Capstone Project: Forecasting Carbon Emissions

A Sample Report

Executive Summary

This project proposes the Seasonal ARIMA (SARIMA) model for the prediction and forecasting of CO_2 emissions from natural gas (NG) based electricity generation in the United States. It is suggested that the model should only consider data beginning in the mid-1990s, which was a period of significant change in the US energy system that precipitated the displacement of coal with NG in electricity generation. The suggested model boasts high conformance with observed data and low error. However, it is subject to a number of limitations, including a lack of consideration for regional policy changes, technological shifts, and other pertinent factors such as the development of renewables. It is recommended that stakeholders consider these variables in building improved long-term forecasting models, as well as include the full range of economic and environmental implications of various energy sources in developing future energy policy.

Problem Summary

Climate change is one of the most pressing challenges facing the planet today. Global warming is primarily driven by CO_2 emissions from various industrial activities and power generation, which remain important for economic development and our day-to-day lives. As the US continues to develop strategies to minimize carbon emissions while maintaining a healthy economy, it is becoming increasingly important to identify the sources of CO_2 and forecast future emissions so as to reduce climate impact and move towards a cleaner energy sector. The key objective of this project is to build a **forecasting model for the CO_2 emissions from natural gas (NG)** in the US in order to provide policy recommendations to reduce emissions. The time series analysis presented here provides insight into the relative contributions of natural gas emissions to the total electricity sector emissions in the US, demonstrates trends in emissions over time, and projects future NG-based emissions. The discussed analysis and forecast models will help to understand the projected changes in natural gas-based emissions and serve as a basis for recommendations for future environmental policy.

Solution design

A number of time-series models and methods were explored as part of the solution design, including seasonal naïve models, models using the full and partial datasets, and the more complex PROPHET model. The final proposed solution is the **Seasonal ARIMA (SARIMA) model** using the 1995 data, which has been optimized with AutoARIMA parameter tuning. The 1995- dataset was chosen because the US energy system experienced a shift around this period (further discussed in the next section), resulting in a decline in coal use and a move towards NG.

Figure 1 shows the best model, which yields very low RMSE (2.73 and 5.29 for training and test data, respectively). This model was effective in capturing the trend and seasonality of the data and was an overall very close match to the original dataset. SARIMA with coal emissions as an exogenous variable was also tested (see Appendix 1), which yielded a slightly higher RMSE for the test data due to additional errors from coal emissions forecasting. Overall, this model was very similar to SARIMA without the exogenous variable and did not provide much additional insight given the available dataset.

Furthermore, as discussed in the following section, it is more likely that the increased availability of NG precipitated the phase-out of coal, implying that NG uptake may be an effective predictor of coal emissions, but not the other way round.

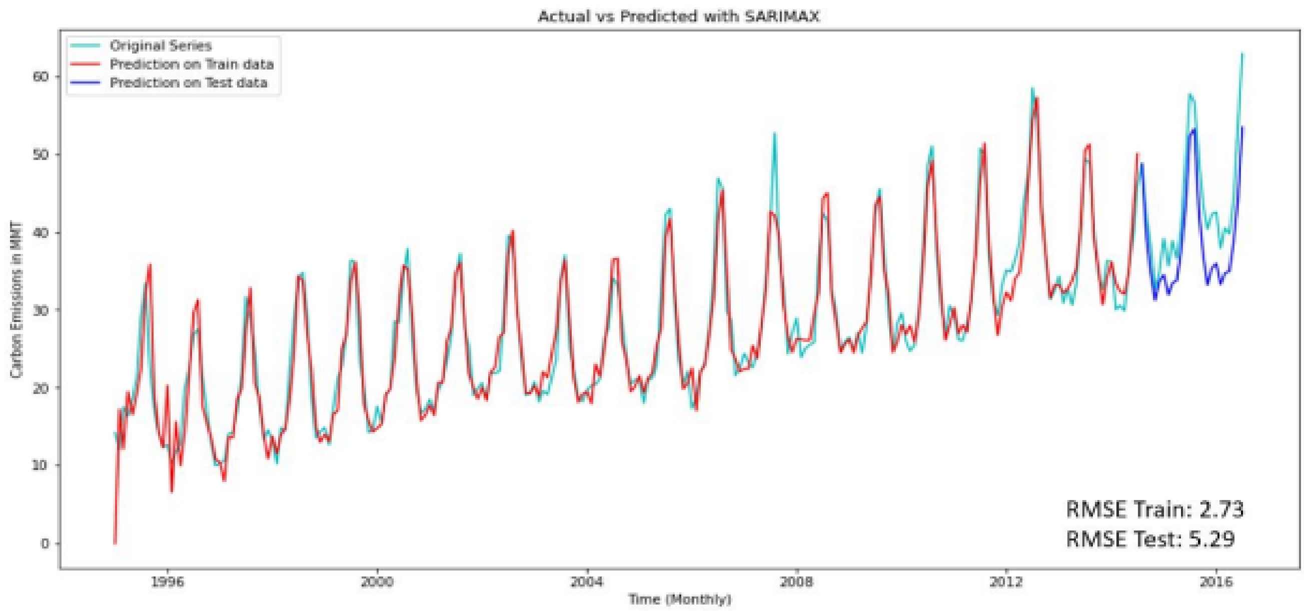


Figure 1: Best overall model SARIMA (1,1,1,12) using 1995- data.

The PROPHET model, illustrated in Appendix 2, also produced a very close match to the observed dataset with a very low RMSE. While it appears to be a robust predictive model, it works best with larger datasets and daily data. Because monthly data is not well-supported by this module, it poses challenges for tasks such as cross-validation. Therefore, I would recommend conventional modeling methods given the limitations of this dataset.

To draw accurate insights from the NG predictive model, it is important to examine the broader emission patterns in the electricity sector along with contextual information from secondary sources. It is also important to distinguish between emissions data and the amount of electricity generated from the given source. Different sources have varying carbon intensities (emissions per kilowatt-hours of electricity). In general, coal tends to be one of the more carbon-intensive fuel sources as it is associated with higher emissions than oil or NG for the same amount of electricity generated. **Figure 2** shows that the total electricity sector emissions trended upwards throughout the 1970s and into the 2000s and fell thereafter. Coal-based emissions historically accounted for the largest share of the total emissions, even during the decline beginning in the 2000s. **Figure 3** further illustrates that coal maintained the highest proportion of electricity generation for the bulk of the examined time period. However, beginning in the 1990s, the proportion of NG has progressively increased, eventually overtaking coal as the main electricity source.

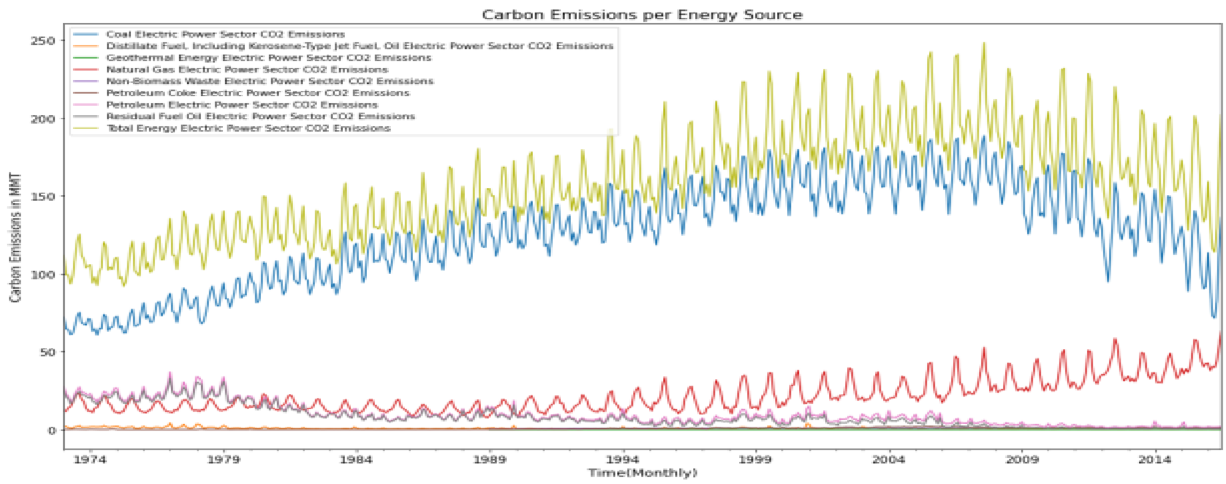


Figure 2: Carbon emissions per energy source in the US electricity sector

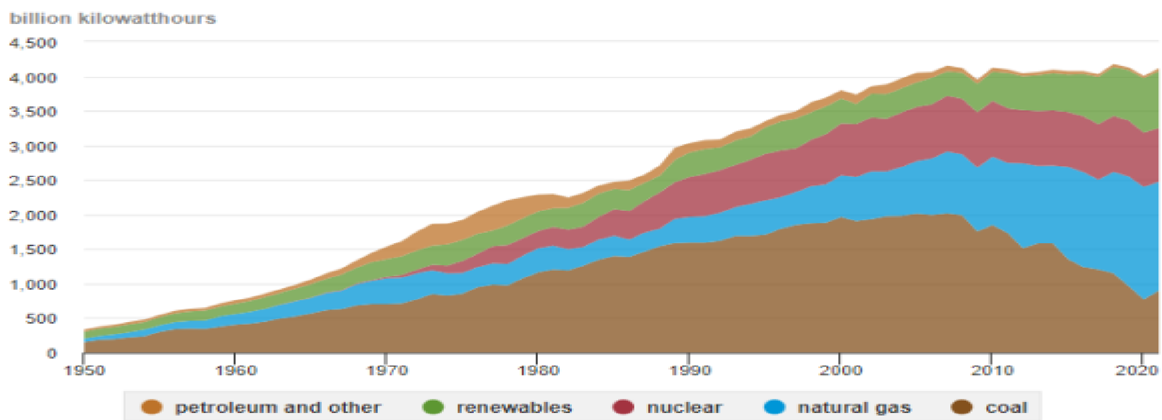


Figure 3: US electricity generation by major energy source, 1950-2021. Source: US Energy Information Administration (EIA)

The 1990s shift in the US electricity sector can be partly explained by the initial rise in imported and domestically produced NG, followed by a sharp rise in domestic NG production throughout the 2000s (shown in **Figure 4**). The sharp rise in production was caused by the application of hydraulic fracturing ("fracking") and horizontal drilling technology to extract the previously discovered shale gas deposits (USGS, 2019). Prior to this technological innovation, gas was only extracted from conventional deposits, while shale deposits remained inaccessible. The expedient and cost-efficient exploitation of shale gas effectively outcompeted the **less economical coal power generation** (Katusa, 2012). In other words, the rise of fracking NG in the United States played a key role in the displacement of coal and its associated emissions that were observed in our data.

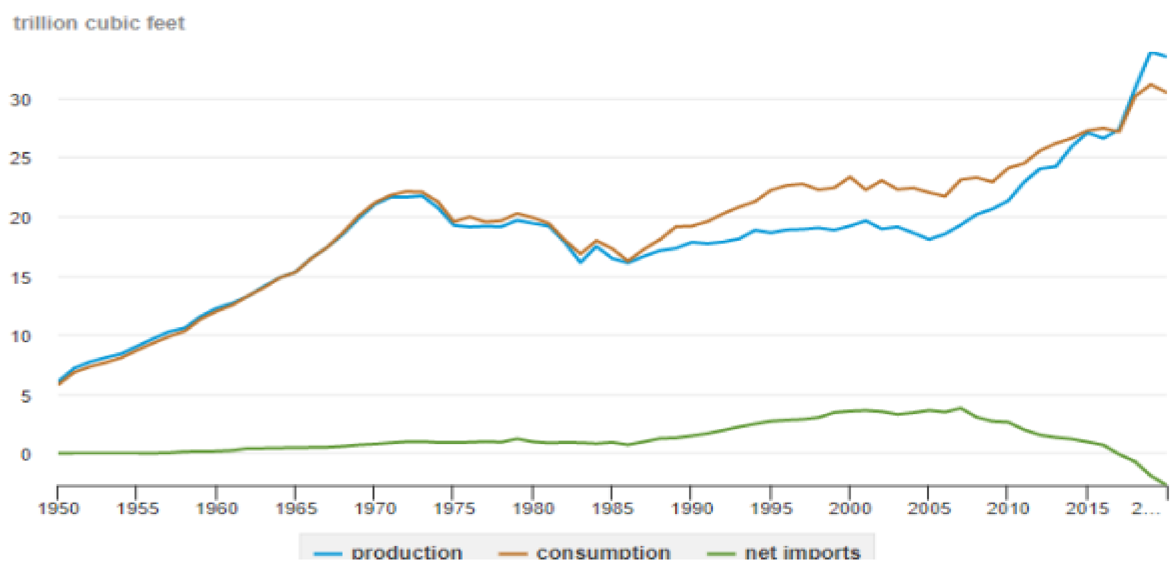


Figure 4: US Natural gas consumption, production, and net imports, 1950-2020. Source: US EIA

Limitations and Recommendations for Further Analysis

Assuming the continued exploitation of shale gas, our model, which predicts the continued increase of NG-based carbon emissions, is likely to be accurate and useful in the short-medium term (i.e., 2-10 years). However, changes in regional policy can quickly nullify the key assumption. For instance, fracking has already been banned in several states due to environmental concerns, with further restrictions in other states and at the federal level (Bermel and Kahn, 2021). Furthermore, like all nonrenewable resources, shale gas deposits are limited, with experts divided on how much gas is actually available (and extractable without significant environmental damage).

For the purpose of NG emissions forecasting, stakeholders need to consider the social, environmental, and economic trade-offs associated with the continued expansion of fracking and/or imports of NG, as they will likely influence the future role of NG in the US energy system. Another aspect worth noting is **the emergence of biomass in the late 1980s**, and the substantive growth of wind and solar as major electricity sources beginning in the 2000s (see Appendix 3).

While these sources were not reflected in our carbon emissions dataset, they could be crucial to consider in long-term emissions forecasting as they may displace carbon-emitting sources, especially if backed by favorable policies and investment.
Overall, a flat univariate time series model is not able to reflect the complex contextual factors leading to the rise of NG in the US nor foresee major shifts (such as the one seen in the 1990s). Radical shifts in energy systems are not only possible, but are made increasingly likely by rapid innovation, such as the large-scale implementation of smart grids, or the development of new energy sources (e.g., better nuclear power).

Therefore, for more accurate long-term projections, it would be necessary to include **variables pertaining to other energy sources** and emerging technologies, as well as keep in mind the physical limitations of non-renewable sources.

Recommendations for Policy

The above analysis clearly demonstrates a decline in overall emissions from the electricity sector despite the rise in NG-based power generation. This is inevitably due to the phase-out of the more carbon-intensive coal-based power plants and their displacement by the more efficient NG power plants. Thus, in the short term, I would recommend a continued phase-out of coal coupled with the continued development of NG. However, keeping in mind the carbon emissions and other environmental implications of NG (especially from fracking), I would recommend continued investment in cleaner energy sources, such as wind, solar, and geothermal, as well as new technologies to reduce energy intensity/demand and increase efficiency.

It is also worth noting that emissions data from electricity generation do not fully encompass the carbon emissions from the given energy source. Fracking, in particular, has been associated with the leakage of methane, which is a highly potent greenhouse gas. Conventional extraction of gas, as well as oil and coal, is also associated with emissions from mining operations and leakages. Therefore, for better-informed policymaking, I would recommend better record-keeping of emissions associated with extraction, processing, transportation, and other activities pertaining to the energy sources in question.

Bibliography

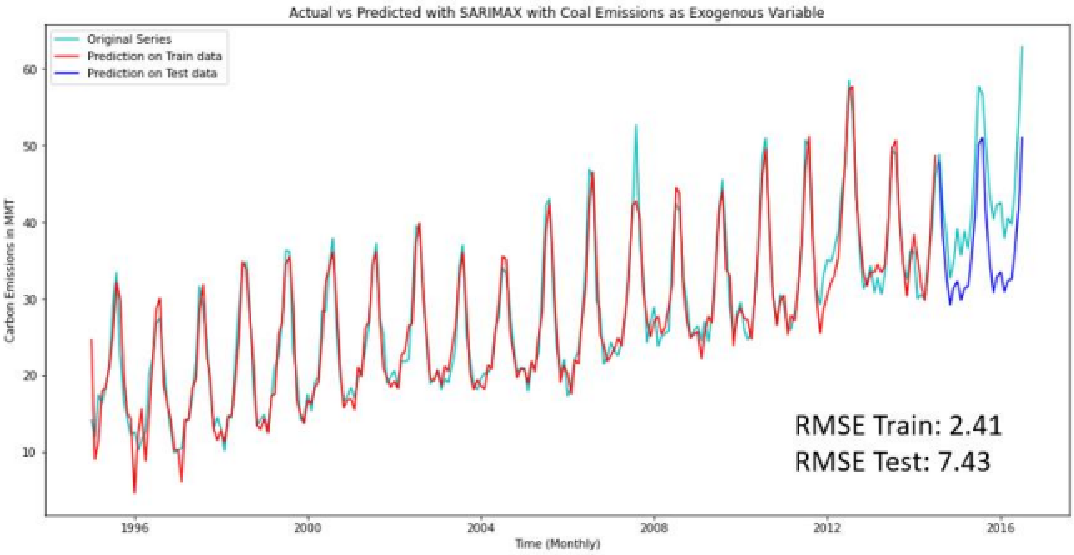
- Bermel, C., and Kahn, D. 2021. Newsom to announce a ban on new fracking projects. Newsom to announce ban on new fracking projects - POLITICO
- Katusa, M. 2012. Shale gas takes on coal to power America's electrical plants. Shale Gas Takes On Coal To Power America's Electrical Plants (forbes.com)
- National Petroleum Council. 2011. Prudent Development: Realizing the potential of

North America's abundant natural gas and oil resources.

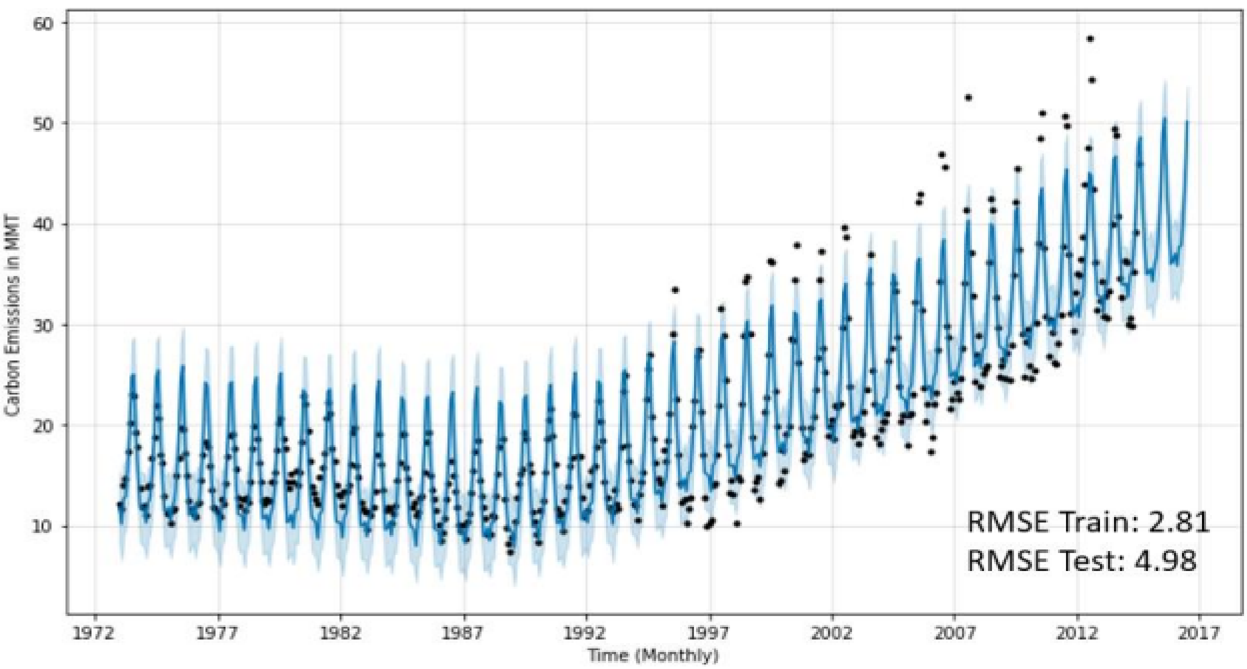
- The United States Energy Information Administration (EIA). 2021. Electricity Explained:
  Electricity in the United States. Electricity in the U.S. - U.S. Energy Information Administration (EIA)
- United States Geological Survey (USGS). 2019. Hydraulic Fracturing. Hydraulic Fracturing | U.S. Geological Survey (usgs.gov)

Appendix

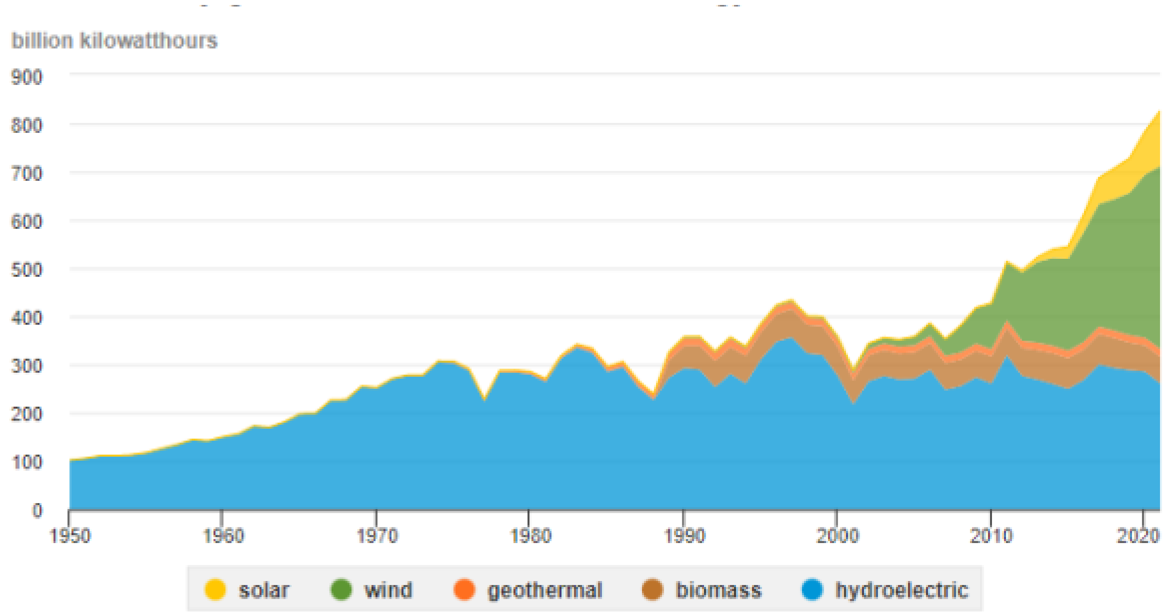
Appendix 1: SARIMA with Coal Emissions as an Exogenous Variable



Appendix 2: Prediction using the Prophet



Appendix 3: US electricity generation from renewable energy sources, 1950-2021



Source: US Energy Information Administration (EIA)