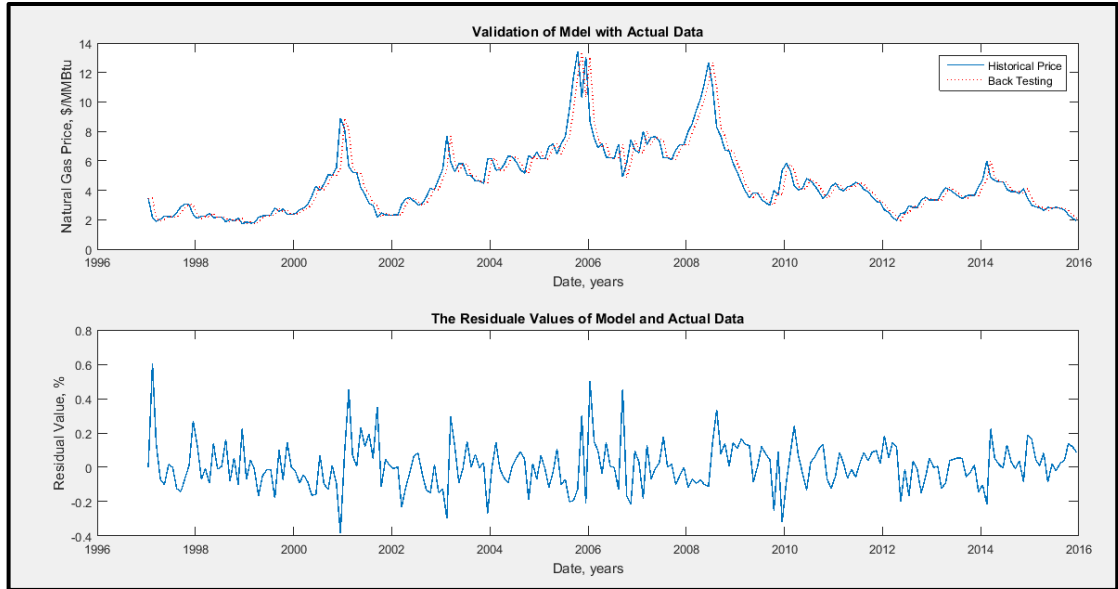


Figure 14. GBM model back test results (constant parameters)

After back testing the model, prediction values for 1, 6, 12, and 24 months have been produced. Same method was applied for the forecasting: the model has been run for 1000 times and the average value for each step has been considered as the solution. The lower and higher possible prices with a 95% confidence level have been determined. Figure 15 shows the histograms of the forecasted price values for 1, 6, 12, and 24 steps ahead. The distribution is a log normal distribution. This behavior was predictable from the definition of the SDEs. For example, the predicted price for two years later, 24 steps ahead from now, is 1.45 \$/MMBtu, which tolerates between 0.47 and 4.31 \$/MMBtu with a 95% confidence level. The distribution of the forecasted price in the next step is closer to the normal distribution, but the distribution of far future behaves more like a log normal distribution.

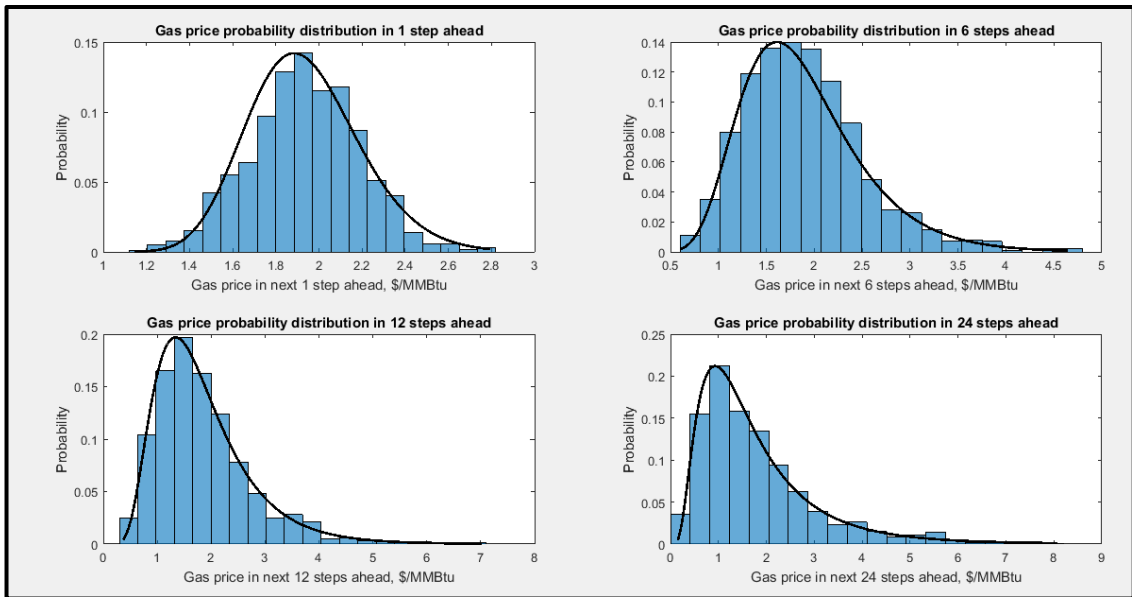


Figure 15. The obtained solutions for 1, 6, 12, and 24 steps ahead by GBM model (constant parameters)

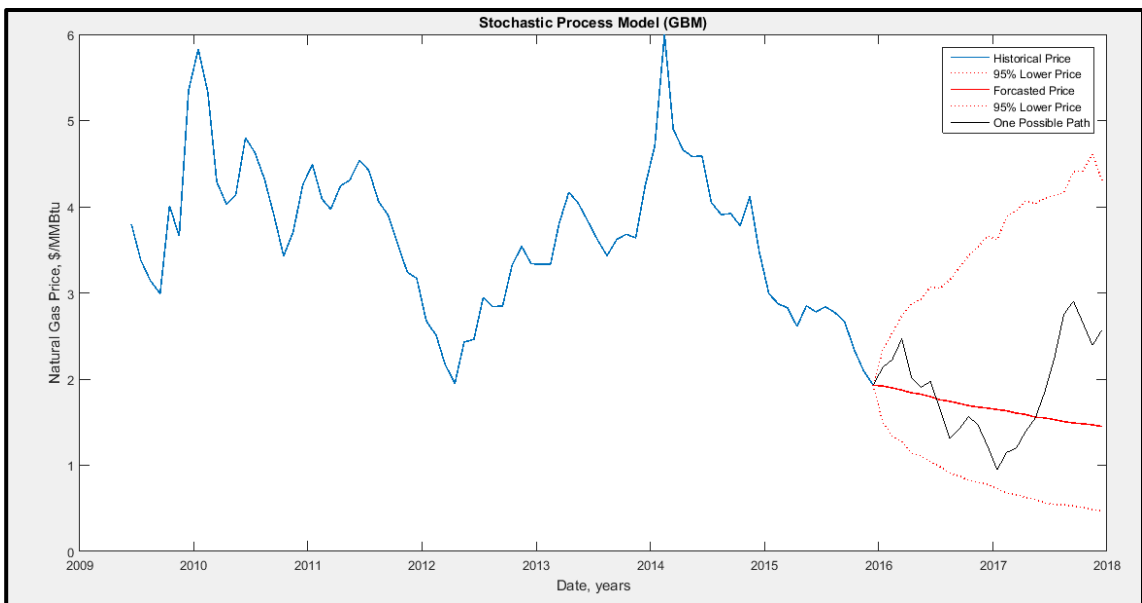


Figure 16. Gas price prediction based on GBM model with constant statistical parameters (constant parameters)

Figure 16 illustrates the forecasted price with 95% confidence bonds. The black solid line is one of the realized paths of the simulated prices. The model is trying to

forecast the price in the future. The gas price has a falling trend for 24 steps ahead. As time goes by, the prediction is less accurate.

The same procedure has been followed for modeling the gas price with variable mean and standard deviation. For the first 100 data, the parameters have been considered as constant parameters and then the new data were added to the previous data and new parameters were calculated. These parameters were used as the input parameters in the model. There is no significant difference observed between the results of Figure 17 through Figure 19 and the results from the constant parameter model.

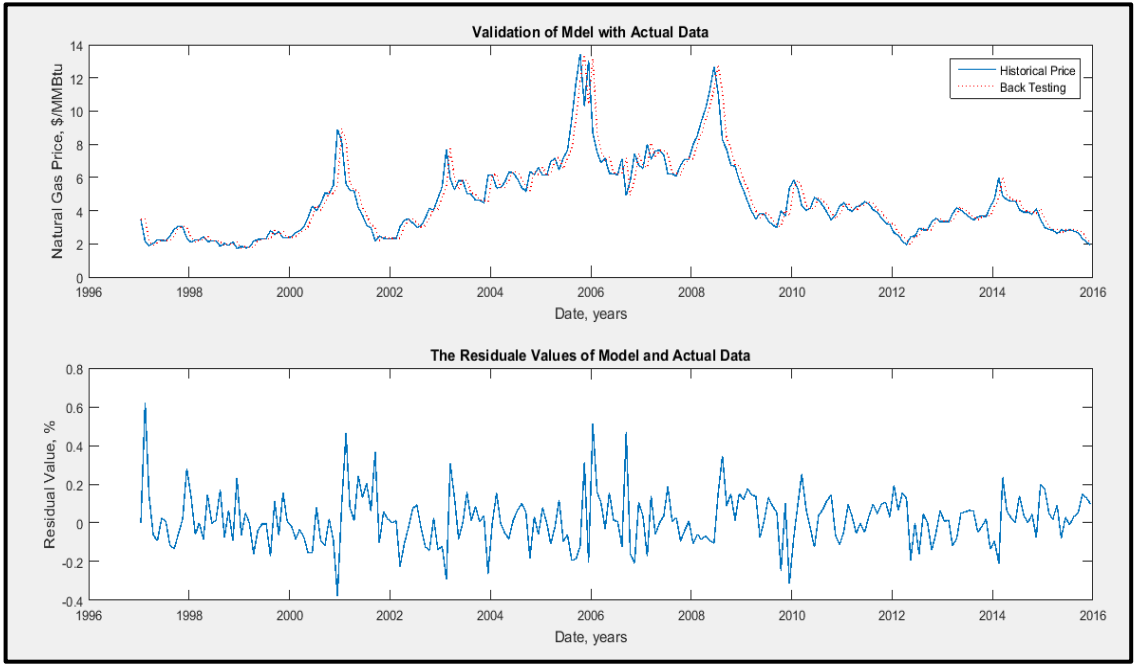


Figure 17. GBM model back test results (variable parameters)

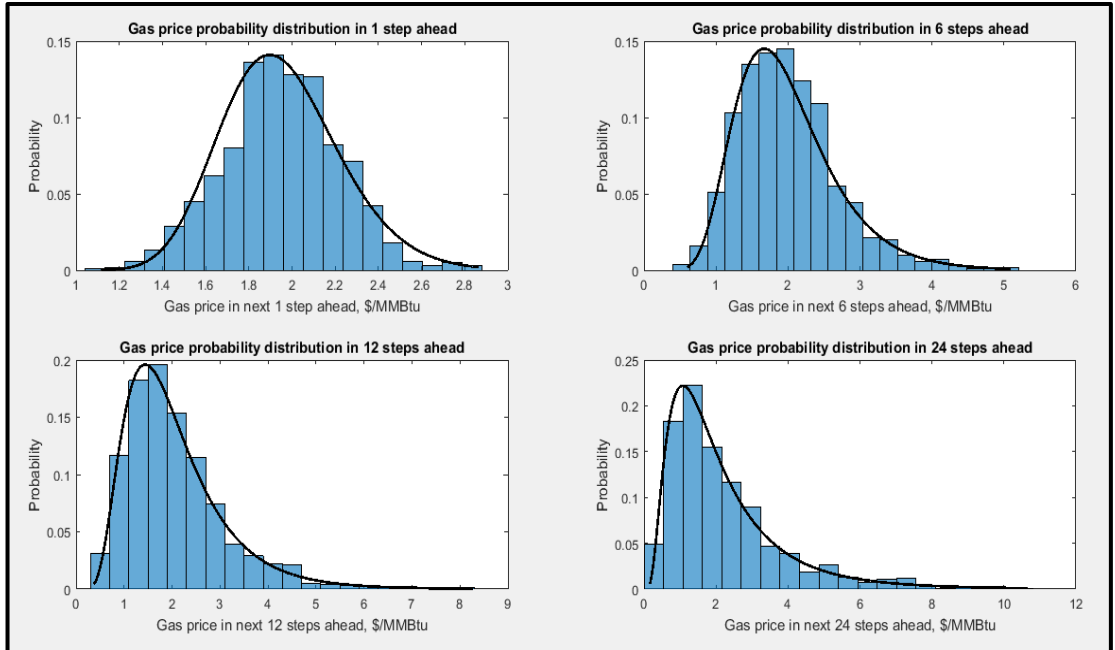


Figure 18. The obtained solutions for 1, 6, 12, and 24 steps ahead by GBM model (variable parameters)

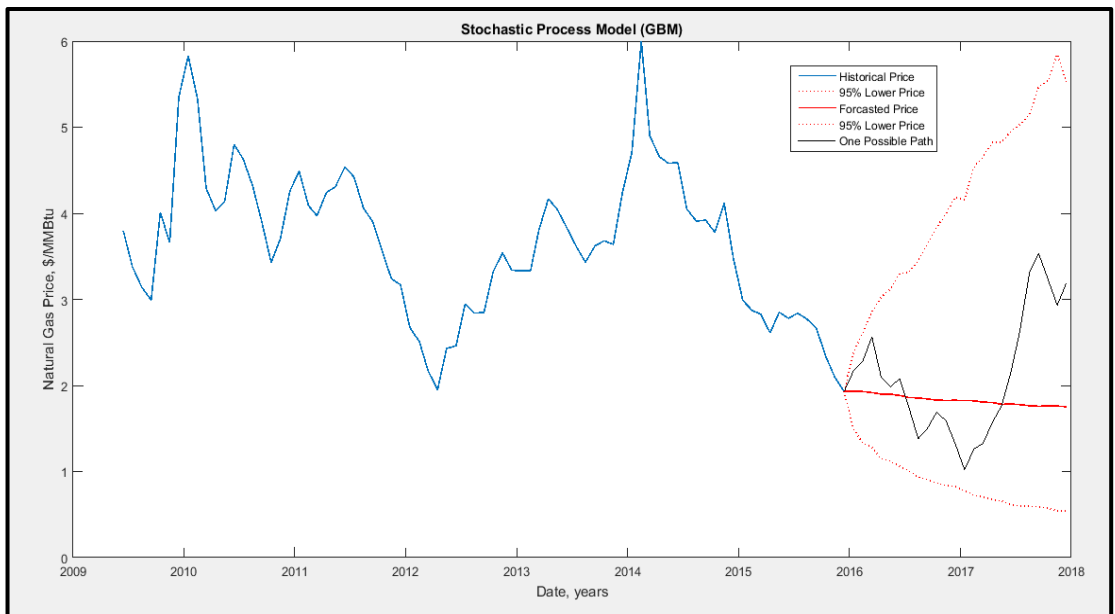


Figure 19. Gas price prediction based on GBM model with constant statistical parameters (variable parameters)

4.1.2. Constant Elasticity of Variance (CEV) models

As mentioned before, CEV models are a special case of drift-rated models, which is defined as follows:

$$dX_t = \mu(t)X_t dt + D(t, X_t^\alpha)\sigma(t)dW_t \quad \text{Eq. 32}$$

If the factor α is equal to unity, the CEV model will be converted to the GBM model. For this section, we assume $\alpha = 0.5$ which builds the diffusion function based on the square root of the price variable. The volatility term is a function of square root of the price, which means compared to the GBM model, the volatility in higher prices is less. If the chosen value for parameter α is greater than unity, the volatility of higher prices will increase dramatically. The same procedure and approach for this model were performed. Figure 20 to Figure 25 demonstrate the results for this model.

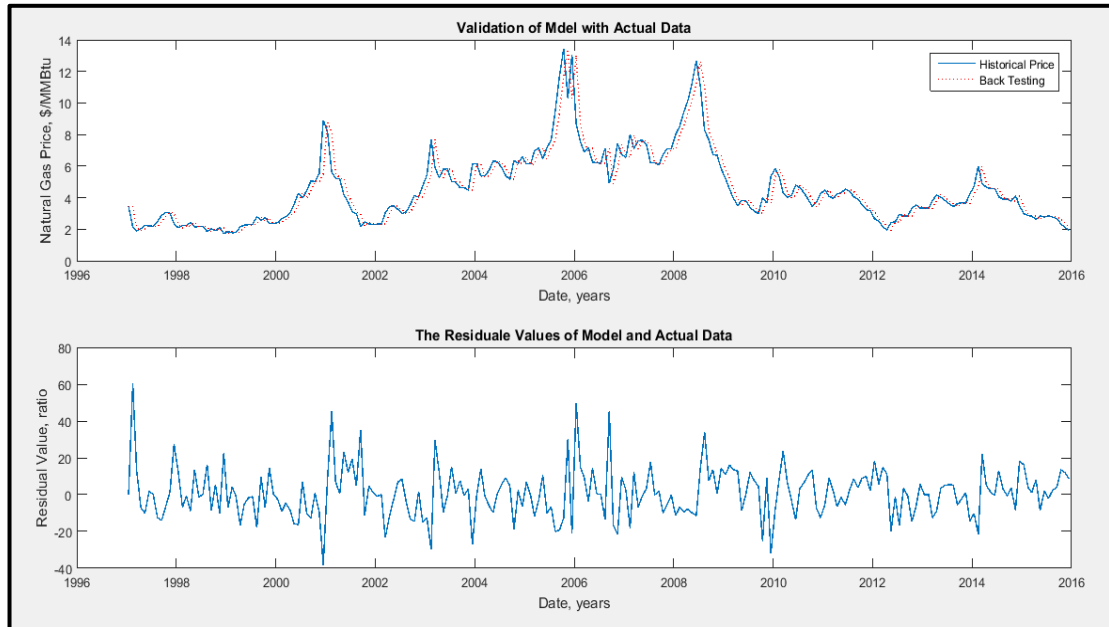


Figure 20. CEV model back test results (constant parameters)

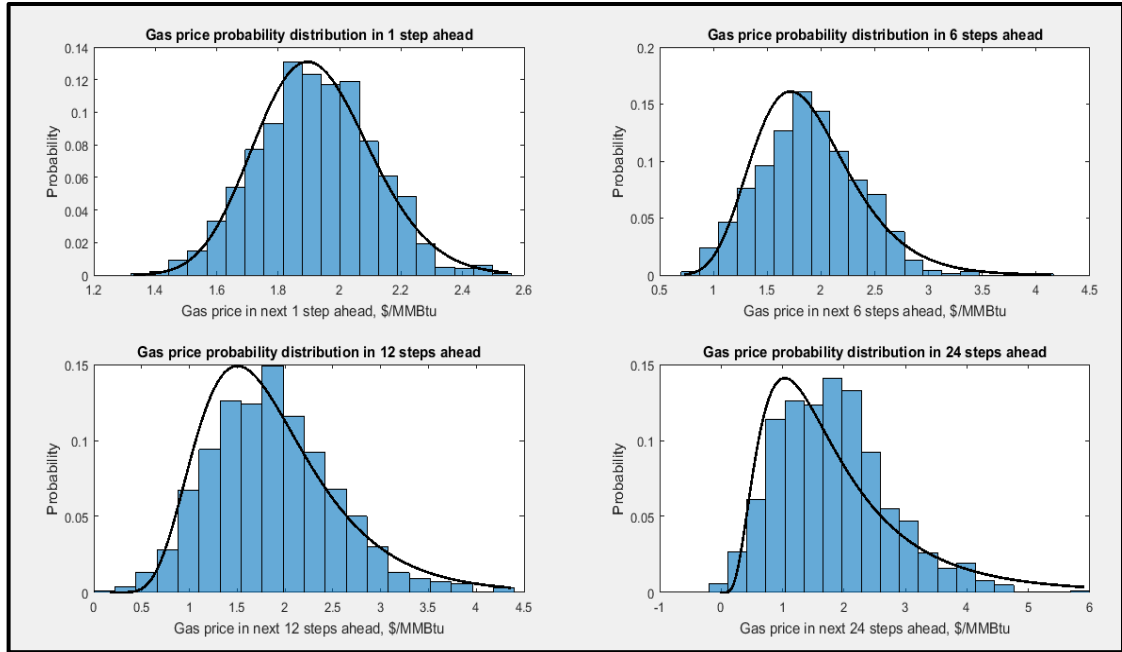


Figure 21. The obtained solutions for 1, 6, 12, and 24 steps ahead by CEV model (constant parameters)

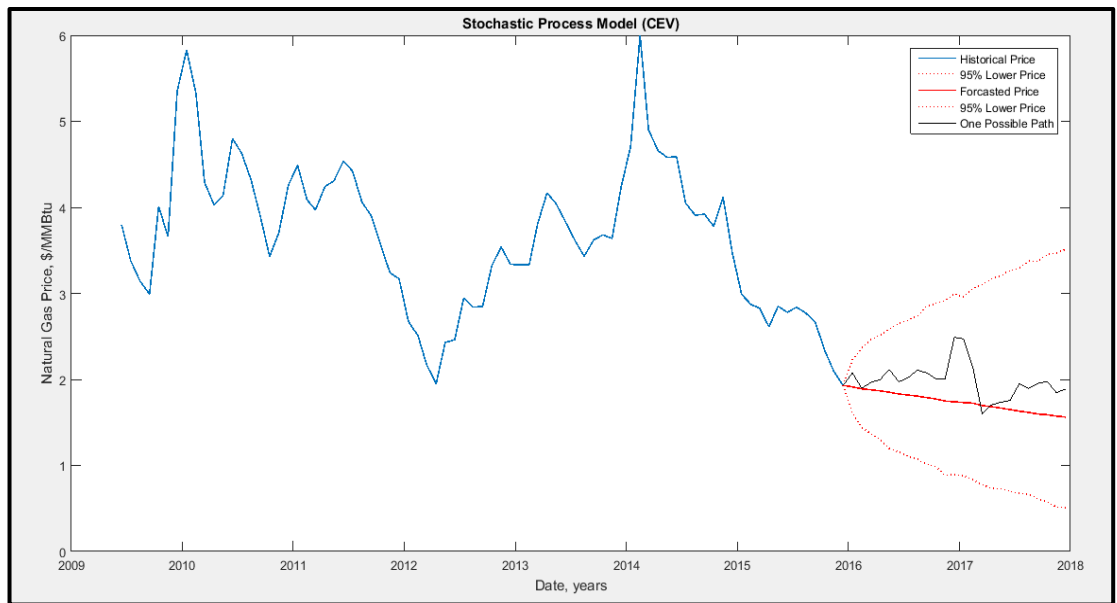


Figure 22. Gas price prediction based on CEV model with constant statistical parameters (constant parameters)

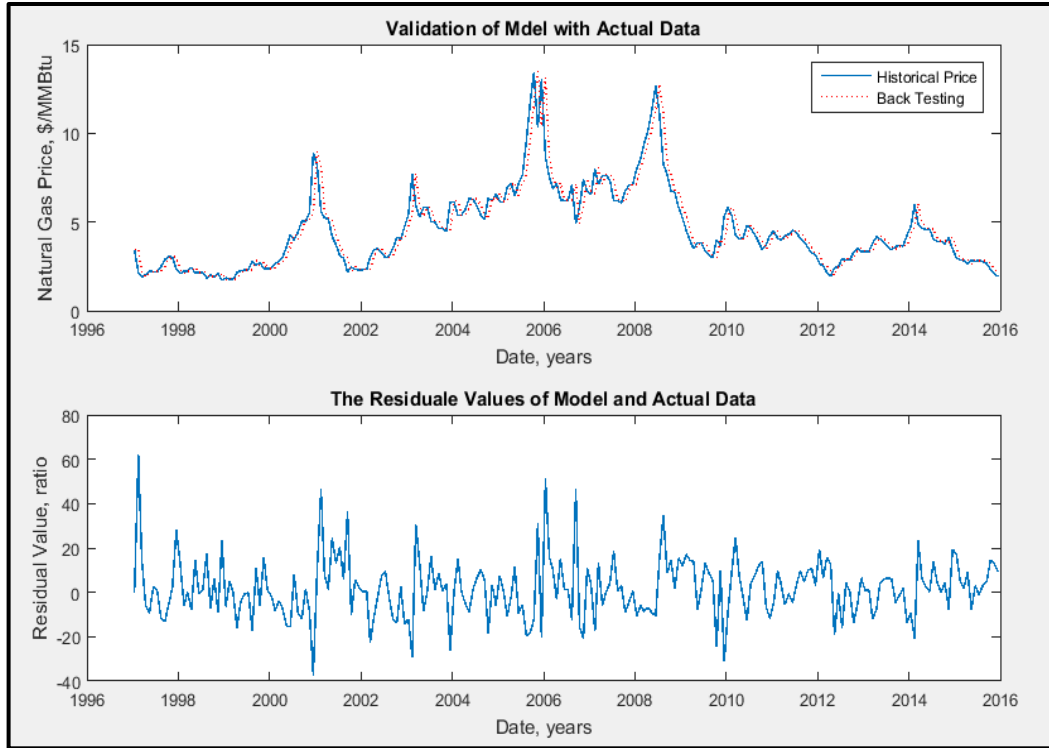


Figure 23. CEV model back test results (variable parameters)

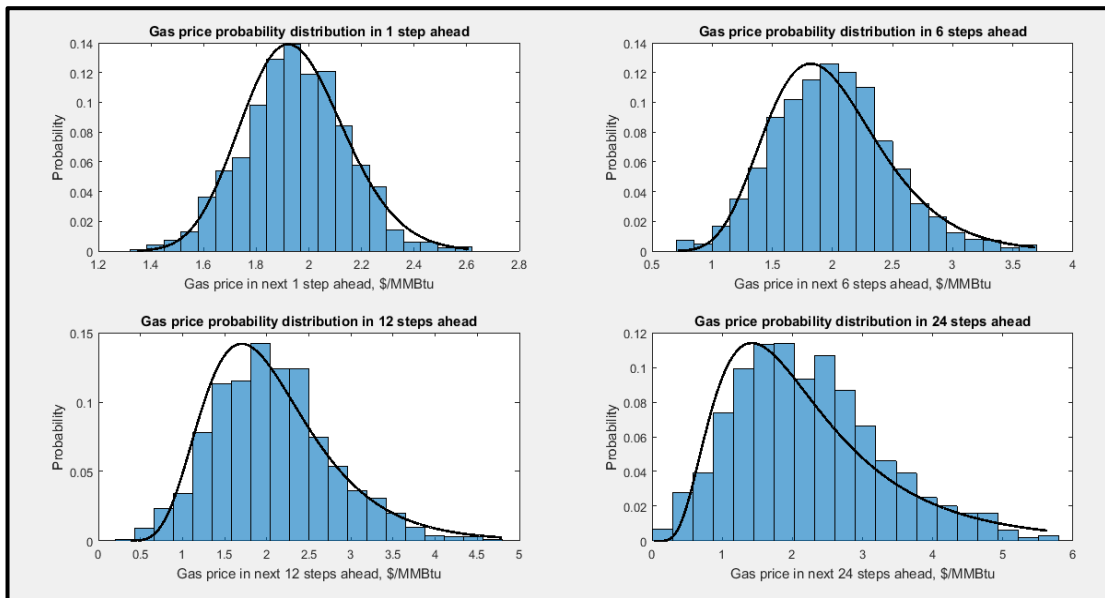


Figure 24. The obtained solutions for 1, 6, 12, and 24 steps ahead by CEV model (variable parameters)

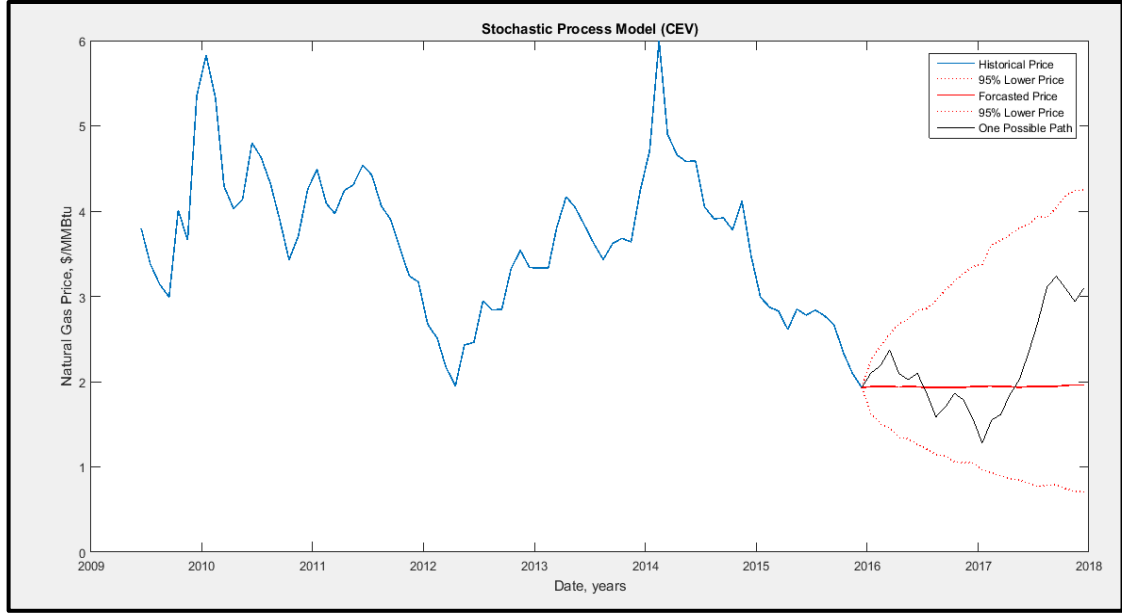


Figure 25. Gas price prediction based on CEV model with constant statistical parameters (variable parameters)

4.1.3. Linear Drift-Rate Models

This model is described based on the following SDE:

$$dX_t = (A(t) + B(t)X_t)dt + D(t, X_t^\alpha)\sigma(t)dW_t \quad \text{Eq. 33}$$

In this model, if the equalities of $\alpha = 1$, $A(t) = 0$, and $B(t) = \mu(t)$ are considered, the model would be as same as the GBM model. Definitely, with $\alpha = 0.5$, the CEV model will be constructed. In order to find a solution for this model we can take $A(t) = B(t) = \frac{1}{2}\mu(t)$ and $\alpha = 1$. For analyzing the data, the same procedure was applied. The following figures show the results:

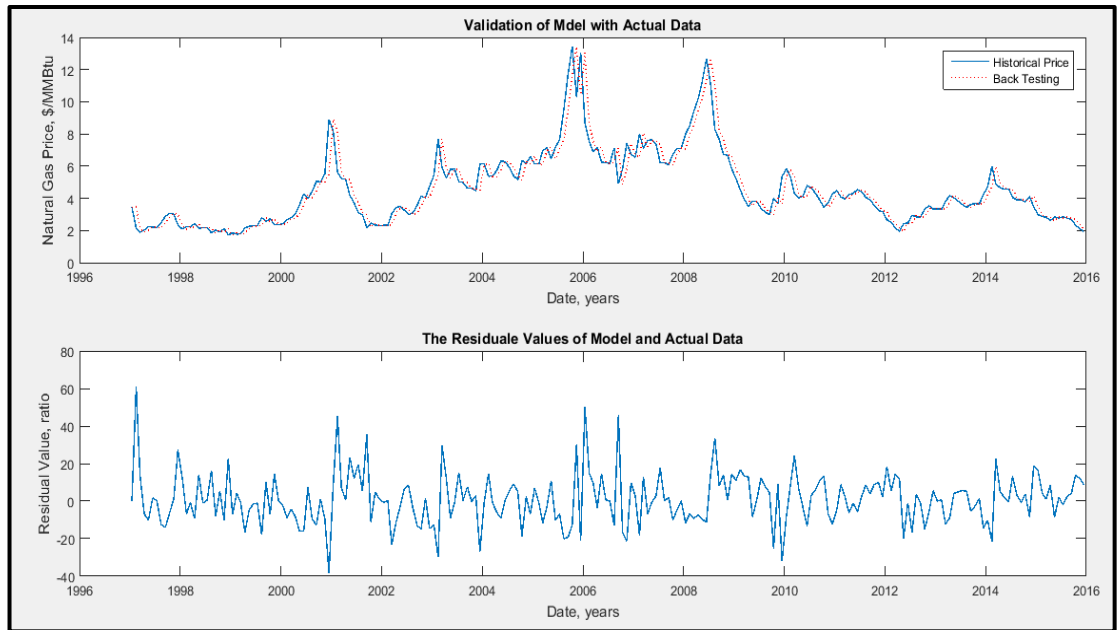


Figure 26. LDR model back test results (constant parameters)

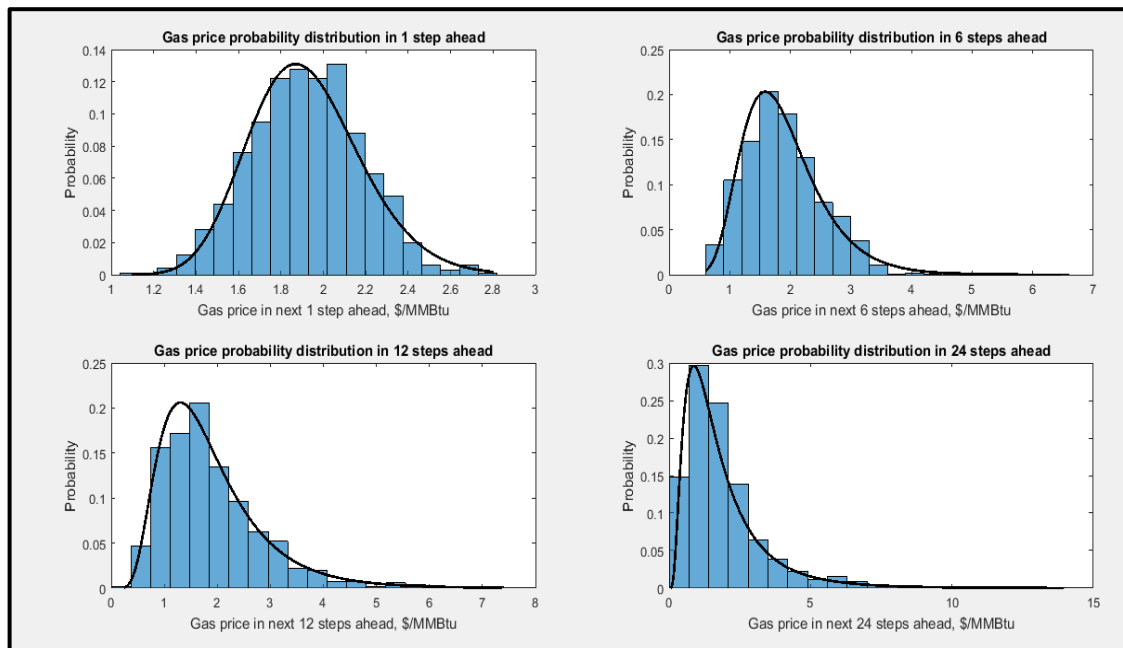


Figure 27. The obtained solutions for 1, 6, 12, and 24 steps ahead by LDR model (constant parameters)

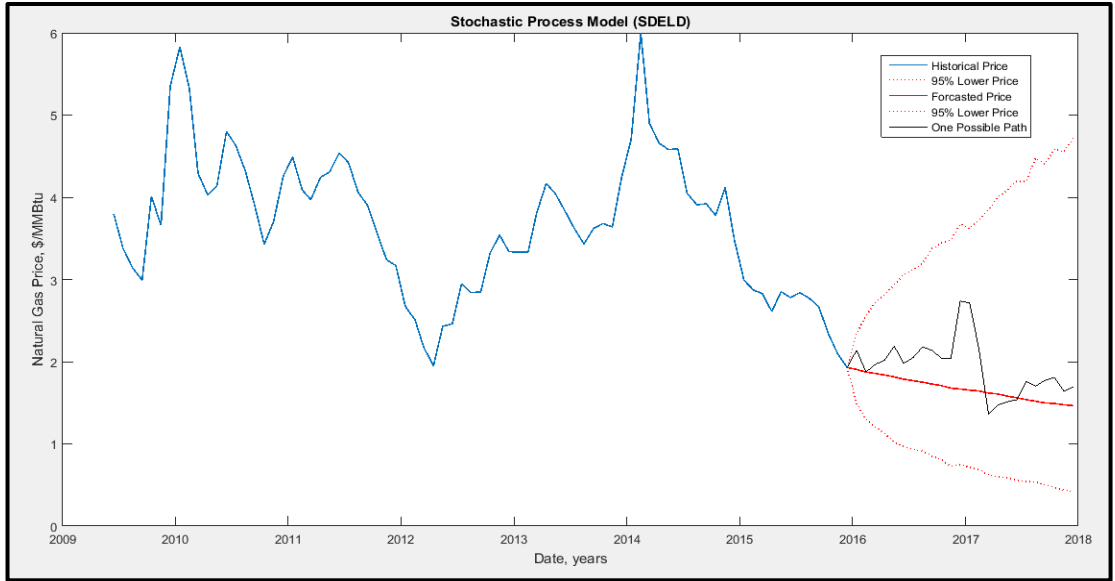


Figure 28. Gas price prediction based on LDR model with constant statistical parameters (constant parameters)

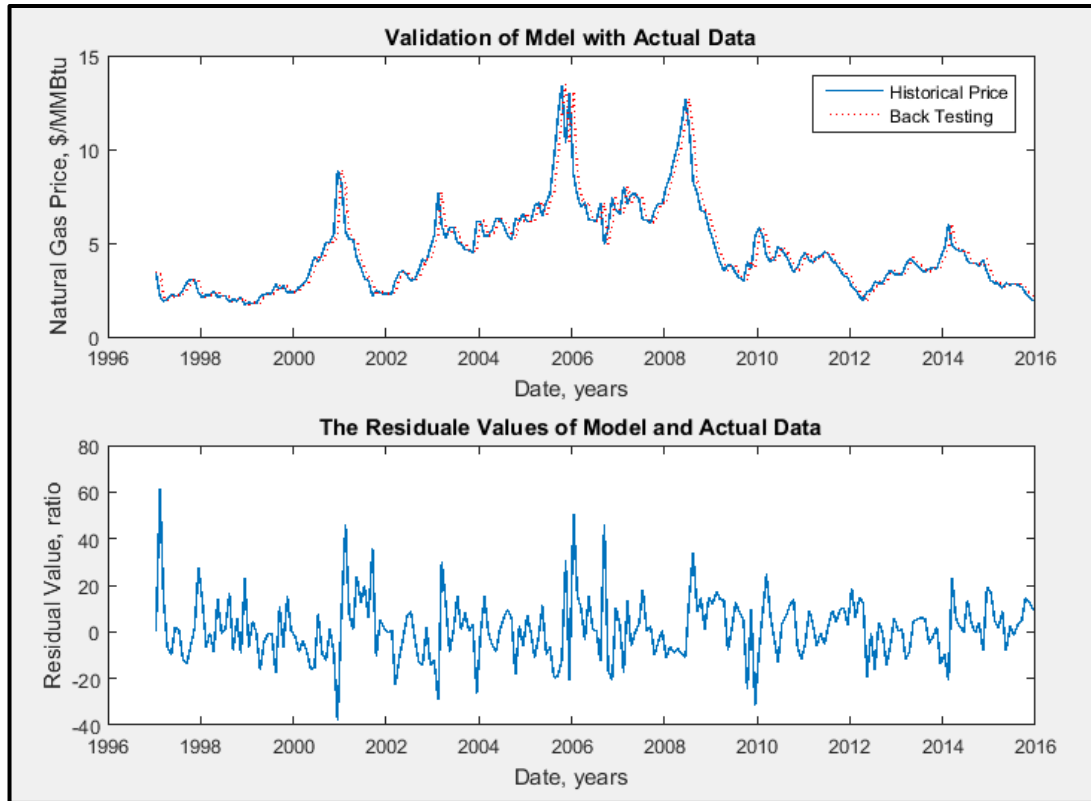


Figure 29. LDR model back test results (variable parameters)

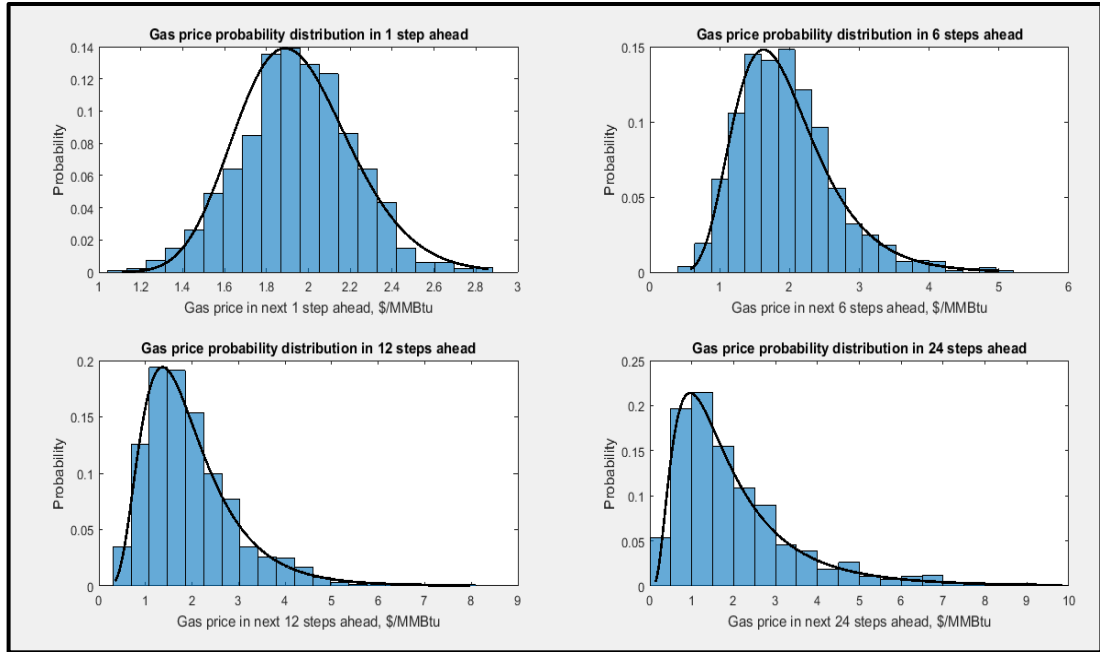


Figure 30. The obtained solutions for 1, 6, 12, and 24 steps ahead by LDR model (variable parameters)

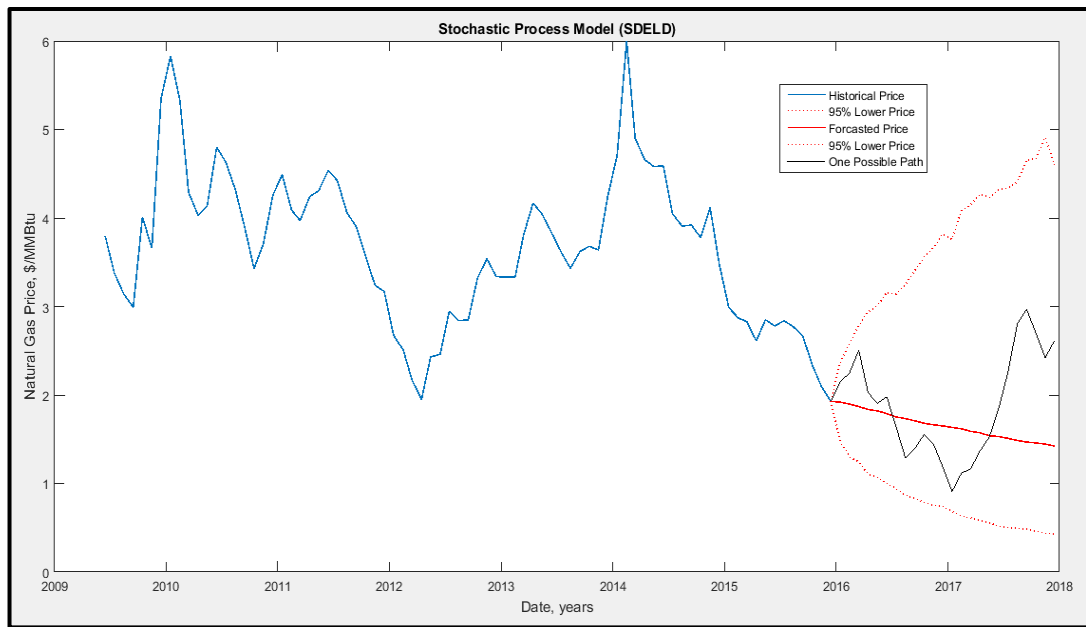


Figure 31. Gas price prediction based on LDR model with constant statistical parameters (variable parameters)

4.1.4. Summary of SDE Models

Table 1 shows a summary of the forecasting results for the presented SDE models. In this table, the minimum and maximum possible prices with a 95% confidence level are given for 1 step and 24 steps ahead. According to the models, the gas price is fluctuating between 0.47 and 5.52 \$/MMBtu for two years. The most possible price for December 2017 is between 1.42 and 1.96 \$/MMBtu.

Table 1. Summary of SDE models forecasting

		1 step ahead forecast			24 steps ahead forecast		
		Min.	Mean	Max.	Min.	Mean	Max.
Static	GBM	1.50	1.91	2.34	0.47	1.45	4.3
	CEV	1.61	1.91	2.23	0.51	1.56	3.51
	LDR	1.49	1.91	2.34	0.42	1.46	4.71
Dynamic	GBM	1.49	1.93	2.39	0.55	1.75	5.52
	CEV	1.49	1.94	2.38	0.71	1.96	4.25
	LDR	1.48	1.92	2.37	0.43	1.42	4.59

4.2. Time Series Models

In this section, the natural gas monthly spot price at Henry Hub is considered as a time series and it was attempted to fit the best model of ARIMA/GARCH by using Econometrics Toolbox™ of MATLAB® software. ARIMA model as a general form of univariate time series has been selected. To establish the model first the mean and variance are considered as constant values, and then the parameters were estimated. In the next step, the statistical parameters were considered as variables and were estimated by the GARCH model. The combined model, ARIMA/GARCH, was established. The following sections indicate the estimation process step by step for these models.

4.2.1. ARIMA/GARCH Model

4.2.1.1. Estimation of Parameter d

To implement the time series models described above in the natural gas spot market, we need to make sure that the given time series is stationary. For non-stationary time series we could apply some linear or non-linear transformations to achieve a stationary trend. As the first step, the autocorrelations (ACF) and partial autocorrelation functions (PACF) of the time series should be calculated which indicate the stationary or non-stationary behavior of the variable. The second step is to use standard tests such as KPSS (Kwiatkowski, Phillips, Schmidt, & Shin, 1992), and augmented Dickey-Fuller tests (Perron, 1988) in order to confirm the stationary behavior of the variable statistically. To convert a non-stationary time series into a stationary one, we may make a new time series in terms of differences of one lag. This process defines the differencing parameter “ d ” in ARIMA model.

As shown in Figure 32, the autocorrelation functions for the original time series, HH spot price, do not converge to zero and are significant for large number of lags. They decay into the range very slowly and could not reach the domain even after 38 lags but in the partial correlation graph after the second lag they become significantly small. It can be inferred from this behavior that this variable is not stationary. Therefore, a new time series was generated by differencing natural gas price, which is illustrated in Figure 33.

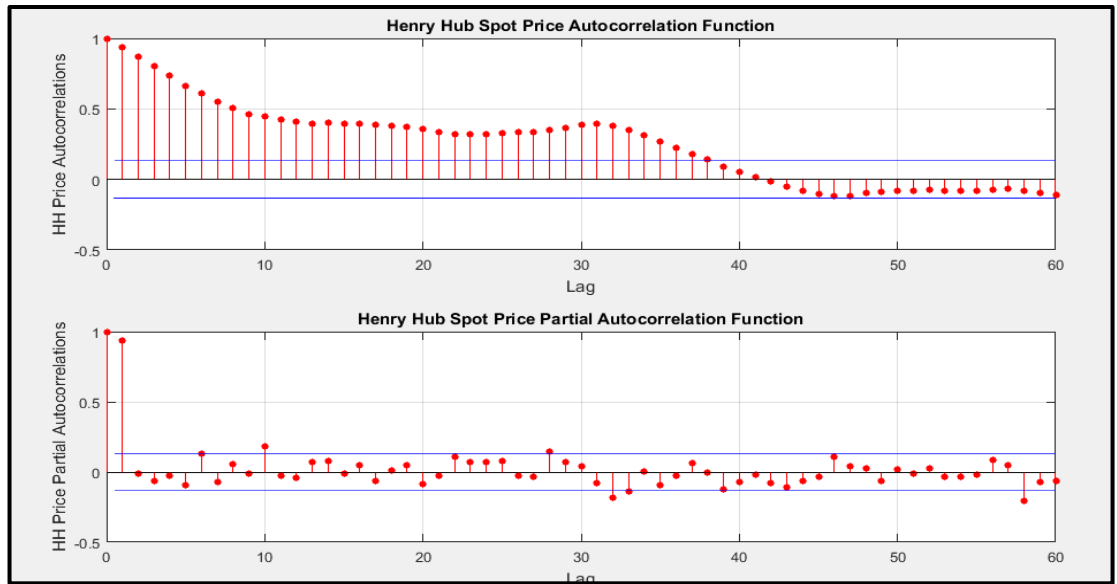


Figure 32. Autocorrelation and Partial Autocorrelations functions for Henry Hub Spot Price

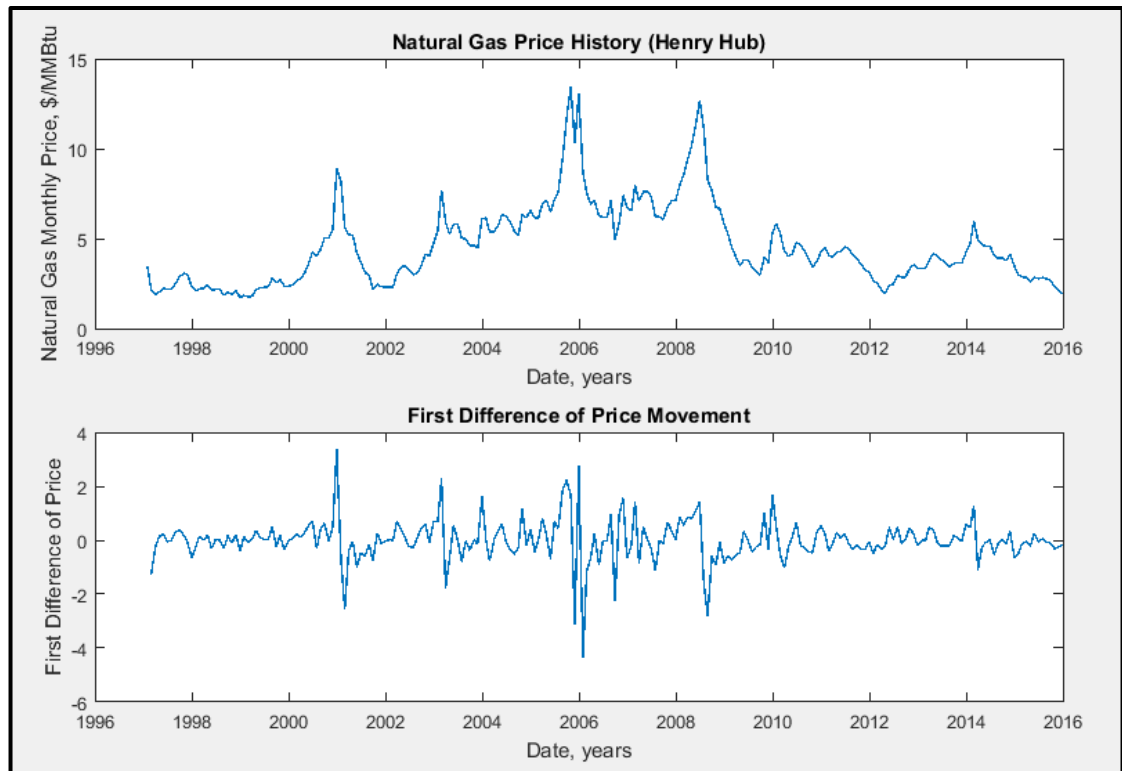


Figure 33. Natural Gas Spot Price (a) and First Difference of Price (b) Historical Data

Figure 35 reveals that the new time series produced by differencing of natural gas price time series is stationary. The KPSS and ADF test results summarized in Table 2 also confirm the explanations of the AFC and PAFC graphs that the spot price trend is not a stationary time series but its first difference is. Therefore, it is suggested to use the first difference (i.e. $d = 1$) in order to establish an ARIMA-GARCH model.

Table 2. Result summary for stationary tests

Variable	Test	Stat.	Crit. Value	Likelihood Log	Significance Level	MSE
Spot Price	KPSS	3.26	0.1460	-508.48	95%	5.110
	ADF	-1.25	-1.9421	-271.49	95%	0.643
First Difference of Price	KPSS	0.03	0.1460	-272.115	95%	0.612
	ADF	-15.42	-1.9421	-270.200	95%	0.643

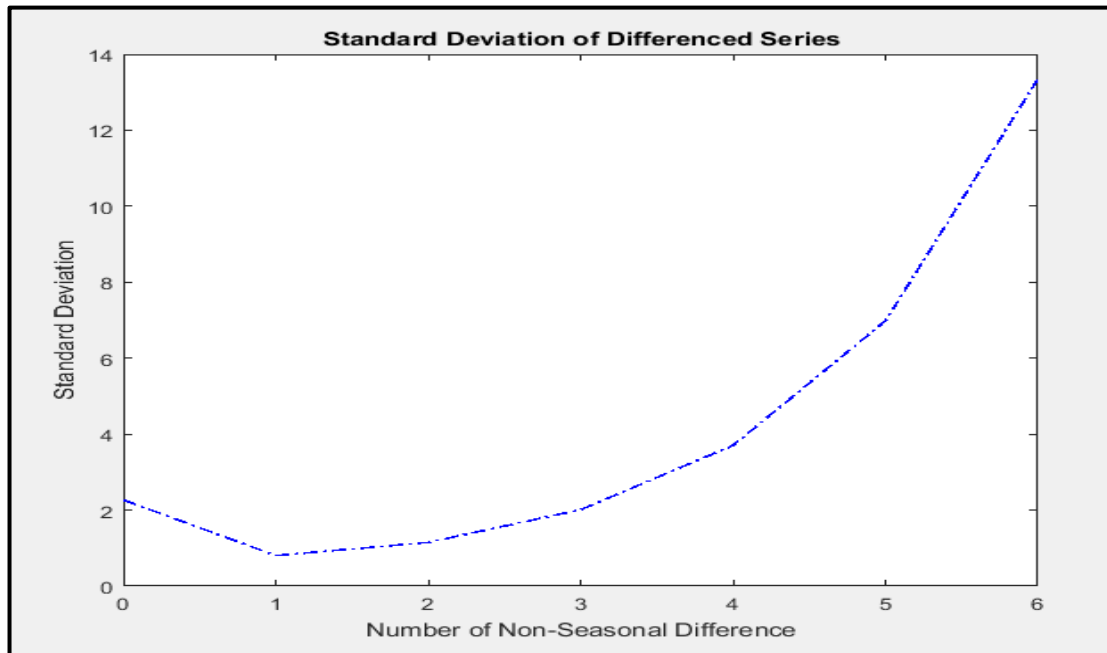


Figure 34. Standard Deviation of Differenced Time Series

Another method for estimation of the optimum number of differences is to find the minimum standard deviation of the produced time series, which is demonstrated in Figure 34. This method also suggests using $d = 1$.

4.2.1.2. Estimation of Parameters r and m

Figure 35 which illustrates the ACF and PACF for the model variable indicates that the 5th and 9th lags of the time series have a close relationship with the variable. The negative values of these significant lags in PACF denote that a slightly over-differenced variable is occurred which could be corrected by considering the lags as MA lags. Therefore, ARIMA (5,1,9) was chosen as a descriptive model in which the coefficient for the lags are zero expect the 5th and 9th lags.

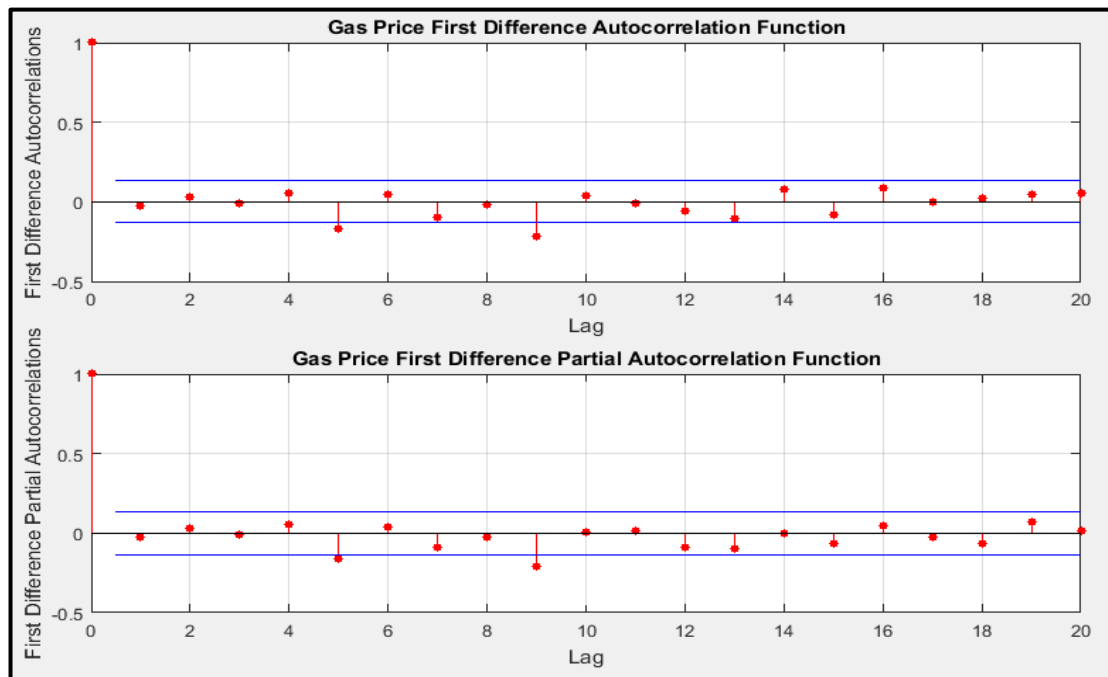


Figure 35. Autocorrelation and Partial Autocorrelations functions for First Difference of Spot Price

In order to find the optimum parameters of the ARMA model, various combinations of the lags were defined, and then Akaike and Bayesian information criteria were examined for those to choose the best model (Box, Jenkins, & Rein, 2015). According to the AIC results in Table 3, the ARIMA (9, 1, 9) model with the 5th and 9th lags in both AR and MA section is the optimum model. On the other hand, the BIC results in Table 4 suggest choosing ARIMA (5, 1, 9) model with the 5th lag in AR and the 9th lag in MA section. As the results of both methods for these two models are very close, ARIMA (5, 1, 9) was preferred to be selected which has two parameters less than the other model. A simpler model is always better if the accuracy is not decreased much.

Table 3. Akaike Information Criteria Results for Different Lags in ARIMA Model

		MA Lags			
		0	5	5, 9	9
AR Lags	0	549.92	545.76	536.65	539.56
	5	545.39	545.25	536.20	535.41
	5, 9	536.23	538.10	531.86	535.16
	9	540.16	537.04	537.46	542.08

Table 4. Bayesian Information Criteria Results for Different Lags in ARIMA Model

		MA Lags			
		0	5	5, 9	9
AR Lags	0	556.77	556.04	550.36	549.85
	5	555.68	558.97	553.34	549.13
	5, 9	549.94	555.24	552.44	552.31
	9	550.45	550.76	554.61	555.80

4.2.1.3. Estimation of ARCH and GARCH parameters

As shown in Figure 36, the variance of first difference of the natural gas price is not constant and varies by time. These changes in statistical parameters led us to define a GARCH model for fitting the variances.

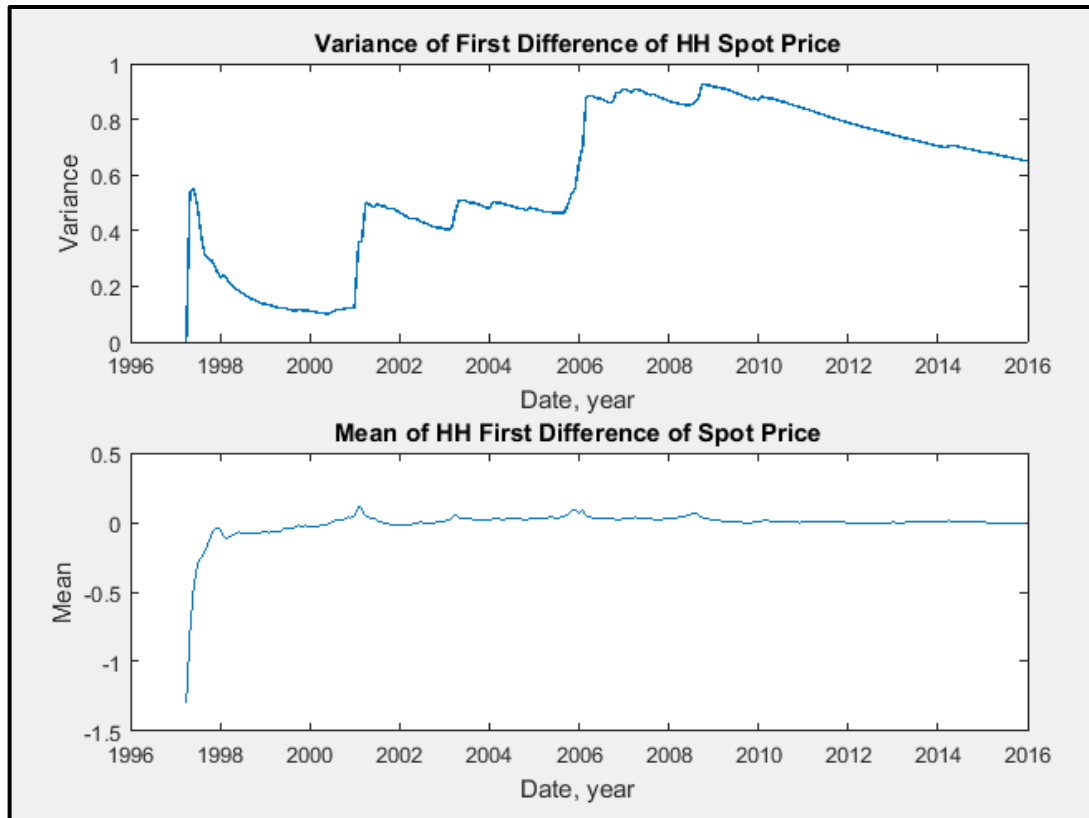


Figure 36. Variance and Mean of Time Series during Time

The first approach is interpretation of ACF and PACF of the modeling variable. Figure 37 indicates that a significant arch effect can be seen in 2 lags of the modeling variable. The number of lags is the summation of ARCH and GARCH lags ($p + q$) together. GARCH (1, 1) was defined for variance modeling and now our model is completed as ARIMA (5, 1, 9) / GARCH (1, 1). There are some statistical tests such as

Engle's test that indicate the conditional heteroscedasticity of the time series (Engle, 1982). This test approves the obtained results as well.

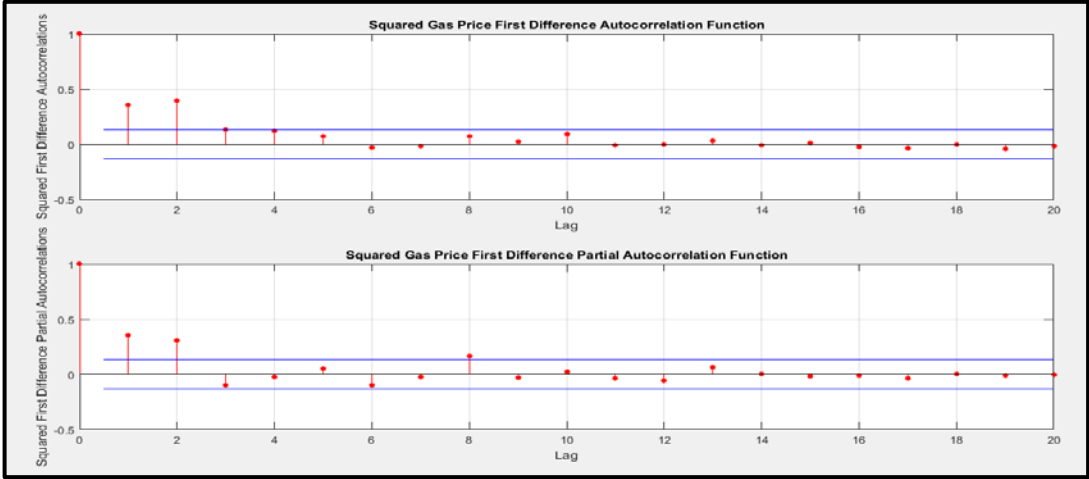


Figure 37. ACF and PACF of Squared First Difference of HH Spot Price

Table 5 shows the model parameters' values, standard errors, and t-statistic associated with the parameters. Plugging the parameters in the models results in:

$$\left\{ \begin{array}{l} X_t = X_{t-1} - 0.0279 - 0.0567 (X_{t-5} - X_{t-6}) + a_t + 0.2205 a_{t-9} \\ a_t = \sigma_{t|t-1} \cdot \varepsilon_t \\ \sigma_{t|t-1}^2 = 0.0499 + 0.5288 \sigma_{t-1|t-2}^2 + 0.4712 a_{t-1}^2 \end{array} \right. \quad \text{Eq. 34}$$

Table 5. ARIMA/GARCH model's parameters

Parameter	Value	Std. Error	t-Statistic
ARIMA Const.	-0.0279	0.0230	-1.2162
AR(5)	-0.0567	0.0447	-1.2677
MA(9)	-0.2205	0.0344	6.4023
GARCH Const.	0.0499	0.0204	2.4433
ARCH(1)	0.4712	0.0725	6.5023
GARCH(1)	0.5288	0.0653	8.0958

To validate and check the model, the model was run, starting from 1999 until the last date of the available data. The model was established for the early data and then was rebuilt for the emerging data over time for forecasting the price in one step ahead. The results are illustrated in Figure 38. In Figure 38(a), the dotted red line is the forecasted prices and the solid blue line is the actual historical data. In Figure 38(b), the residual values are illustrated, which indicate that the model is outperforming when new data are coming into the model. The residual values are converging to zero over time although the model has poor performance at the spikes.

Finally, the model was run to predict the price for 12 months in the future with the confidence level of 95%. The results are demonstrated in Figure 39. The prediction shows a descending trend until it reaches the lowest point at a price of 2 \$/MMBtu in May 2016 and then starts to move upward until it hits the price of 2.3 \$/MMBtu in September 2016. It also shows that the natural gas price will fluctuate between 1.5 and 3.2 \$/MMBtu in a 95% level of confidence.

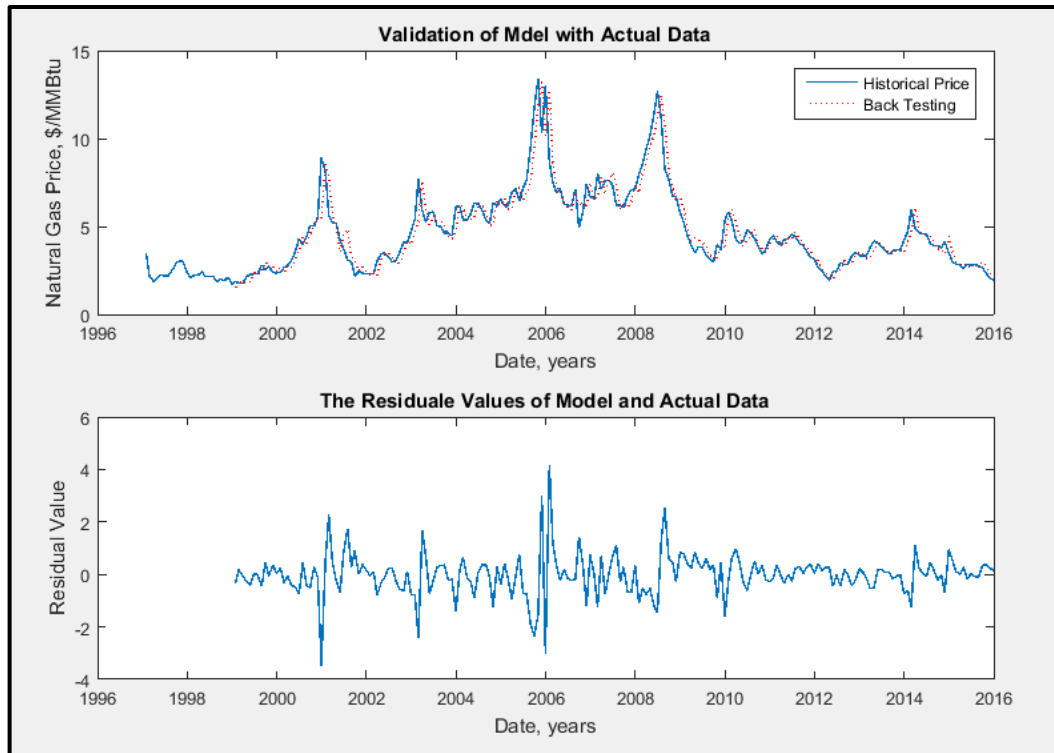


Figure 38. (a) Model Back Test Results, (b) Residual Value for Model

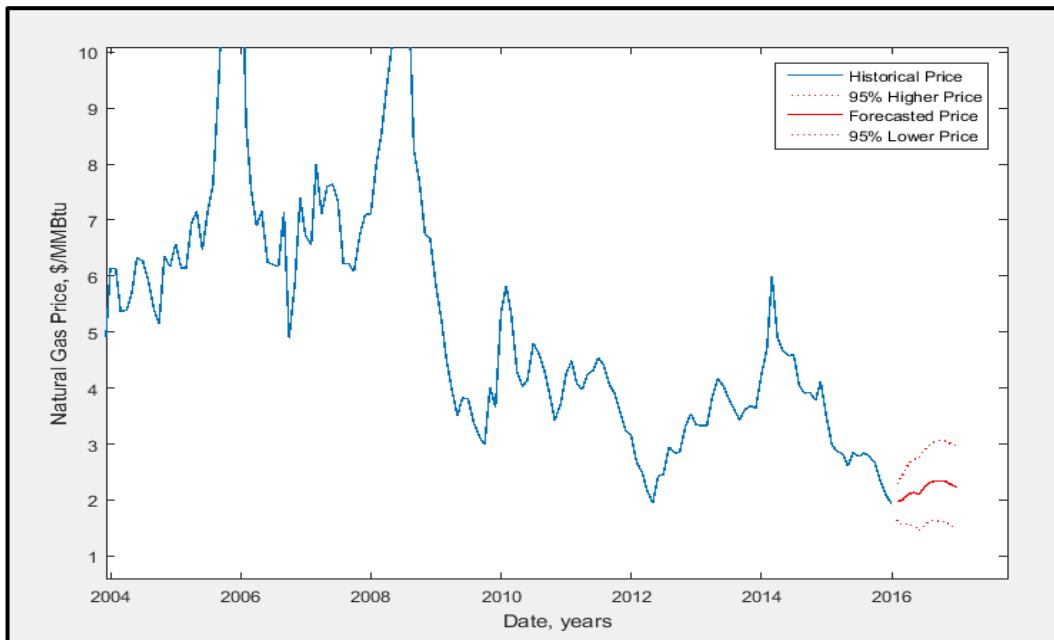


Figure 39. Model Forecasting for 12 Month in Future

4.3. Neural Network Models

In this section, univariate nonlinear auto-regression neural network model (NAR) and multivariate nonlinear auto-regression neural network model (NARX) have been examined to describe the natural gas price behavior. 70% of the data is considered as the training dataset while the remaining 30% of data was equally reserved for testing and validation.

4.3.1. Univariate Neural Network Autoregression (NAR) Model

To determine the number of lags, a model including one lag and one hidden layer was built. The autocorrelation for the error terms is demonstrated in Figure 40. These results are consistent with the obtained results in the ARIMA Modeling section. Therefore, 9 lags were selected to eliminate the auto-correlated errors from the diagram and the model was run for new parameters. Figure 41 shows that 9 lags could not capture the autocorrelation errors completely. Therefore, more lags were added to the model. Figure 42 shows that 10 lags for this model could be appropriate enough and would capture the autocorrelation effect on the errors.

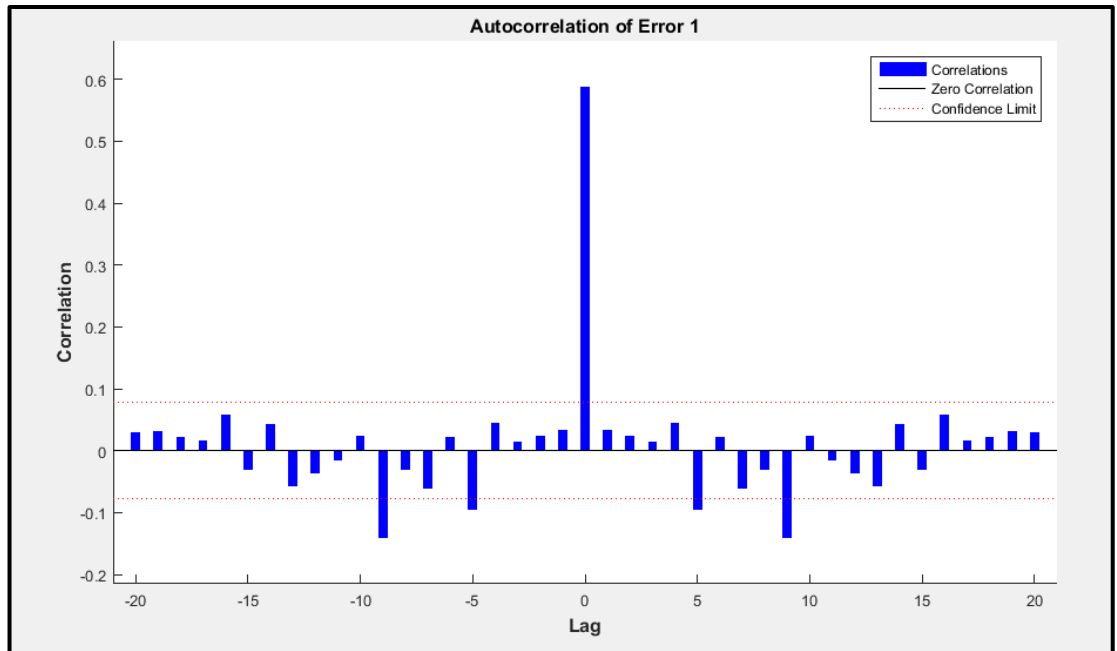


Figure 40. Autocorrelation of error for 1-layer 1-lag NAR model.

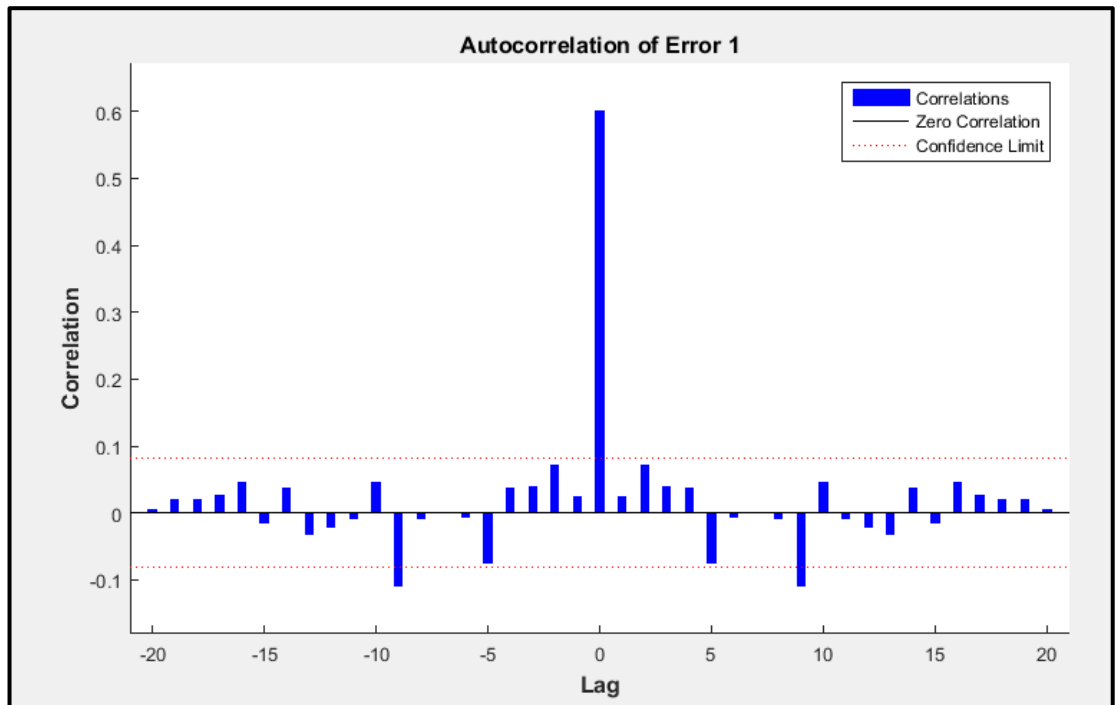


Figure 41. Autocorrelation of error for 1-layer 9-lag NAR model.

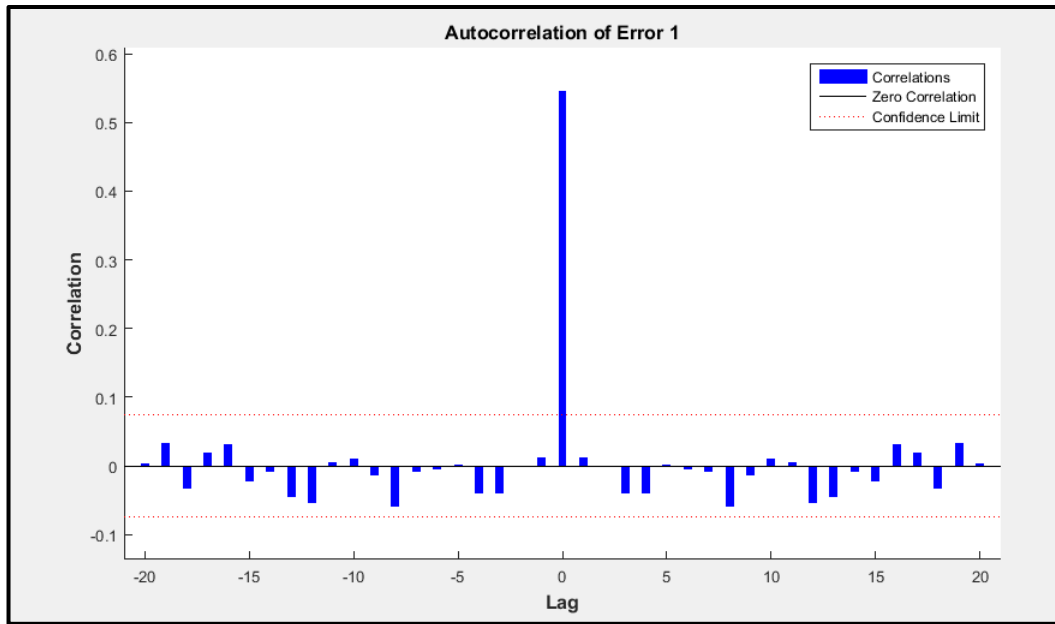


Figure 42. Autocorrelation of error for 1-layer 10-lag NAR model.

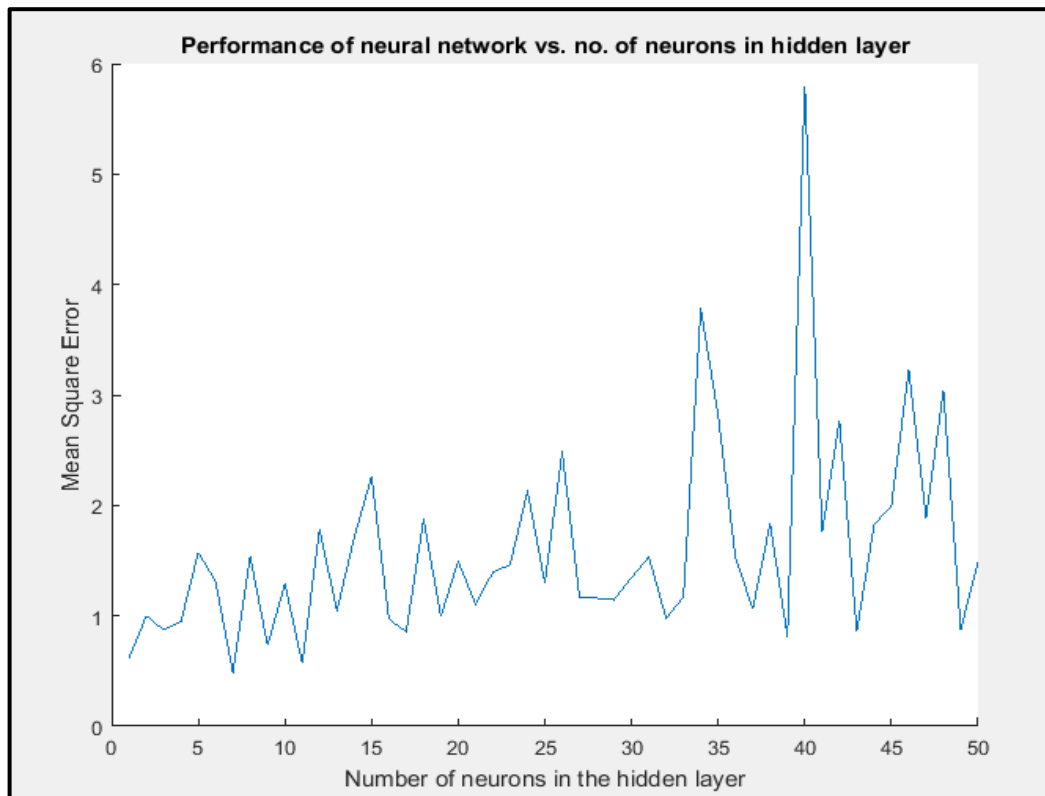


Figure 43. Performance of neural networks vs. number of neurons in the hidden layers for NAR model.

To determine the number of neurons in the hidden layer for neural networks, the model was run for 1-neuron to 50-neurons in the hidden layer, the MSE for each layer has been plotted in Figure 43. The plot shows that adding more neurons in the hidden layer does not improve the results considerably. Therefore, one neuron model was selected as the describing model for this section. A model with 6 neurons in the hidden layer shows slightly better results but the enhancement is not good enough to choose. The Figure 47 through Figure 49 show the results for 10-6-1 neural networks. The results are almost similar to the outputs for 1-layer model.

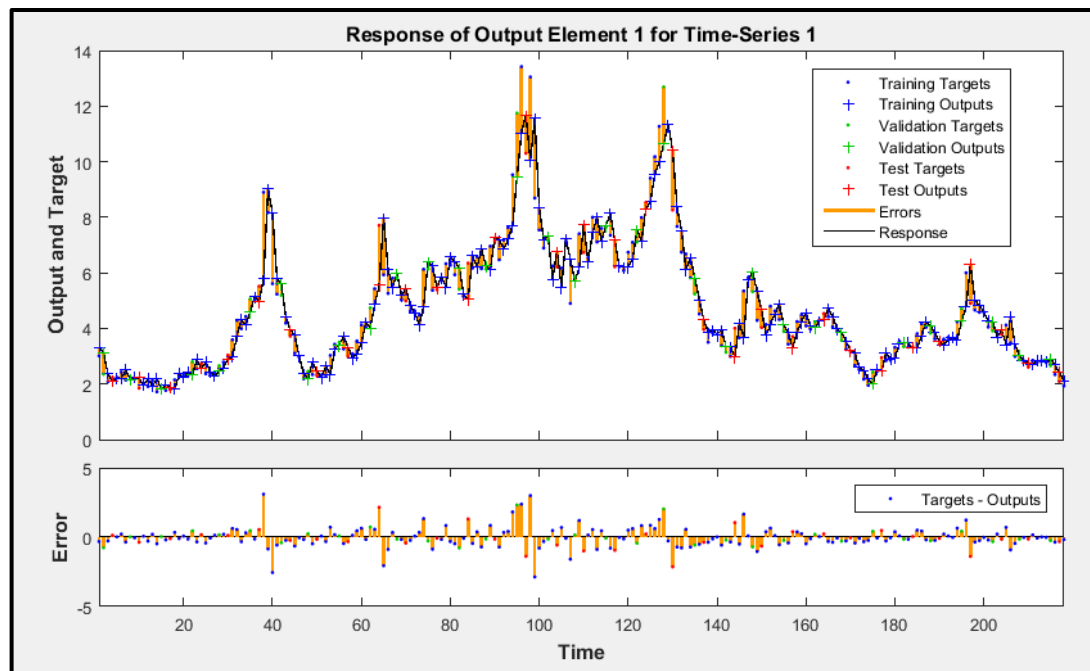


Figure 44. Back testing of the NAR model with 10 lags and 1 neuron in the hidden layer

The response is plotted to check if the model is appropriate enough to capture the nonlinear behavior of the time series. The model simulates the movements very well unless a rapid change is occurred. The error diagram shows the absolute errors during the simulation time.

Figure 45 illustrates the regression between the target values and the simulated values for the training, testing, and validation data sets. In terms of regression lines, the simulated values show very promising results. The R^2 values for the training, testing, and validation data are close which means that the model is not over fitted. This result is also confirmed by the performance trend of the model for these data shown in Figure 46.

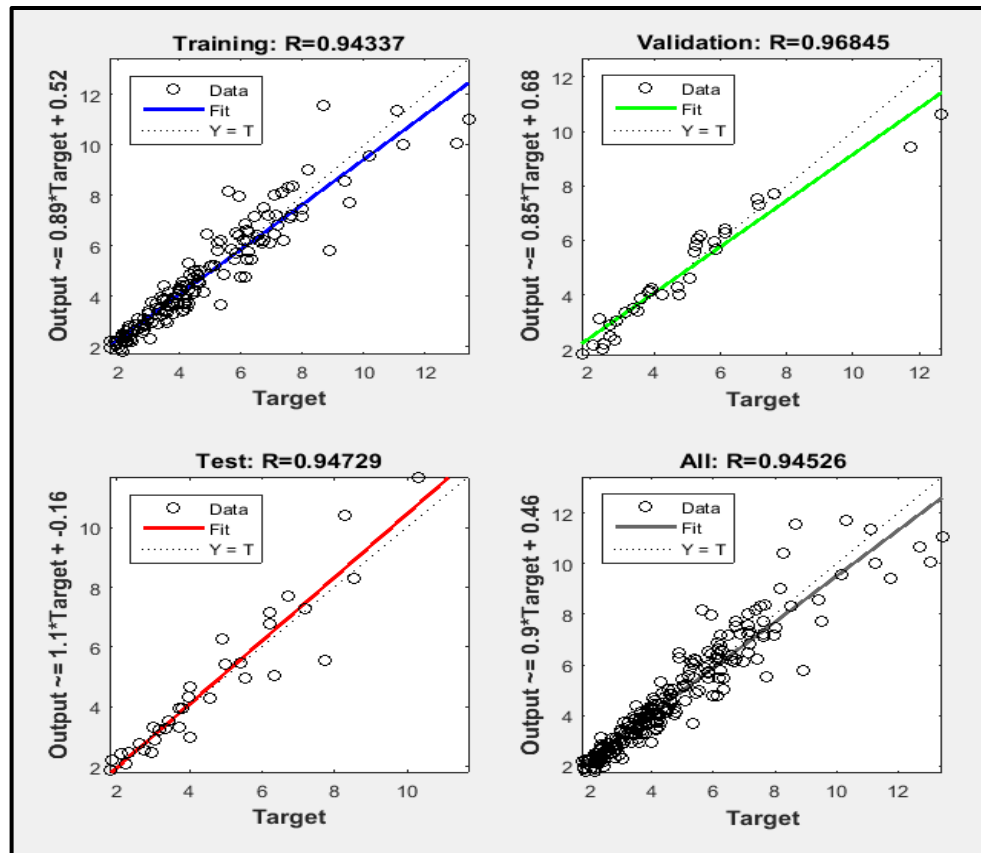


Figure 45. Regression results and R squared for 10-1-1 NAR model.

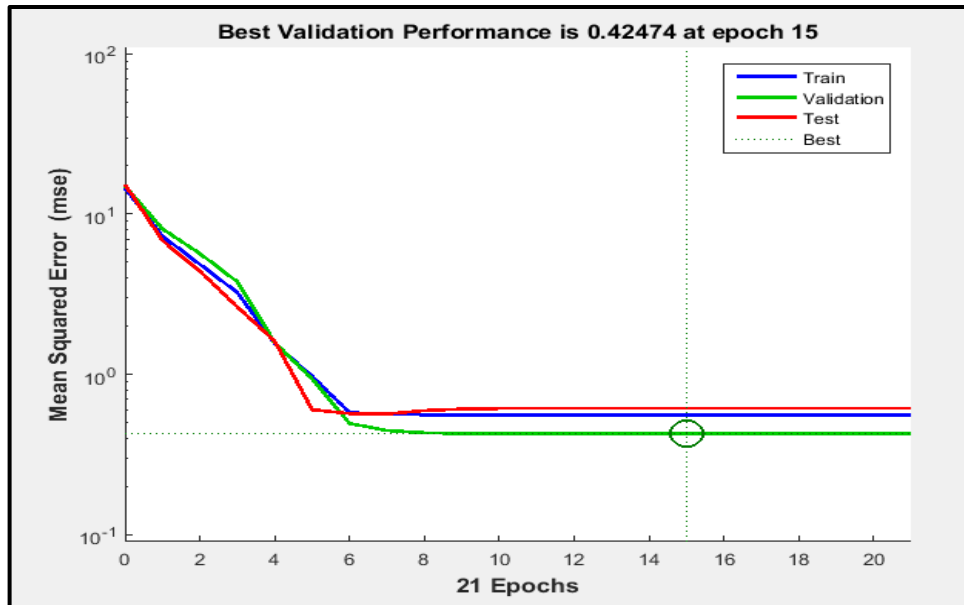


Figure 46. Performance progress for 10-1-1 NAR model.

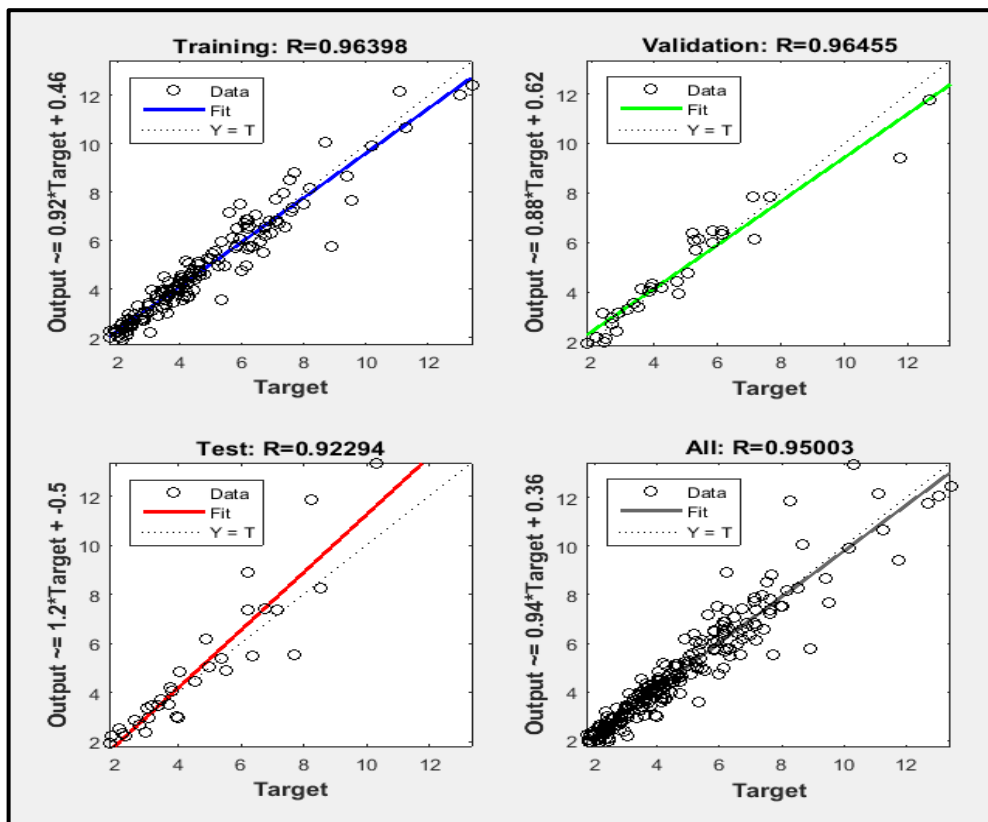


Figure 47. Regression results and R squared for 10-6-1 NAR model.

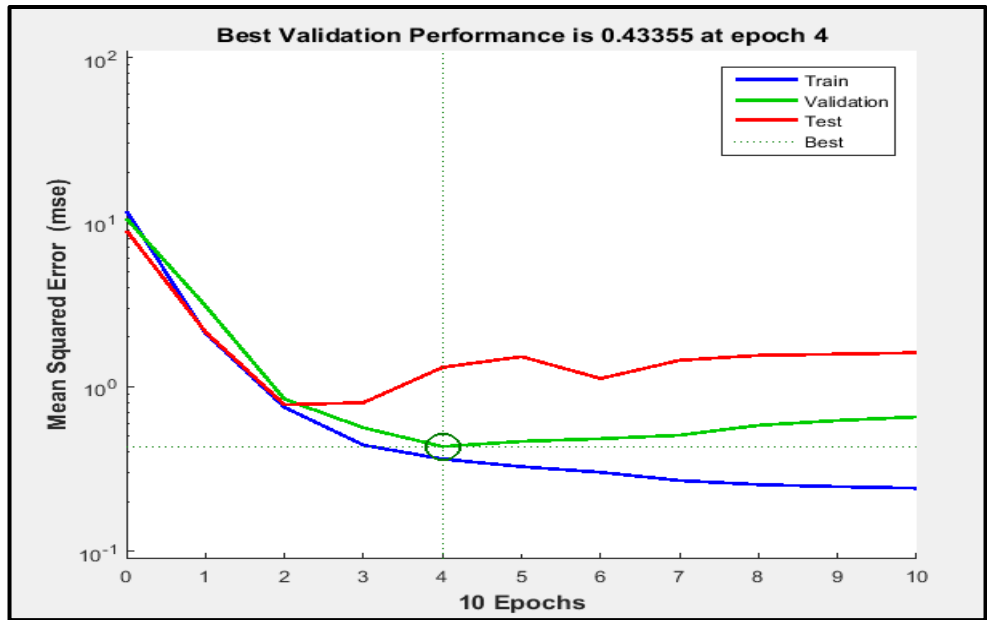


Figure 48. Performance progress for 10-6-1 NAR model

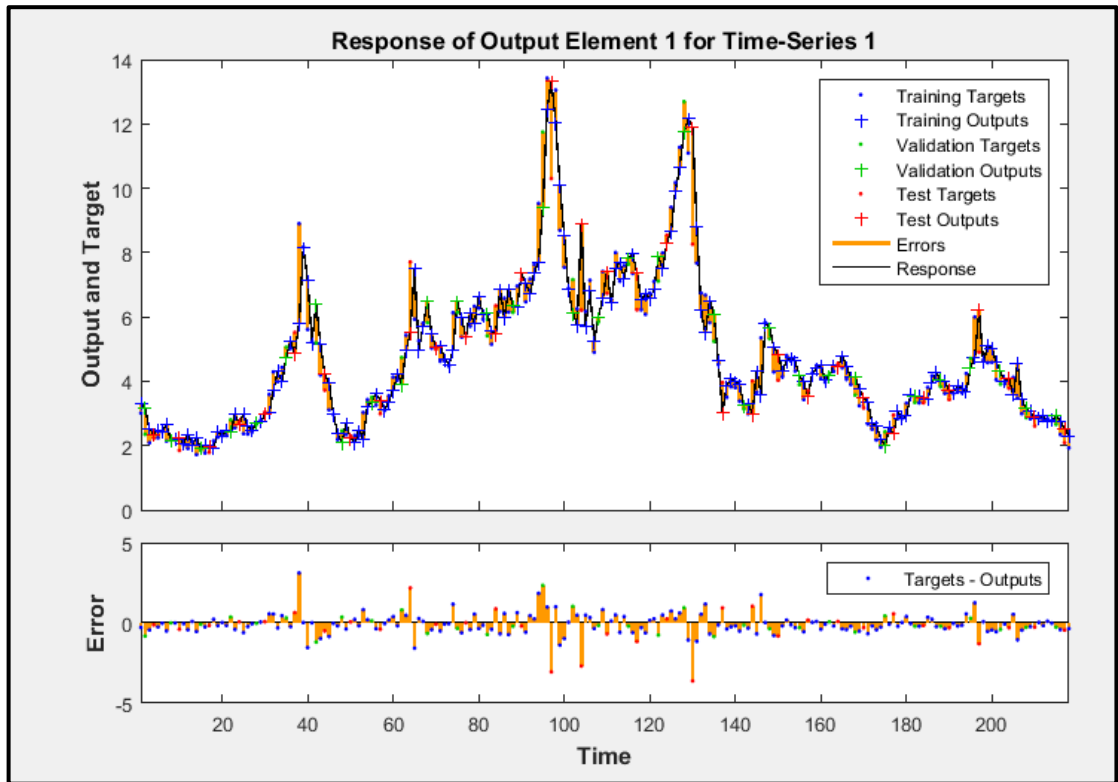


Figure 49. Back testing of the NAR model with 10 lags and 6 neurons the hidden layer

4.3.2. Multivariate Neural Network Autoregression (NARX) Model

The last model that was created is a NARX model; A neural network time series model which contains exogenous variables to justify the responses. The variables are West Texas instrument oil price, total gas production, total gas consumption, storage capacity, storage volume, injection to and withdraw from the storages, cooling days, heating days, extreme minimum temperature, extreme maximum temperature, and mean temperature. As performed before for the NAR model, to determine the numbers of lags for gas price, an autocorrelation between the lags is plotted in Figure 50. The plot leads us to choose 9 lags for this model. The starting lags for the exogenous variables were considered to be one. This investigation revealed that no more lags are required for these variables.

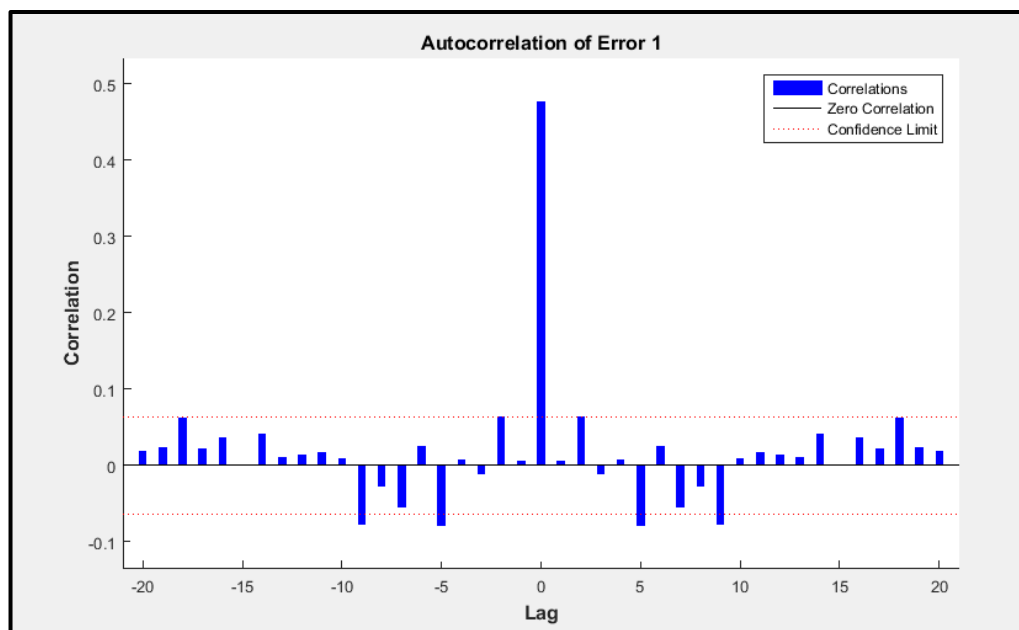


Figure 50. Autocorrelation for 1 neuron 1-lag NARX

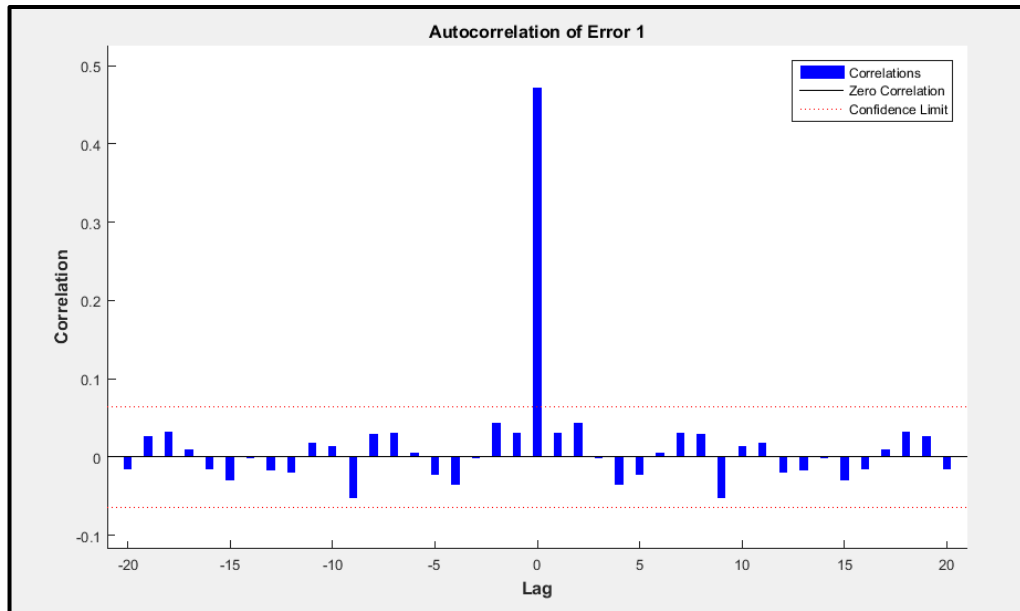


Figure 51. Autocorrelation of error for 1-neuron 9-lag NARX model.

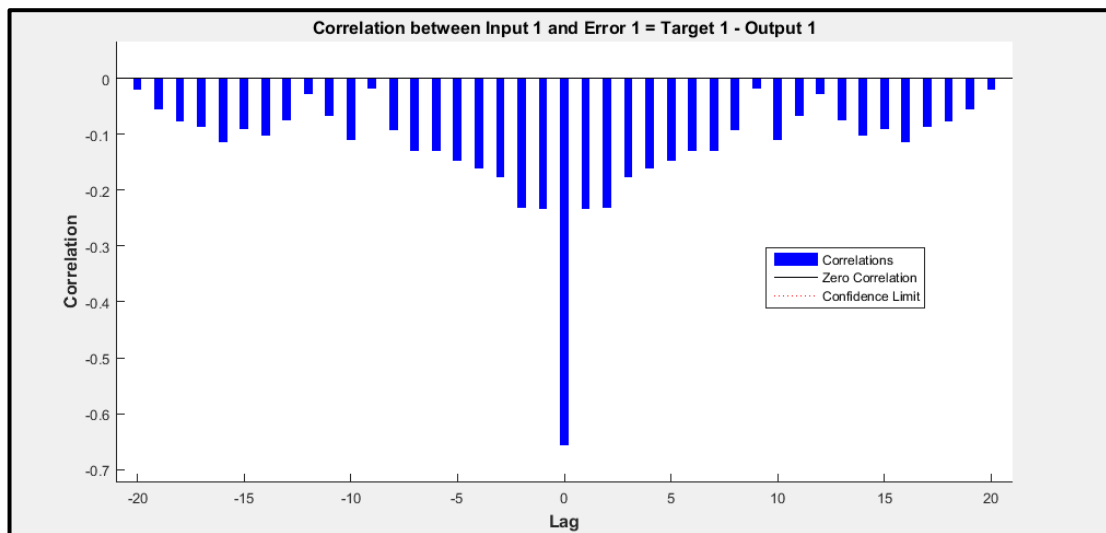


Figure 52. Autocorrelation of error for 1-layer 9-lag NARX model.

Figure 51 and Figure 52 show that 9 lags for gas price is good enough to capture the correlation between the errors. To determine the number of neurons in the hidden layers, the model was simulated for different situations plotted in Figure 53. This trend

has a very similar behaviour to the NAR model illustrated in Figure 43. The model was run for both 1 and 6 neurons in a hidden layer neural network which seems to be the best possible solution.

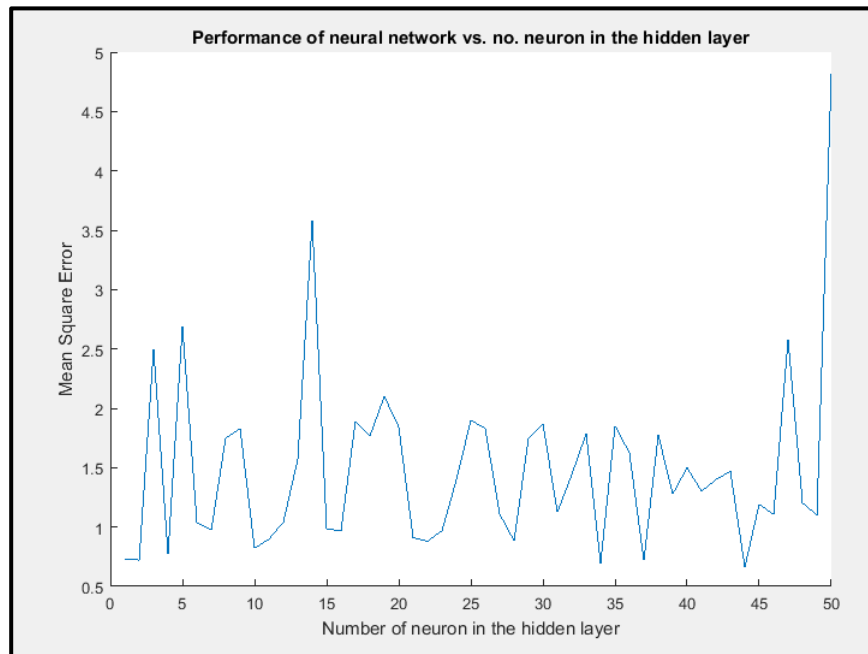


Figure 53. Performance of neural networks vs. no. of neurons in the hidden layer in NARX model.

The following figures show that the improvement by using 6 neurons in the hidden layer instead of 1 neuron is acceptable. The R^2 values for the training, testing, and validation sets demonstrate that the model is not over fitted.

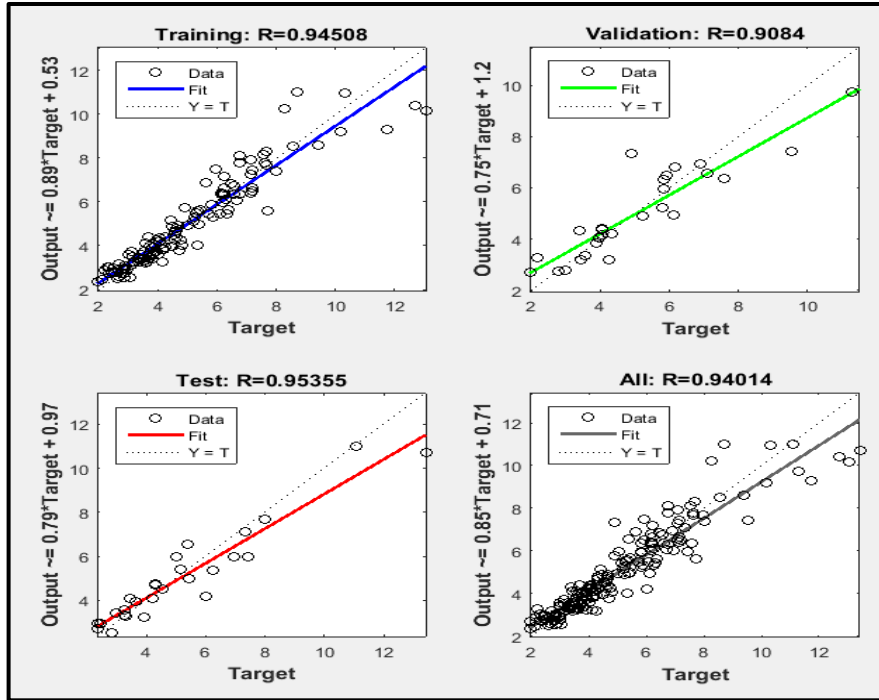


Figure 54. Regression results and R squared for 21-1-1 NARX model.

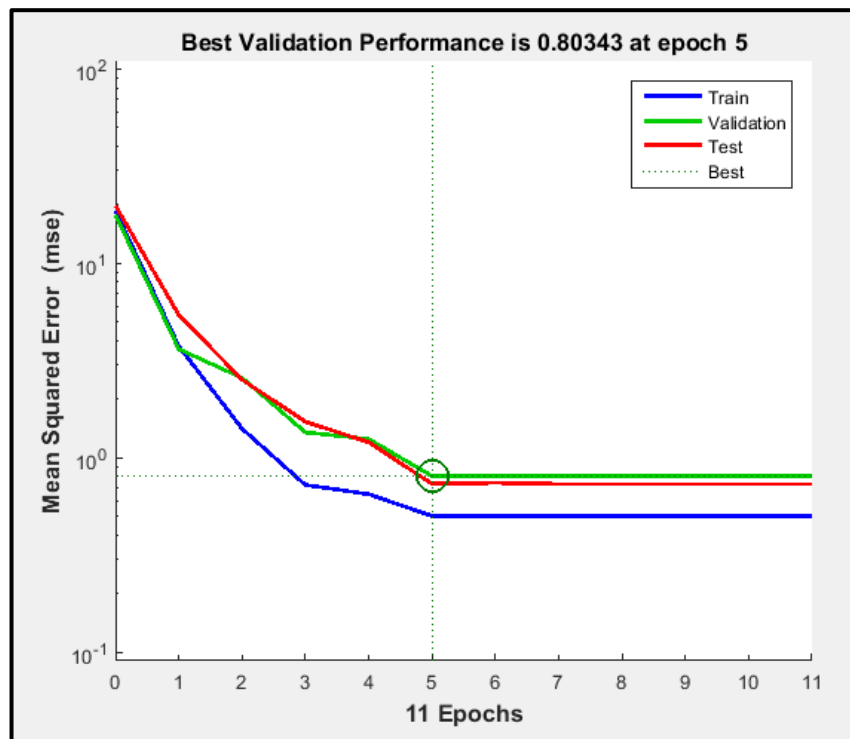


Figure 55. Performance progress for 21-1-1 NARX model

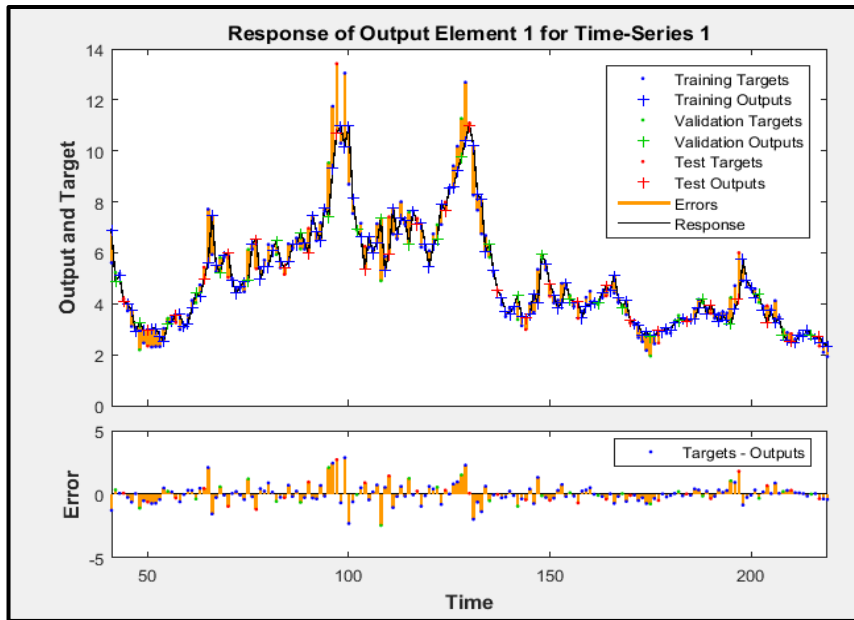


Figure 56. Back testing of the NARX model with 9 lags and 1 neuron in the hidden layer.

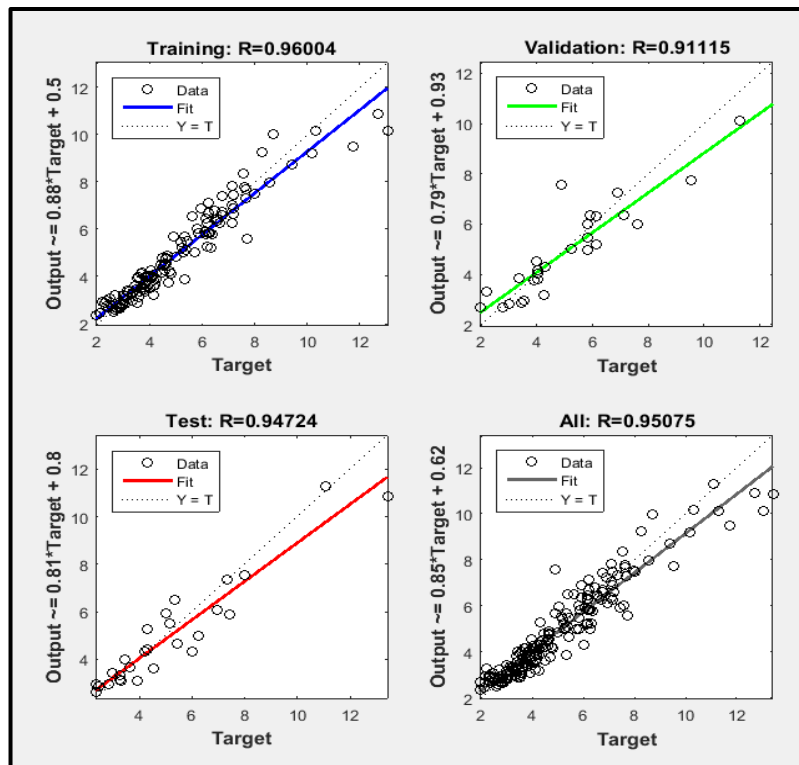


Figure 57. Regression results and R squared for 21-4-1 NARX model.

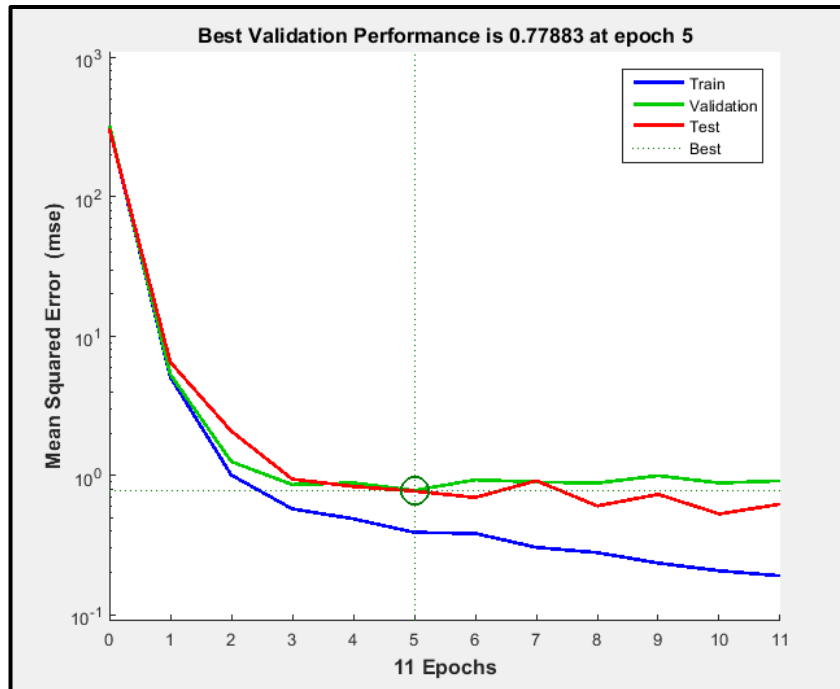


Figure 58. Performance progress for 21-4-1 NARX model

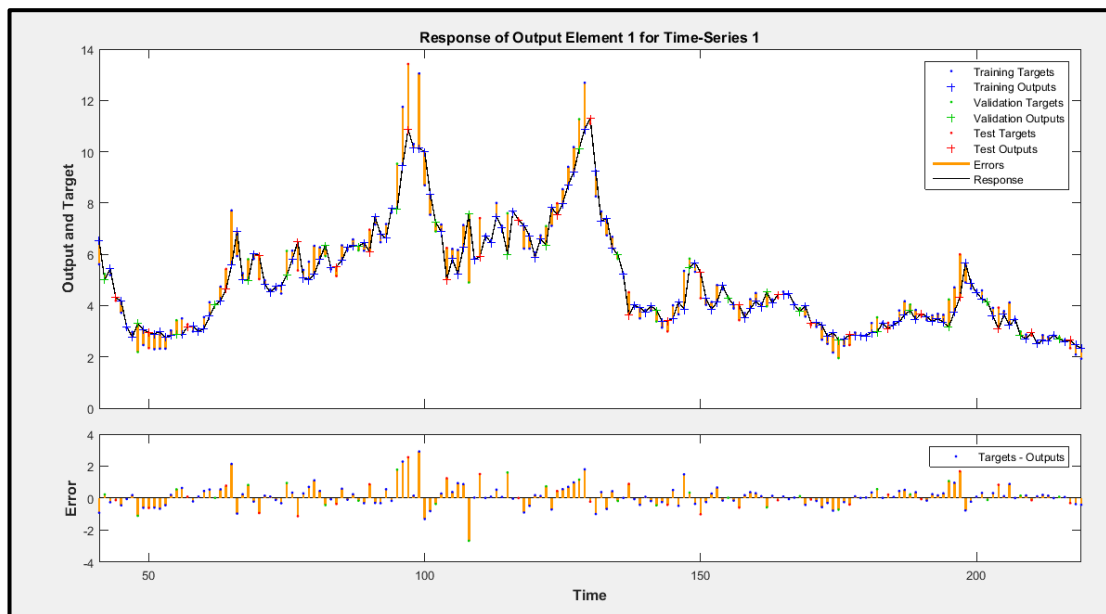


Figure 59. Back testing of the NARX model with 9 lags and 4 neurons in the hidden layer.

4.3.3. Sensitivity Analysis

In order to investigate the effect of exogenous variables on model response, a sensitivity analysis has been performed. Based on the model and real data, a base case was defined. Each variable was multiplied by factors of 0.5, 0.8, 1.2, and 2, and then introduced into the model. The following figures show the changes on the response due to the changes in the exogenous variables. A mean error was defined to determine the deviation of the response with respect to the base case based on the following formula:

$$SAE_c = \frac{1}{N} \sum (\hat{y}_b - \hat{y}_c)^2 \quad \text{Eq. 35}$$

where SAE_c is the sensitivity analysis error, \hat{y}_b is the base case response, \hat{y}_c is the case study response, and N is the number of data.

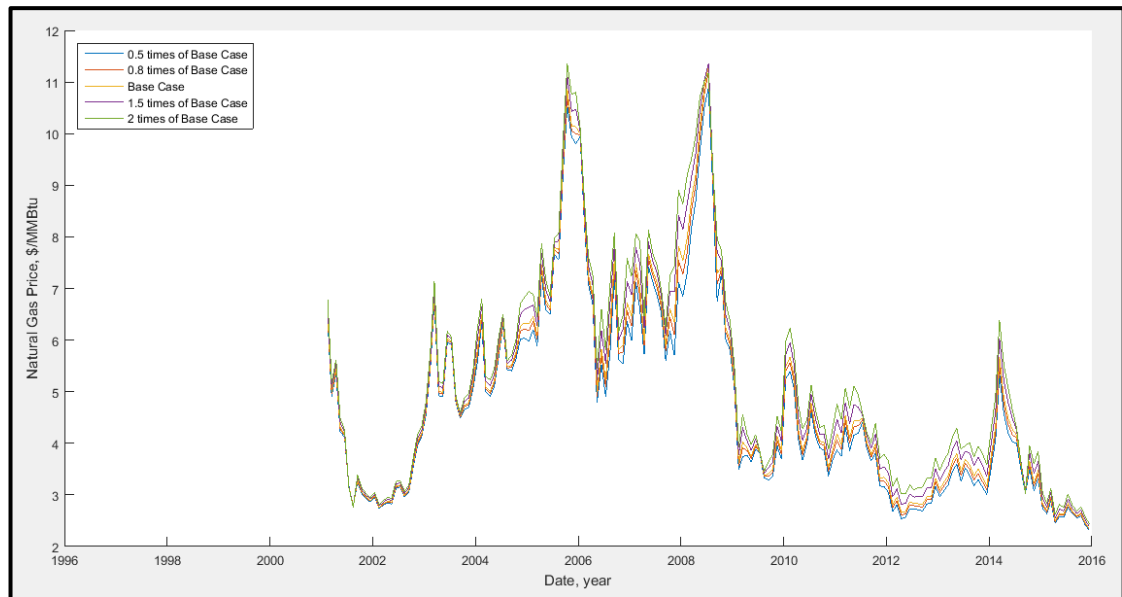


Figure 60. Sensitivity analysis for West Texas oil price.

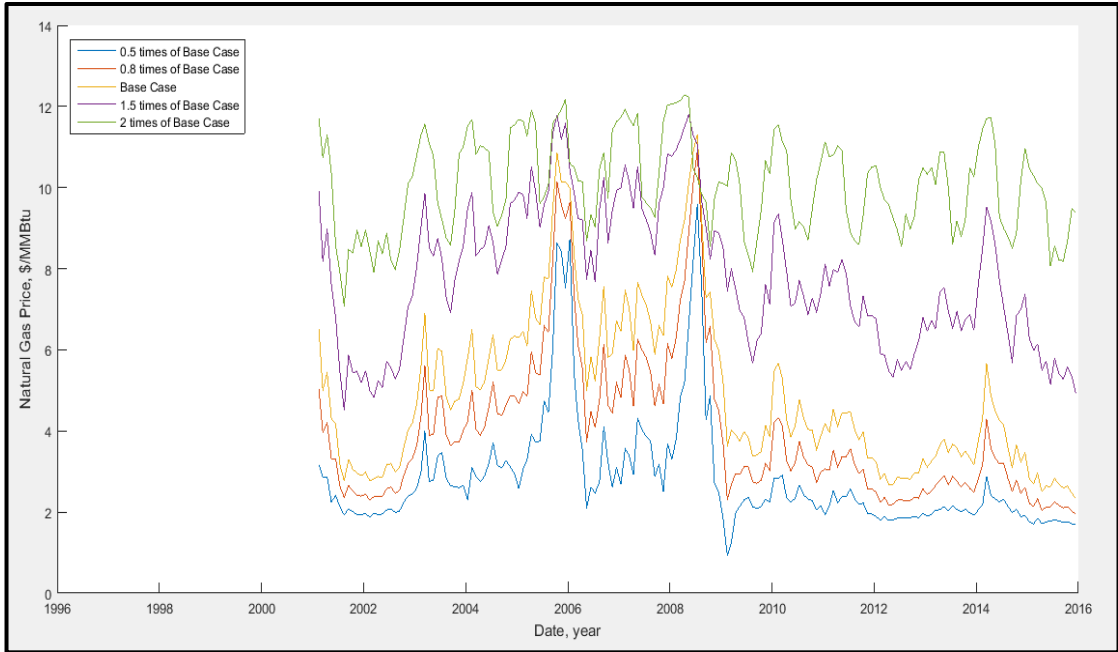


Figure 61. Sensitivity analysis for total production of natural gas.

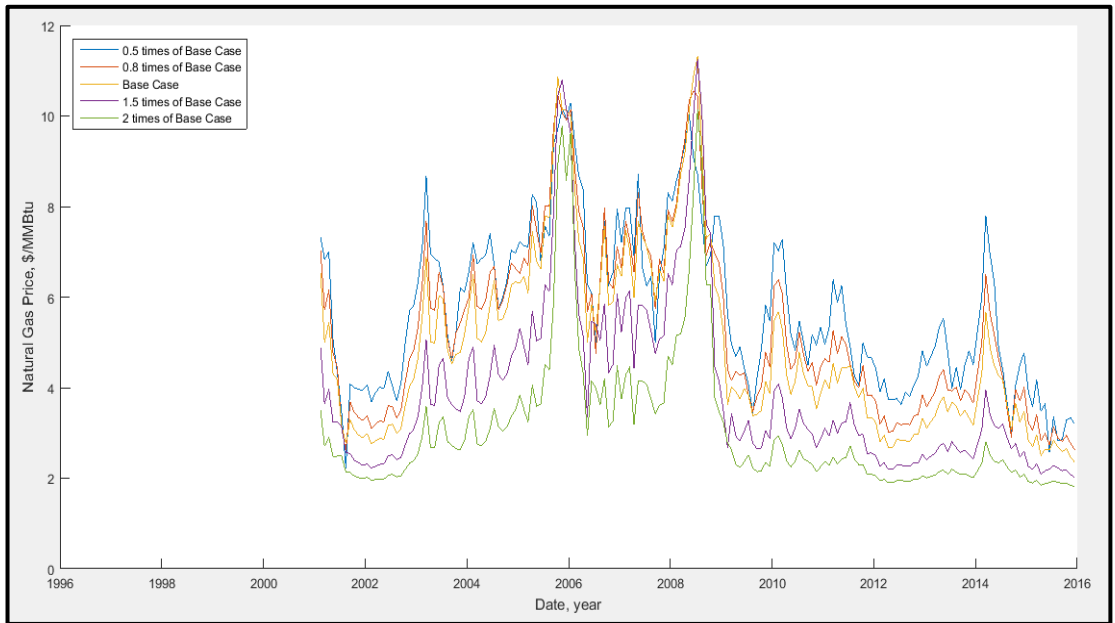


Figure 62. Sensitivity analysis for natural gas storage capacity.

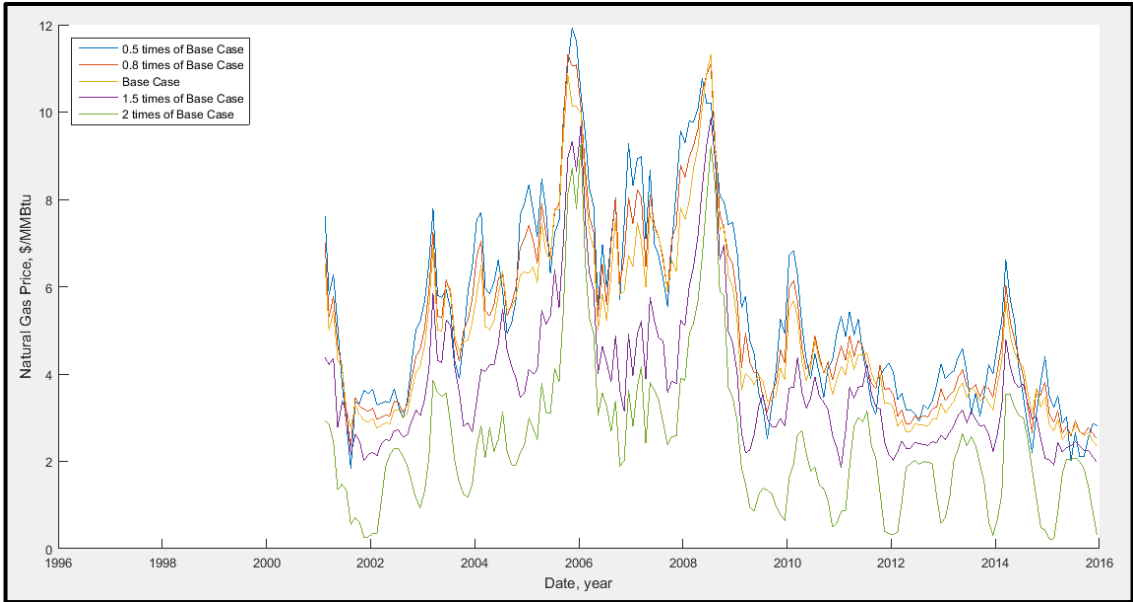


Figure 63. Sensitivity analysis for natural gas storage volume.

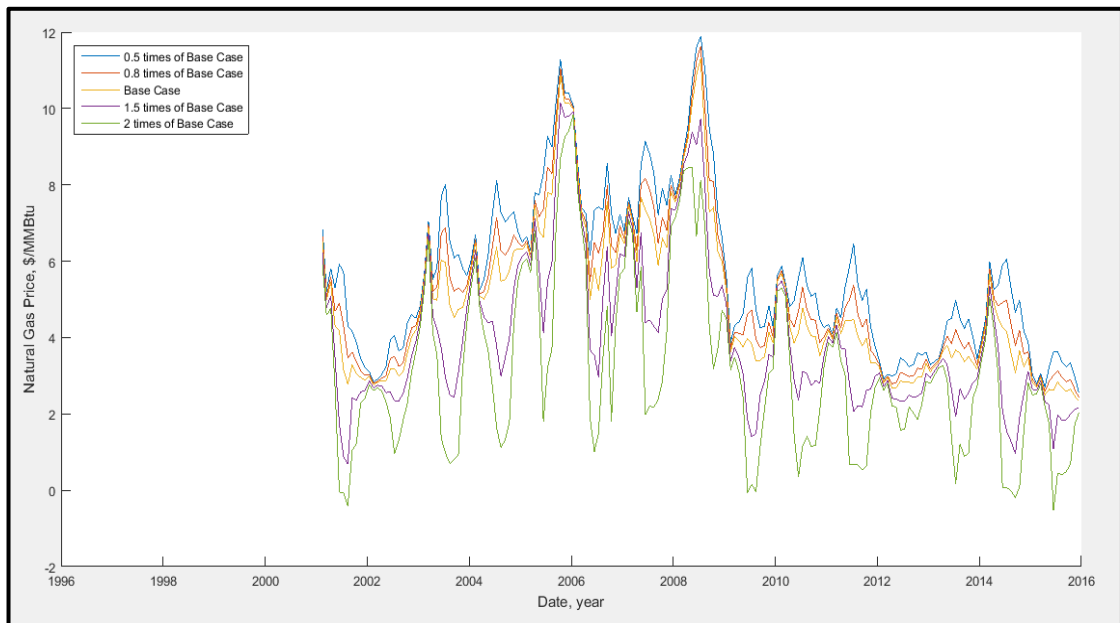


Figure 64. Sensitivity analysis for natural gas storage injection.

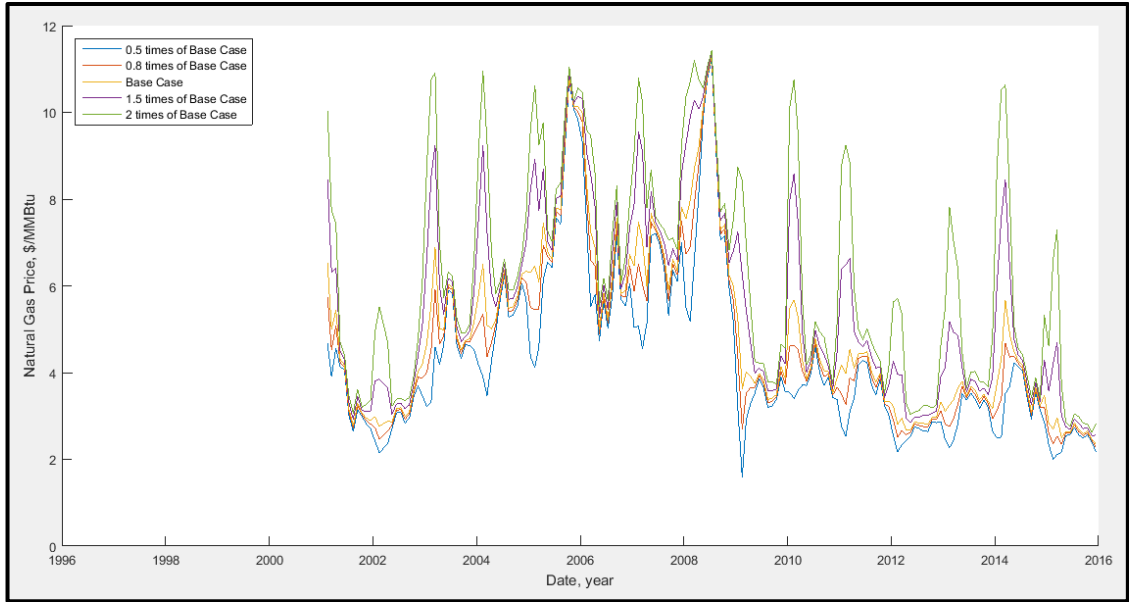


Figure 65. Sensitivity analysis for natural gas storage withdrawal

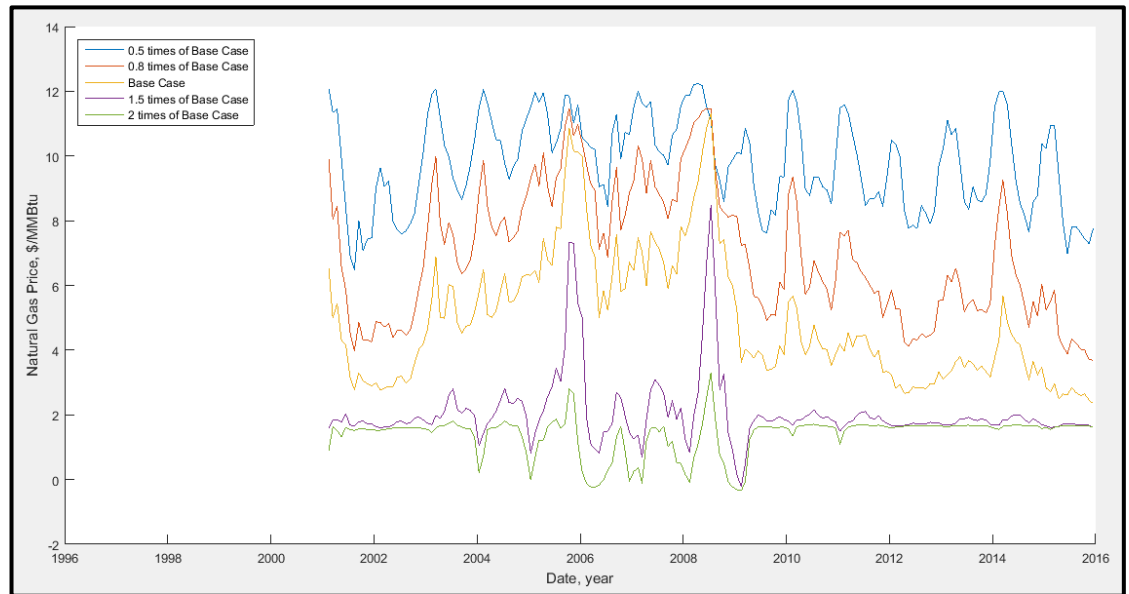


Figure 66. Sensitivity analysis for total consumption of natural gas.

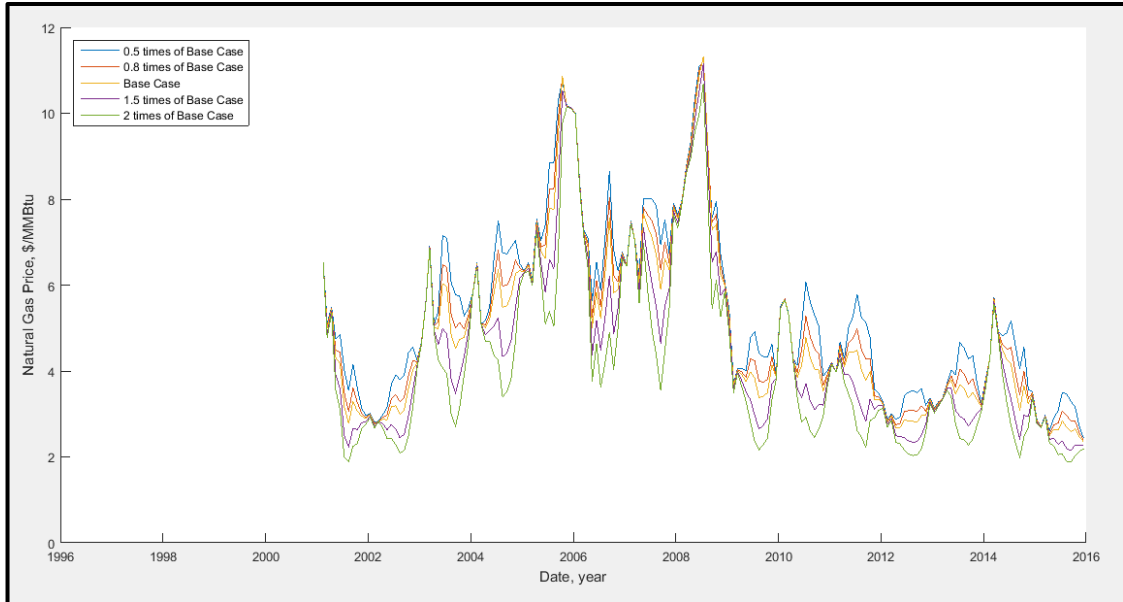


Figure 67. Sensitivity analysis for cooling degrees day Temperature.

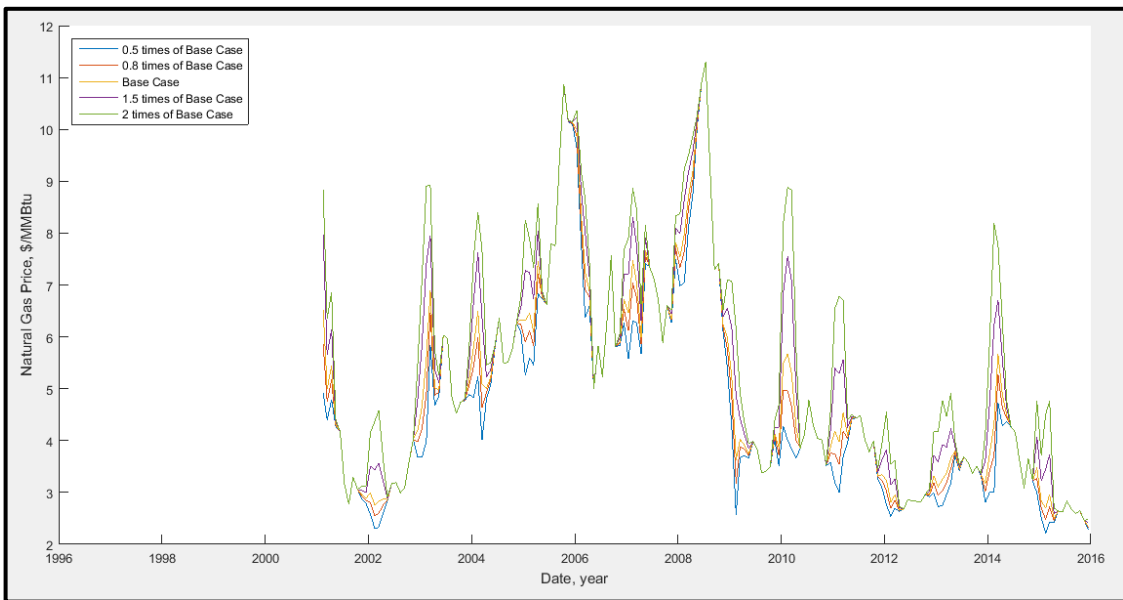


Figure 68. Sensitivity analysis for heating degrees day temperature.

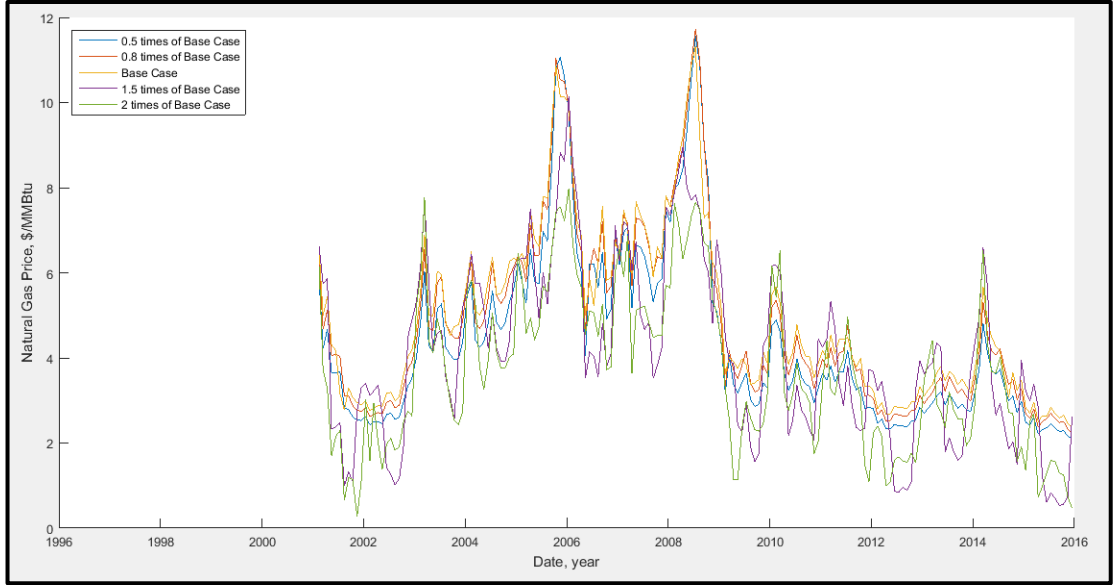


Figure 69. Sensitivity analysis for extreme maximum temperature at New Orleans, LA.

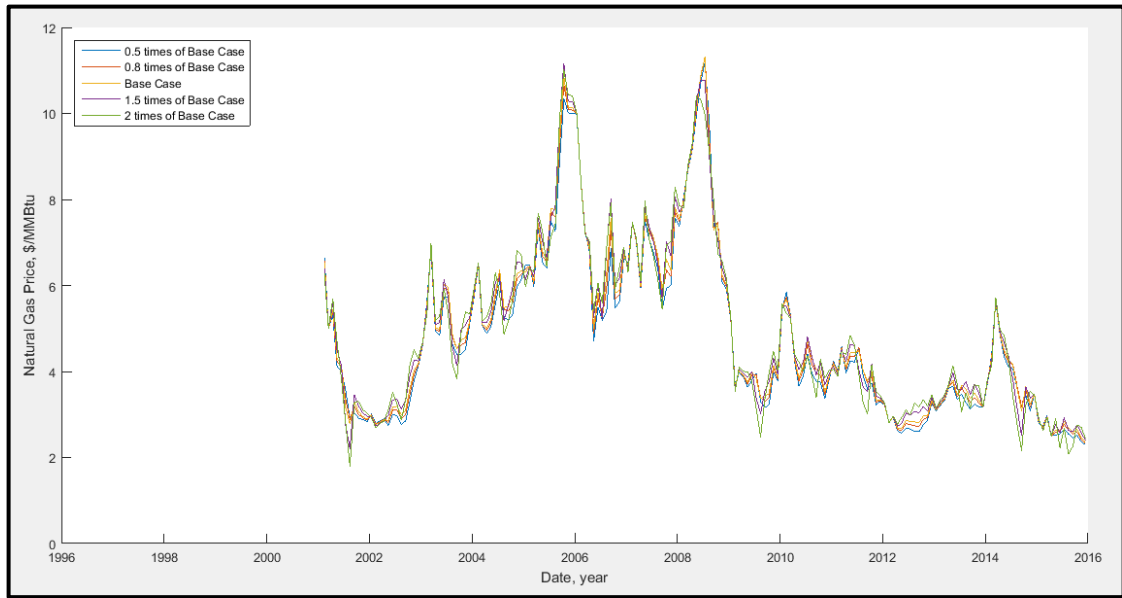


Figure 70. Sensitivity analysis for extreme minimum temperature at New Orleans, LA.

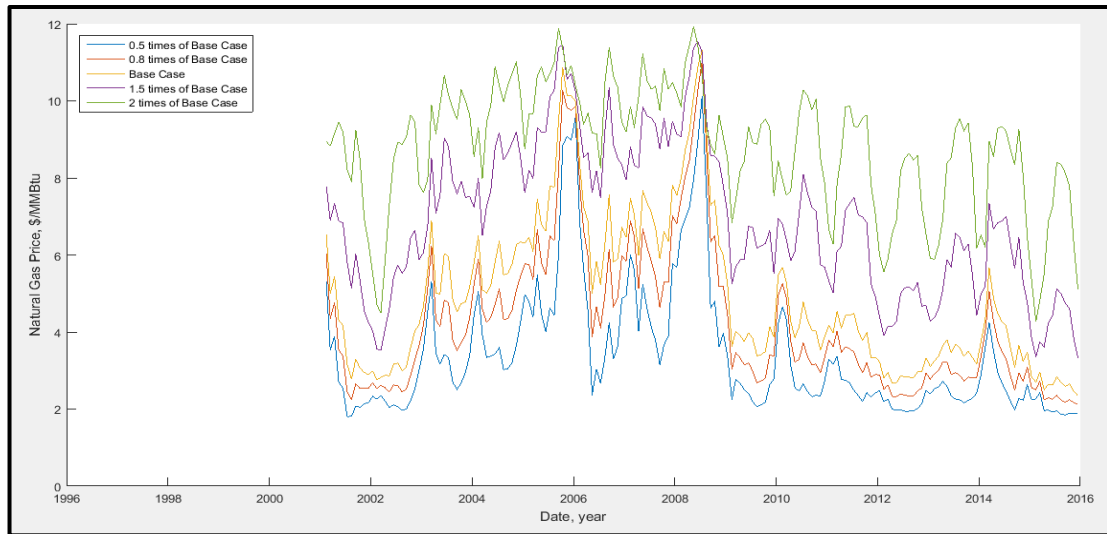


Figure 71. Sensitivity analysis for mean temperature at New Orleans, LA.

As can be seen in the figures, the most sensitive parameters are total consumption, total production, and mean monthly temperature. The least important exogenous parameters are CDD/HDD temperatures, extreme minimum temperature and WTI oil price.

Table 6. The measured mean error in sensitivity analysis

Exogenous Variables	SAE _c			
	0.5X Base	0.8X Base	1.5X Base	2X Base
WTI Oil Price	0.05	0.01	0.04	0.17
Total Production	5.22	1.05	8.88	28.93
Storage Capacity	1.25	0.22	1.21	4.05
Storage Volume	0.79	0.17	1.72	7.01
Storage Injection	0.90	0.16	1.36	5.18
Storage Withdrawal	0.77	0.14	1.04	3.89
Total Consumption	26.06	4.82	10.53	17.26
Cooling Degree Day	0.36	0.06	0.30	1.03
Heating Degree	0.22	0.04	0.27	1.04
Extreme Max. Temp.	0.41	0.10	1.85	2.33
Extreme Min. Temp.	0.04	0.01	0.04	0.13
Mean Temperature	2.75	0.55	4.54	16.94

4.4. Result Summary

This section provides a summary of the results obtained using the mentioned models. To better explain the difference between these models, R-squared is calculated for each model. R-squared values do not vary significantly among the SDE models. Table 7 summarizes the calculated R-squared for these models. It can be observed in this table that neural network based models fit the data better than the other models. Among the different neural network models, NARX 21-6-1, is the best model in term of the test performance.

Table 7. Fit results of models.

Model	R-Squared
SDE Models	% 84.3
ARIMA/GARCH	% 84.6
NAR 10-1-1	% 94.5
NAR 10-6-1	% 95.0
NARX 21-1-1	% 93.9
NARX 21-6-1	% 94.6
NARX 14-6-1	% 94.2

The obtained results for one step predictions is available in Table 8. The results obtained using different models are in a good agreement except the one obtained using the NAR model. The real value for gas price is 2.28 \$/MMBtu which is closer to the NAR model prediction. According to the back test of the models, the NARX model has a better fit to the historical data, but the results in Table 8 shows otherwise.

Table 8. Predicted natural gas price for different models.

Model	Predicted Price
GBM	1.93
CEV	1.94
LDR	1.92
ARIMA	1.95
NAR 10-6-1	2.01
NARX 21-6-1	1.92

4.5. Natural Gas Price Forecaster Package

In order to produce the results based on the models that were proposed in this research, a software package has been developed. Graphical User Interfaces (GUIs) were designed to be able to use the models easier and more efficient. This package helps to reproduce the results and also examine the models with different parameters. Figure 72 shows the main menu for the software package. In the main menu, the historical price of natural gas is illustrated. From the radio button section, we can choose the type of models. Pressing “Model Details” button enables us to select the model.

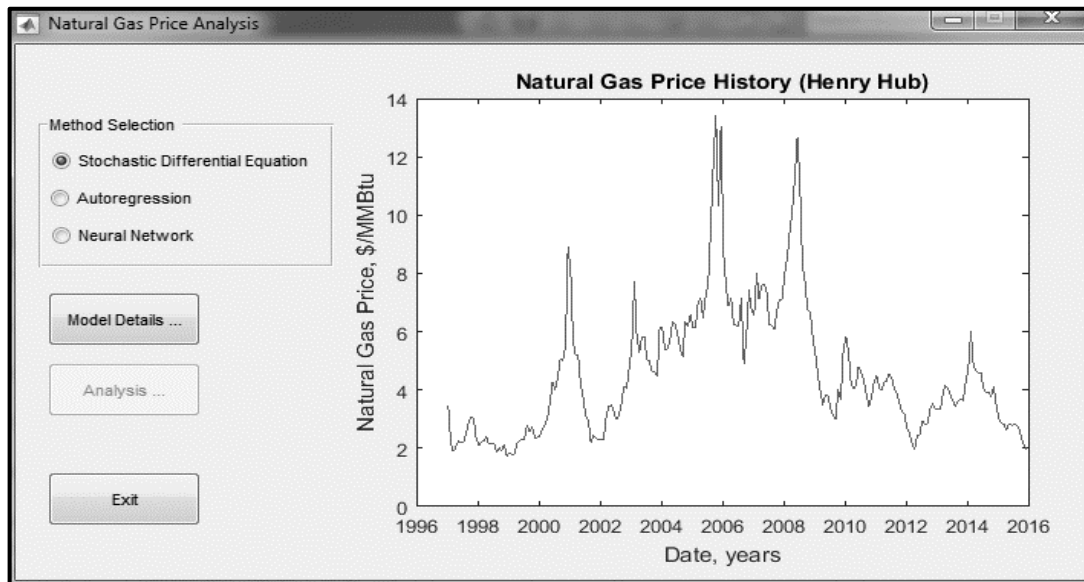


Figure 72. Software package's main menu.

After selecting the method, another window will appear for choosing the specific model. The following figures show the menus which depend on the method selection. Figure 73 has the options to select GBM, CEV, or LDR model. In this window, we can also determine if the statistical parameters are static or dynamic. Figure 74 shows the parameters for autoregression and moving average lags as well as heteroscedasticity parameters. Here we can define these parameters.

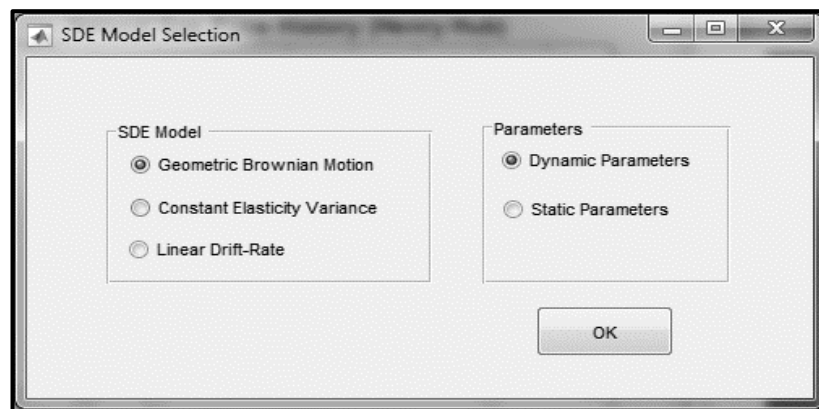


Figure 73. SDE model selection window.

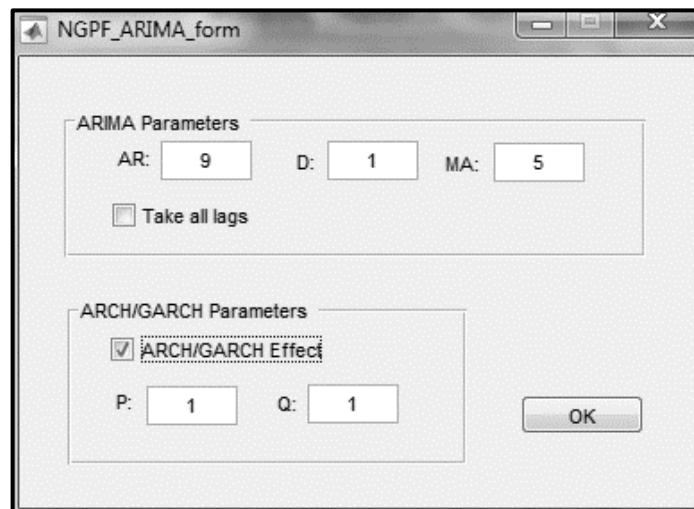


Figure 74. ARIMA / GARCH model selection window

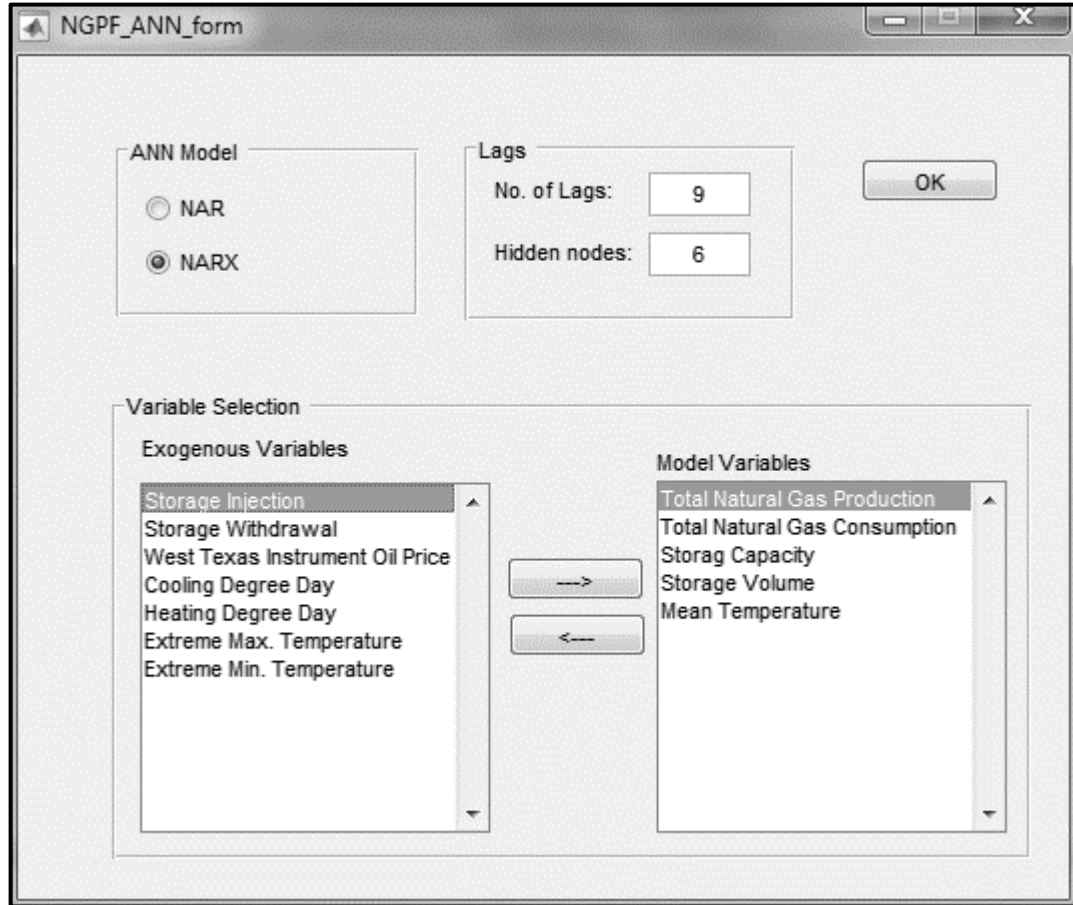


Figure 75. Neural networks model selection window

Figure 75 is the menu for setting the parameters for neural networks models. In this menu, we can determine the exogenous variables, which are introduced into the NARX model.

The next step is to define what type of curves and results are desirable for us. By pressing the “Analysis” button, depending on which model was selected, a new model will appear to set the favorable outputs. Figure 76 to Figure 78 show these windows. The results are the same as presented in the previous sections.

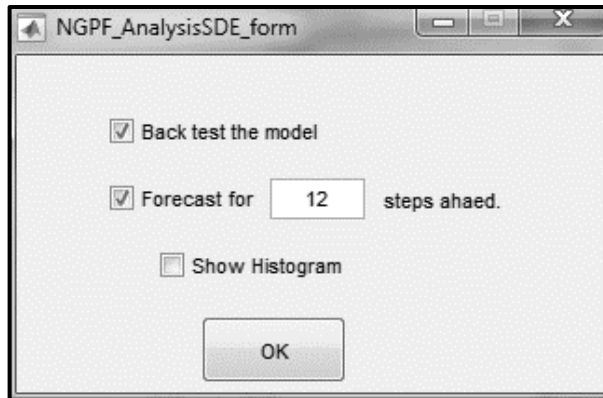


Figure 76. Output setting window for SDE model.

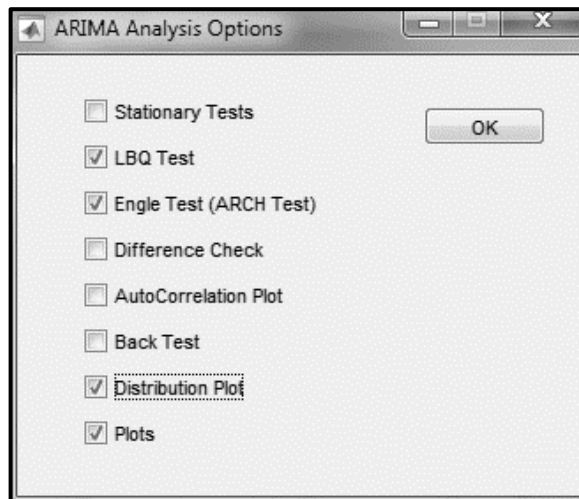


Figure 77. Output setting window for ARIMA model.

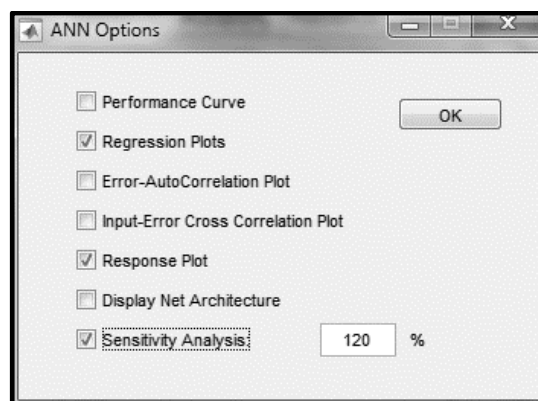


Figure 78. Output setting window for ANN.

Chapter 5: Conclusions

Three different classes of gas price prediction models are examined in this study to determine the optimum model. Introducing exogenous variables is proposed as a viable approach towards accurate price prediction. However, the results in this study indicate that the accuracy in the prediction is not significantly improved by introducing exogenous variables. Moreover, the price spike characterization is not improved despite the increased complexity of the model due to exogenous variables. Therefore, the univariate models may sometimes be selected due to their simplicity as compared to the complex multivariate models.

It is shown in the modeling result that the appropriate number of the lags for the model is 9, particularly in the autoregression models. Due to the seasonality nature of the natural gas price, which is a function of consumption, the cycle is expected to match the calendar cycles, namely, seasonal (3 months) or annual (12 months) periods.

The results indicate that, the NAR neural network provides a better fit to the given data as compared to the other proposed models. The three-layer NAR model with 6 hidden neurons was found to have the best performance in terms of one step ahead price prediction.

The accuracy of the NARX model with 6 neurons is found to be higher than that of the other models. Although, this model provides a reasonable fit to the given data, it fails to capture the price spikes effectively.

The sensitivity analysis shows that CDD/HDD temperatures, extreme minimum temperature, and WTI oil prices have an insignificant effect on the results. On the other

hand, total consumption, total production, and mean weather temperature impact the results significantly.

In summary, three different price modeling classes are examined and compared with each other from the accuracy point of view. Moreover, an accurate yet simple model is proposed in this work that is capable of gas price forecasting in the volatile oil/gas market. Here are some of the model features:

- The same data set is used in all of the simulations which makes the results comparable.
- This work is the first attempt to model natural gas price by multivariable neural autocorrelation technique.
- The sensitivity analysis shows the importance of the exogenous variables in natural gas price simulation.
- It was found that total natural gas consumption and production as well as the temperature changes are the most controlling exogenous variables.
- A user-friendly and easy-to-use software package is developed to simulate gas price behavior using various models with different parameters.

Recommendation for future work in this field should encompass:

- Examine other exogenous variables like weather shock, disaster's effect, and political incidents.
- Using the lags of exogenous variables in the simulation.

References

- Agarwal, S., Saferpour, H. R., & H., D. C. (2014). Adaptive Learning Model for Predicting Negotiation Behaviors through Hybrid K-means Clustering, Linear Vector Quantization and 2-Tuple Fuzzy Linguistic Model. *Procedia Computer Science*, 36, 285-292.
- Agbon, I. S., & Araque, J. C. (2003). Predicting Oil and Gas Spot Prices Using Chaos Time Series Analysis and Fuzzy Neural Network Model. Society of Petroleum Engineers. doi:10.2118/82014-MS
- Al-Fattah, S. M., & Startzman, R. A. (n.d.). Predicting Natural Gas Production Using Artificial Neural Network. Society of Petroleum Engineers. doi:10.2118/68593-MS
- Al-Kaabi, A. U., & Lee, W. J. (1993). Using Artificial Neural Networks To Identify the Well Test Interpretation Model. *Society of Petroleum Engineers*. doi:10.2118/20332-PA
- Anderson, R. B., Kölbel, H., & Rálek, M. (1984). *The Fischer-Tropsch Synthesis*. Academic Press.
- ASHRAE. (1993). *ASHRAE Handbook — Fundamentals (SI). Energy Estimating Methods (chapter 28)*.
- Avalos, R. G. (2012). Pipeline constraints in wholesale natural gas markets: effects on regional pricing and market integration. *Dissertation*. Bozeman: Montana State University, College of Agriculture.
- Ayatollahi, S. (2016). A rigorous approach for determining interfacial tension and minimum miscibility pressure in paraffin-CO₂ systems: Application to gas injection processes. *Journal of the Taiwan Institute of Chemical Engineers*.
- Bachmeier, L. J., & Griffin, J. M. (2006). Testing for market integration crude oil, coal, and natural gas. *The Energy Journal*, 55-71.
- Bahadori, A. (2014). *Natural Gas Processing: Technology and Engineering Design*. Gulf Professional Publishing.
- Barlow, M. T. (2002). A diffusion model for electricity prices. *Mathematical Finance*, 12(4), 287-298.
- Barlow, M., Gusev, Y., & Lai, M. (2004). Calibration of multifactor models in electricity markets. *International Journal of Theoretical and Applied Finance*, 7(2), 101-120.

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307-327.
- Bopp, A. E. (1990). An analytical approach to forecasting natural gas prices. *AGA Forecasting Review: American Gas Association*, 33-52.
- Bower, R. S., & Bower, N. L. (1985). Weather normalization and natural gas regulation. *The Energy Journal*, 101-115.
- Box, G. E., Jenkins, G. M., & Rein, G. C. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- BP. (2014). *Statistical Review of World Energy*.
- Brown, S. P., & Yücel, M. K. (2008). What drives natural gas prices? *The Energy Journal*, 45-60.
- 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.
- Caldwell, R. H., & Heather, D. I. (1996). Why Our Reserves Definitions Don't Work Anymore. Society of Petroleum Engineers. doi:10.2118/30041-PA
- Chandra, V. (2006). *Fundamentals of natural gas: an international perspective*. PennWell Books.
- Chiou-Wei, Zan, S., Chen, C.-F., & Zhu, Z. (2008). Economic growth and energy consumption revisited—evidence from linear and nonlinear Granger causality. *Energy Economics*, 30(6), 3063-3076.
- Chiroma, H., Abdulkareem, Sameem, Abubakar, Adamu, & Mohammed, J. U. (2013). Computational intelligence techniques with application to crude oil price projection: a literature survey from 2001-2012. *Neural Network World*, 23(6), 523-551.
- Considine, T. J. (2000). The impacts of weather variations on energy demand and carbon emissions. *Resource and Energy Economics*, 22(4), 295-314.
- Conti, J. J., Holtberg, P. D., Diefenderfer, J. R., Napolitano, S. A., Schaal, A. M., Turnure, J. T., & Westfall, L. D. (2015). *Annual Energy Outlook*. EIA, DOE. Energy information administration.
- Doris, R. F., & Economides, M. J. (1999). Prediction of short-term natural gas prices using econometric and neural network models. *SPE Hydrocarbon Economics and Evaluation Symposium*. Society of Petroleum Engineers.

- Dougherty, E. L. (1987). Oil Prices: Can We Predict Where They Are Going? . Society of Petroleum Engineers. doi:10.2118/16850-MS
- EIA. (2009). *U.S. Natural Gas Pipeline Network, 2009* . Retrieved from http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/ngpipeline/ngpipelines_map.html
- EIA. (2014). *Natural Gas Annual*. Energy Information Administration (EIA), Office of Oil, Gas, and Coal Supply Statistics. U.S. Department of Energy (DOE).
- EIA. (2016). *US Energy Information Administration*. Retrieved from http://www.eia.gov/dnav/ng/ng_pri_fut_s1_m.htm
- Ekweanua, U. E., Sharma, S. C., Wu, X., Zhu, Z., & Callard, J. G. (2014). Correlating Natural Production, Price, Import and Export in a Neural Network Framework. Society of Petroleum Engineers. doi:10.2118/169867-MS
- Elkhafif, M. A. (1996). An iterative approach for weather-correcting energy consumption data. *Energy economics*, 18(3), 221-230.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.
- Gallager, R. G. (2013). *Stochastic processes: theory for applications*. Cambridge University Press.
- Geman, H., & Ohana, S. (2009). Forward curves, scarcity and price volatility in oil and natural gas markets. *Energy Economics*, 31(4), 576-585.
- Gibson, R., & Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45(3), 959-976.
- Glasserman, P. (2003). *Monte Carlo methods in financial engineering* (Vol. 53). Springer Science & Business Media.
- GPSA. (2013). *Engineering data book*. Gas Processors Suppliers Association.
- Habiballah, W. A., Startzman, R. A., & Barrufet, M. A. (1996). Use of Neural Networks for Prediction of Vapor/Liquid Equilibrium K-Values for Light-Hydrocarbon Mixtures. *Society of Petroleum Engineers*. doi:10.2118/28597-PA
- Hajirezaie, S., & et al. (2015). A smooth model for the estimation of gas/vapor viscosity of hydrocarbon fluids. *Journal of Natural Gas Science and Engineering*, 1452-1459.
- Haykin, S. (1999). *Neural Networks, A Comprehensive Foundation*. Prentice Hall.

- Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory*. Psychology Press.
- Hecht-Nielsen, R. (1989). Theory of the backpropagation neural network. *International Joint Conference on Neural Network 1989* (pp. 593-605). IEEE.
- Hoskins, J. C., & Himmelblau, D. M. (1988). Artificial neural network models of knowledge representation in chemical engineering. *Computers & Chemical Engineering*, 12(9), 881-890.
- Hosseiniipour, S., Babalouyan, S., Tehrani, R., & Pourkazemi, V. (2009). The Performance of Filter Rules on the Tehran Stock Exchange. *Fifth Global Academy of Business and Economic Research Conference (GABER)*, (pp. 549-556). Kuala Lumpur.
- Hosseiniipour, S., Babalouyan, S., Tehrani, R., Hosseiniipour, V., & Pourkazemi, V. (2009). Testing Weak Form Efficiency for Tehran Stock Market. *Fifth Global Academy of Business and Economic Research Conference (GABER)*, (pp. 644-651). Kuala Lumpur.
- Hotelling, H. (1931). The economics of exhaustible resources. *The journal of political economy*, 137-175.
- Hsieh, Y. (1990). Natural Gas Price Projection: A Methodology. *American Gas Association Forecasting Review*, 2.
- Hu, Y., & Trafalis, T. B. (2011). New kernel methods for asset pricing: application to natural gas price prediction. *International Journal of Financial Markets and Derivatives*, 2(1-2), 106-120.
- Huston, J. L. (2015). *The Declaration of Dependence: Dividends in the Twenty-First Century*. Archway Publishing.
- IEA. (1995). *The IEA Natural Gas Security Study*. International Energy Agency. Paris: OECD/IEA.
- Inikori, S. O., Kunju, M. K., & Iledare, O. O. (2001). The Responsiveness of Global E&P Industry to Changes in Petroleum Prices: Evidence From 1960 - 2000. Society of Petroleum Engineers. doi:10.2118/68587-MS
- Jarque, C. M., & Bera, A. K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 55(2), 163–172.
- Kamali, A., & Pournik, M. (2015). Rough Surface Closure – A Closer Look at Fracture Closure and Conductivity Decline. *ISRM Regional Symposium - EUROCK 2015*. Salzburg, Austria.: International Society for Rock Mechanics.

- Kamali, A., & Pournik, M. (2016). Fracture closure and conductivity decline modeling—Application in unpropped and acid etched fractures. *Journal of Unconventional Oil and Gas Resources*, 44-45.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3), 159–178. doi:10.1016/0304-4076(92)90104-Y
- Linn, S. C., & Zhu, Z. (2004). Natural gas prices and the gas storage report: Public news and volatility in energy futures markets. *The Journal of Futures Markets*, 24(3), 283.
- Lucia, J. J., & Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange." *Review of derivatives research*. 5(1), 5-50.
- MacAvoy, P. W., & Pindyck, R. S. (1974). *The economics of the natural gas shortage (1960-1980)*. MIT Energy Lab.
- Mishra, P. (2012). Forecasting Natural Gas Price-Time Series and Nonparametric Approach. *Lecture Notes in Engineering and Computer Science*.
- Mokhatab, S., & Poe, W. A. (2012). *Handbook of natural gas transmission and processing*. Gulf Professional Publishing.
- Mokhatab, S., Mak, J. Y., Valappil, J. V., & Wood, D. A. (2013). *Handbook of liquefied natural gas*. Gulf Professional Publishing.
- Morgan, M. D., Mikalson, D., & Herchen, L. L. (2012). A Model of Canadian Oil and Gas Price Fluctuations: 2012 Update. Society of Petroleum Engineers. doi:10.2118/162629-MS
- Mu, X. (2007). Weather, storage, and natural gas price dynamics: Fundamentals and volatility. *Energy Economics*, 29(1), 46-63.
- NOAA. (2016). *Climate Data Online*. Retrieved from <http://www.ncdc.noaa.gov/cdo-web/datasets>
- Nogales, F. J., Contreras, J., Conejo, A. J., & Espínola, R. (2002). Forecasting next-day electricity prices by time series models. *Power Systems, IEEE Transactions on*, 17(2), 342-348.
- Ogwo, O. U., Mbeledogu, I. U., & Nwokedi, M. (2007). Equitable Gas Pricing Model. Society of Petroleum Engineers. doi:10.2118/111897-MS
- Omidvar, A., Garakani, M., & Safarpour, H. R. (2013). Context based user ranking in forums for expert finding using WordNet dictionary and social network analysis. *Information Technology and Management*, 1-13.

- Panella, M., Barcellona, F., & D'eclesia, R. L. (2012). Forecasting energy commodity prices using neural networks. *Advances in Decision Sciences*.
- Perron, P. (1988). Trends and random walks in macroeconomic time series: Further evidence from a new approach. *Journal of economic dynamics and control*, 12(2), 297-332.
- Pilipovic, D. (2007). *Energy Risk: Valuing and Managing Energy Derivatives: Valuing and Managing Energy Derivatives*. (2nd ed.). McGraw Hill Professional.
- Pindyck, R. S. (1978). The optimal exploration and production of nonrenewable resources. *The Journal of Political Economy*, 841-861.
- Powers, L. W., & Stevenson, W. M. (1987). Perceptions the Key to Oil Prices. Society of Petroleum Engineers. doi:10.2118/16838-MS
- Prabhu, N. U. (2007). *Stochastic processes: basic theory and its applications*. World scientific.
- Qian, N., & Sejnowski, T. J. (1988). Predicting the secondary structure of globular proteins using neural network models. *Journal of molecular biology*, 202(4), 865-884.
- Reiter, D. F., & Economides, M. J. (1999). Prediction of short-term natural gas prices using econometric and neural network models. *SPE Hydrocarbon Economics and Evaluation Symposium*. Society of Petroleum Engineers. doi:10.2118/52960-MS
- Richardson, J., & Nordhaus, R. (1995). The National Energy Act of 1978. *Natural Resources & Environment*, 10(1), 62-88.
- Roberts, M. L. (1989). OPEC Production and Crude Prices: Review and Outlook. Society of Petroleum Engineers. doi:10.2118/18914-MS
- Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning*. *Cognitive science*, 9(1), 75-112.
- Safarpour, H. R., Eneyo, E. S., & Shahrak, M. Z. (2013). Presenting an Integrated Model for Studying the Key Factors of Customer Trust in Developing Countries' E-Commerce Environment. *International Journal of Research in the Academic Disciplines in Higher Education*.
- Safarpour, H. R., Fathi, A., Hubbard, K. M., & Eneyo, E. S. (2013). Single Supply of Electrical Energy in the Supply Distribution Model: Using Gams to Model the Effects of Network Parameters on the Level of Consumer Demand. *International Journal of Research in the Academic Disciplines in Higher Education*.

- Safarpour, H. R., Shahrak, M. Z., Eneyo, E. S., & Hubbard, K. M. (2014). Social Networks Interactions in Developing Countries E-Commerce. *International Journal of Engineering Research and Technology*.
- Schulz, H. (1999). Short history and present trends of Fischer–Tropsch synthesis. *Applied Catalysis A: General*, 186(1), 3-12.
- Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3), 923-973.
- Shahrak, M. Z., Mohammadi, S., & Safarpour, H. (2014). Middle East User Navigation in Online Social Networks and Interactions in E-Commerce, an Analogy. *Advances in Computer Science: an International Journal*, 32-36.
- Skov, A. M. (1995). An Analysis of Forecasts of Energy Supply, Demand, and Oil Prices. Society of Petroleum Engineers. doi:10.2118/30058-MS
- Song, Z.-W., Linn, S. C., & Zhu, Z. (2013). The Response of US Natural Gas Futures and Spot Prices to Storage Change Surprises and the Effect of Escalating Physical Gas Production. *Journal of International Money and Finance*.
- Speight, J. G. (2007). *Natural gas: a basic handbook*. Elsevier.
- Taylor, S. J. (2007). *Modelling financial time series*.
- Texas. (1976). *Natural Gas: A Case for Decontrol*. Austin, TX: Governor's Energy Advisory Council.
- Tohidi-Hosseini, S.-M. (2016). Toward Prediction of Petroleum Reservoir Fluid Properties: A Rigorous Model for Estimation of Solution Gas-Oil Ratio. *Journal of Natural Gas Science and Engineering*.
- Trippi, R. R., & Turban, E. (1992). *Neural networks in finance and investing: Using artificial intelligence to improve real world performance*. McGraw-Hill, Inc.
- U.S. EIA | IEO. (2013). *With Projections to 2040*.
- Villar, J. A., & Joutz, F. L. (2006). *The relationship between crude oil and natural gas prices*. Energy Information Administration, Office of Oil and Gas.
- Wang, X., & Economides, M. (2013). *Advanced natural gas engineering*. Elsevier.
- Wood, D., Mokhatab, S., & Economides, M. J. (2008). Technology options for securing markets for remote gas. *Proc. 87th Annual Convention*. GPA.
- Xu, z. (2004). Stochastic Models for Gas Price. *Thesis*. Alberta: The University of Calgary.

- Yi, X., & Wang, Z. (2013). The Study on Oil Prices' Effect on International Gas Prices Based on Using Wavelet Based Boltzmann Cooperative Neural Network. *Proceedings of the 2012 International Conference of Modern Computer Science and Applications*. Springer Berlin Heidelberg.
- Yıldırım, M. H., Bayrak, Ö. T., & Weber, G.-W. (2014). Survey and Evaluation on Modelling of Next-Day Electricity Prices. *Modeling, Dynamics, Optimization and Bioeconomics I*. (pp. 723-737). Springer International Publishing.
- Zhu, Z., & Maxwell, D. (2011). Natural Gas Prices, LNG Transport Costs, and the Dynamics of LNG Imports. *Energy Economics*, 33(2), 217-226.

Appendix A: Nomenclature

ANN	:	Artificial Neural Network
ARCH	:	Autoregressive Conditional Heteroscedasticity
ARIMA	:	Autoregressive Integrated Moving Average
a_t	:	White noise at time t
B^r	:	Backward operator at lag r
CCD	:	Cooling degree days
CEV	:	Constant Elasticity of Variance
d	:	Integrating parameter in ARIMA model
$d_k(n)$:	Desired value of neuron k at iteration n
E	:	Boltzmann energy function
$E(X)$:	Expected value of X
$e_k(n)$:	Calculated error of neuron k at iteration n
GARCH	:	General Autoregressive Conditional Heteroscedasticity
GBM	:	Geometric Brownian motion
HDD	:	Heating degree days
LDR	:	Linear Draft-Rate
MSE	:	Mean Squared Error
m	:	Moving average parameter in ARIMA model
\mathbb{N}	:	Set of natural numbers
NAR	:	Nonlinear Autoregressive
NARX	:	Nonlinear Autoregressive with exogenous variables

p	:	GARCH model parameter
q	:	ARCH model parameter
\mathbb{R}	:	Set of real numbers (state space)
r	:	Autoregression parameter in ARIMA model
S_t	:	Stochastic process
SAE_c	:	Calculated mean error for model in case c
SDE	:	Stochastic Differential Equation
\mathbb{T}	:	Set of numbers demonstrate time
t	:	Time index
WTI	:	West Texas Instrument
W_t	:	Wiener process at time t
w_{ki}	:	Synaptic weight for signal input i in neuron k
$\Delta w_{ki}(n)$:	Adjusted to synaptic weight at iteration n
X	:	Random variable
X_t	:	Discrete stochastic process (random) variable
$X(t)$:	Continuous stochastic process (random) variable
x	:	Real number
x_i	:	Input signal to neuron
y_k	:	Response of neuron k
$y_k(n)$:	Calculated value of neuron k at iteration n
\hat{y}_b	:	Calculated response for model in base case
\hat{y}_c	:	Calculated response for model in case c

α	:	Leverage parameter of price level and price volatility
α_q	:	Regression coefficient of ARCH at lag q
β_p	:	Regression coefficient of GARCH at lag p
$\gamma(t_1, t_2)$:	Auto-covariance of time series variable at time t_1 and t_2
ε_t	:	Random error at time t
η	:	Learning rate
σ	:	Standard deviation / Volatility
θ_m	:	Moving Average (MA) coefficient at lag m
μ	:	Mean value
σ^2	:	Variance
$\sigma_{t t-1}^2$:	Conditional variance at time t
v_k	:	Induced local field of neuron k
ϕ_r	:	Auto-Regression (AR) coefficient at lag r
$\varphi(\cdot)$:	Activation function of a neuron
Ω	:	Probability space
ω	:	Probability number

Appendix B: Software Package

A software package including MATLAB® m files, Microsoft Excel® files, and related files and codes are associated with this research. The files are compressed in rar format and are available [online](#). The author's email saied@ou.edu is also available to respond to the questions and requests.