Predicting the Price of Natural Gas

Angela Zhang, Martin Zagari, and Omar Elfeky

University of San Diego

Shiley-Marcos School of Engineering

# Abstract

Natural gas is a crucial fuel in modern industrial society, contributing 23% of global primary energy demand, alongside oil, coal and nuclear fuels (REF1). For example, according to the US Energy Association, natural gas fuels 40% of US electricity production (EIA) and that figure continues to increase (EIA). Nearly half of US households rely on natural gas for heating (Vox). Natural gas is also being looked as a “bridging fuel” for climate change. Natural gas releases 48% less CO2 than coal and 34% less than oil having equivalent heat content (REF2). Although it burns cleaner it not as transportable and storable as oil or coal. It is so expensive to store and transport that energy producers routinely vent or “flare … natural gas during well or processing disruptions and when they lack access to pipelines” (Hampton, Oct. 11, 2021). Thus, natural gas price volatility is higher than oil (REF3) since prices are determined by local supply and demand and flexible storage is still problematic (REF4). For this reason, it is important to study the types of cycles and patterns that characterize the price stability of natural gas if it is to become more important in the transition to a net zero economy. Important research questions include: 1) the importance of seasonality vs market factors in gas price, 2) relationships between gas prices and other energy commodities such as oil and coal, 3) volatility trends in gas market prices.

**Abstract:**  
Natural gas is a major global energy commodity. Gas prices around the world face substantial volatility, inducing major downside market risks. Forecasting accuracy is thus a major concern for the consumers. Traditional econometrics models do not perform well due to inherent nonlinear and nonstationary gas price data. We thus propose an Autoregressive Neural Network (ARNN) model for forecasting daily spot gas prices. The model is benchmarked against the traditional Autoregressive Integrated Moving Average (ARIMA) model. Using a cross validation study, the ARNN model showed an improvement of around 33% over ARIMA in terms of mean squared error. This improvement is significant when price forecasts are used in gas purchase decisions.

------------

REF1: International Energy Agengy, Global Gas Security Review 2021, Fuel Report October 2021. Ed. Justin French-Brooks. www.iea.org

REF2: US Energy Information Administration FAQs. https://www.eia.gov/tools/faqs/faq.php?id=73&t=11

REF4, Gas Trading Manual. Chapter 2: Fundamentals of the Gas Market. David Long, Geoff Moore, Gay Wenban-Smith. 2003. Woodhead Publishing. ISBN: 978-1-85573-446

*Keywords:* commodities, electricity, industry. economy

**Table of Contents**

[Abstract 2](#_Toc75716696)

[List of Figures 5](#_Toc75716697)

[List of Tables 5](#_Toc75716698)

[Problem Statement 6](#_Toc75716699)

[Exploratory Data Analysis 6](#_Toc75716700)

[Data Preprocessing and Splitting 7](#_Toc75716701)

[Model Strategies 11](#_Toc75716710)

[Validation and Testing 14](#_Toc75716727)

[Conclusion 16](#_Toc75716729)

[References 17](#_Toc75716730)

[Appendix A 18](#_Toc75716731)

[Appendix B 20](#_Toc75716732)

List of Figures

[Figure 1. <FigureTitle> 8](#_Toc75718924)

List of Tables

[Table 1<TableTitle> 15](#_Toc75718928)

# Problem Statement

This is only the tip of the iceberg. As the strong trend of replacing coal for natural gas continues there are legs to go much further since coal still generates nearly 20% of the electricity produced, the fourth highest source (EIA). Renewable sources of electricity are notoriously volatile producing as low as 0 percent when the sun isn’t shining, or the wind isn’t blowing. Natural gas at a minimum is an energy bridge to renewables producing half as many emissions as coal. Without significant improvements in battery technology or widespread adoption of nuclear; natural gas will be the core baseload electricity source. Modern life and habitation of cold climates relies on the electricity and heating of natural gas, and it is imperative for the oil and gas industry to predict the prices to invest in production to meet demand. Equally critical for high energy users and manufacturers to be able to predict the cost of one of their major inputs. As the recent Experience in Europe has demonstrated, more renewable electricity reduces the use of fossil fuels but does not eliminate it and without careful planning can increase fossil fuel usage. Germany, one of the highest green energy spenders, installed ~54 gigawatts of solar panels enough to technically provide 25% of its electricity needs but only generates 10% (Hancock, E. 2021).

**Natural Gas Price History**

Natural gas prices have exhibited a much different profile than that for other energy sources, particularly oil. At the highest level, natural gas prices exhibited a general upward trend from approximately 1997 through the financial crisis of 2007-2008, and then a general downward trend from 2009 until the present day. Of course, there is substantial volatility in any given week or day. This is different from the generally upward trend in oil prices, meaning that the overall demand for energy and economic growth is likely not the major driver of gas price as it is for oil price. Although the overall natural gas price trend for nearly 25 years has been essentially flat, substantial price movement suggests complex market forces and dependencies that bear exploration.

One important difference between oil and gas is that natural gas is a by-product of oil production, meaning that the demand for and production of oil may very well have a knock-on effect on gas prices. As a result, we will explore the relationship between oil and gas prices, with gas prices potentially being a lagging variable for oil demand and/or production. In particular, the shale oil boom in Texas may have had an outsized impact on the price of natural gas, supressing it for much with last decade with a recent resurgence in prices perhaps being due to a collapse in the shale oil boom.

Another important difference between oil and gas is ease of transport and storage. Despite considerable infrastructure for gas transport, there are still many fewer options for natural gas transport and storage than for oil. (But fortunately, natural gas is much more storable than electricity!). For this reason, exogenous shocks such as weather may impact gas markets differently than they do for oil markets, and natural gas prices are, in fact, more volatile than for oil (but less volatile than electricity). Evidence of this is seen in headlines speculate on gas price spikes based on mild vs severe cold weather, something that doesn’t really occur in oil markets. For this reason, we will also explore the importance of seasonality in natural gas prices that may be obscured by trend and shocks.

Finally we still see “shocks” in gas price corresponding with important world and economic events. One gas price spike appears to occur around September 11, 2001. A series of price spikes occurred before and around the financial crisis, perhaps corresponding first to economic growth, followed by economic uncertainty and the collapse of oil prices. Our models will therefore have to account for a high level of exogenous noise in energy markets, as most economic models do.

To summarize, natural gas prices appear to have followed an extremely long cycle upward and downward trend, with frequent macroeconomic and political shocks along this journey, perhaps with underlying seasonality in price. Our analysis will take these domain factors into account.

# Data Exploration, Pre-processing, and Splitting

## *Zero and Near Zero Variance*

## *Missing Values*

## *Data Imputation*

## *Data Preparation*

Following the importation of necessary packages and the raw datasets, the data was prepared for further evaluation. The initial phase of pre-processing began with inspection of the data and calculation of statistical measures. Other properties such as head and dimension were produced to provide additional information. Statistical measures such as median and mean are useful in outlier detection. However, outliers were not removed from the dataset due to having a potential statistical significance on forecasting results and seasonal behavior. Missing values were collected and handled appropriately. The dataset contained only one missing value which was removed completely. Each data frame was converted to a time series object using the date column. Finally, the resulting time series objects were plotted to visualize both the daily and monthly behavior of natural gas prices.

### *Sampling*

### *Distribution*

### *Data Partitioning*

# Model Strategies

# Validation and Testing

# Conclusion

# References

US Energy Information Administration. Electricity in the United States. <https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php#:~:text=Natural%20gas%20was%20the%20largest,in%202020%E2%80%94about%2019%25>.

Roberts, D. (2018, July 2).Vox. Most American homes are still heated with fossil fuels. It’s time to electrify. <https://www.vox.com/energy-and-environment/2018/6/20/17474124/electrification-natural-gas-furnace-heat-pump>

Hancock, E. (2021, February 3). PV-Tech. Germany’s Federal Network Agency registers 54GW of solar PV.

<https://www.pv-tech.org/germanys-federal-network-agency-registers-54gw-of-solar-pv/>

Hampton, L. (2021. October 11). Reuters . U.S. shale producer APA ends flaring, captures more gas as prices soar. <https://www.reuters.com/business/energy/us-shale-producer-apa-ends-flaring-captures-more-gas-prices-soar-2021-10-11/>

# Appendix A – Reproduceable Markdown Code