Analysis and Forecast of Fossil Fuels and Renewable Energy in the United States

University of California, Santa Barbara

PSTAT 174

Andrew Nguyen

ID: 8882383

Executive Summary

The goal of this study is to determine whether univariate time series models can answer the question if there is a negative correlation between fossil fuels and renewable energies and whether it is possible to determine when electricity production from fossil fuels will reach zero using electricity production data of the United States. All data sets resembled a SARIMA (Seasonal Auto-Regressive Integrated Moving Average) model to forecast values against 2022's recorded values, and the 3 best models from each data set were analyzed on their forecast performance. We concluded that there were considerable forecast errors due to the fact that electricity production rates cannot be predicted based on value alone, but rather predicted given the context of other variables such as climate change or price changes in resources for producing such electricity. Further studies should be conducted into more complex time series that would accommodate multivariate scenarios.

Introduction

With ecological issues rising with each passing year, we must look to different methods of life to conserve the resources of the Earth for future generations to come. One vital resource is power - an essential asset of life that is used to power homes, buildings, and factories, and recently becoming more popular is transportation. With an ever-increasing world population, the demand for electricity increases. There are two main methods to generate electricity: *fossil fuels* and *renewable energy*. Fossil fuels are derived from decomposing plants and animals (located in the Earth's crust) and are burned for energy. Some examples of fossil fuels are coal, natural gas (methane), and oil. On the other hand, renewable energy is obtained via natural resources that are replenished at a higher rate than it is consumed. These sources (hydro, wind, and recently solar) are bountiful and are highly effective methods to generate power. Additionally, one unique form of energy production is nuclear energy, which generates an immense amount of power but at the expense of risk to nuclear radiation, endangering the health of those who get too close to the nuclear plants.

Fossil fuels have proved to be a trustworthy resource for generating power in the United States, with 60% of the country's power generated from fossil fuels. Despite this, fossil fuels are very harmful to the environment, as burning these fossil fuels releases pollutants into the air endangering individuals with respiratory issues, generation of greenhouse gases leads to a direct cause of global warming, and drastic increases in heat lead to wildfires and overall larger risk to humans (State Impact Center 2020). Not only do fossil fuels directly harm our environment, but these resources are also non-replenishable, and takes millions of years for these plants and animals to decompose into usable fuel. Renewable energy inversely creates significantly less carbon dioxide compared to fossil fuels, and since this energy is derived from a renewable resource, it deems a viable method for energy production in the future. In the United States, a significant portion of electricity providers is privately owned across the country.

Figure 1 shows the distribution of the different types of energy produced by these companies as percentages, and that companies such as AEP, First Energy, and Southern Company are still relying on fossil fuels for their main source of electricity, with the exception that California's PG&E has over 50% of total energy sourced from renewable resources. The goal of this study is first to determine if there is a trade-off between renewable energy and fossil fuels in their contribution to total energy production, and a further study on whether fossil fuels will ever reach 0 in the near future, or assess any issues that we may see to arise. With these results, we can infer from private power provider companies which methods of electricity generation should be invested in for the future.

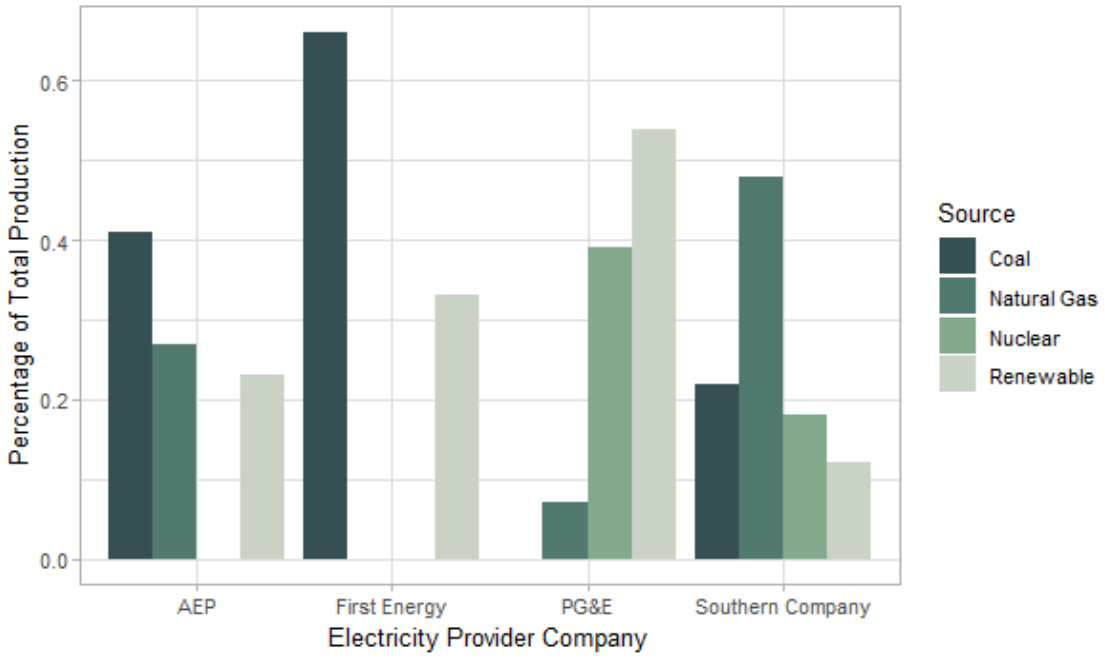


Figure 1: Percentage of Electricity Production Type per Company

In context to today's studies, there have been multiple studies of expected energy production to reach 100% renewable by 2035, as shown in the study National Renewable Energy Laboratory (NREL 2020), who have also run forecasts that have noticed a shift away from fossil fuels and instead to renewable energy, including further analysis of motivations and struggles to make this change (NREL 2020). Some studies address the trade-offs of investing more in renewable energy - by increasing environmental quality by investing in renewable energy, nuclear, and financial development at the expense of GDP (Kartal et al., 2023).

Data is collected from EIA (United States Energy Information Administration), which provides public data to be used for the analysis of energy in the United States for policy making and its interactions with the economy and the environment. This data was selected because of its monthly sampling rate to capture any differences in energy production between seasons as well as the data is updated frequently, with its last observation being in December 2022. With recent data, we can provide more insightful and accurate predictions of energy production behavior for the future.

Data Analysis and Experimental Design

Data Description

Data downloaded from EIA includes a total of 264 measurements of electricity production in units of thousand-megawatt hours (kMWHs). Each observation is collected every month from January 2001 to December 2022. Measurements are also broken down into a variety of fuels, including a vast selection of both fossil fuels and renewable energy sources. The data is presented in a neat format with no missing information or any preprocessing in order to make the data interpretable.

As mentioned, there is a multitude of different methods to produce electricity covered in the data set, but methods coal, natural gas, nuclear, wind, and hydro were ultimately due to their prevalence in the United States. Solar is also a prominent point of interest in this study, but we would require further data before performing any meaningful forecasts. Data were split into two sets: training and test sets, where the test set includes only data from January 2022-2023. The test set is used at the end to assess model performance.

Throughout this study, we will be adopting the Box-Jenkins methodology to determine the best models for each data set. The method consists of 5 steps:

1) **Data Preparation:** Data will be transformed to obtain stable variance as well as differenced to obtain stationary series
2) **Model Selection:** Analysis of ACF and PACF plots to determine possible model parameters
3) **Estimation:** Testing models parameter estimates and selecting models based on AICc
4) **Model Evaluation:** Fitting models and analyzing ACF/PACF of residuals and portmanteau tests
5) **Forecasting:** Forecasting models against testing data to assess model performance

Data Tidying & Empirical Analysis

Before we begin fitting our data to the models, we should first make sure our model is easy to work with. We can see that the data is in reverse order as well as the columns can be renamed for clarity. Additionally, to validate our model performance, we will be training models using data from 2001 to 2021, and forecasting electricity production values for 2022. This way we can assess whether our models are accurate or not by comparing our forecasted values to our test set. Let's begin with some basic plotting of our data sets.

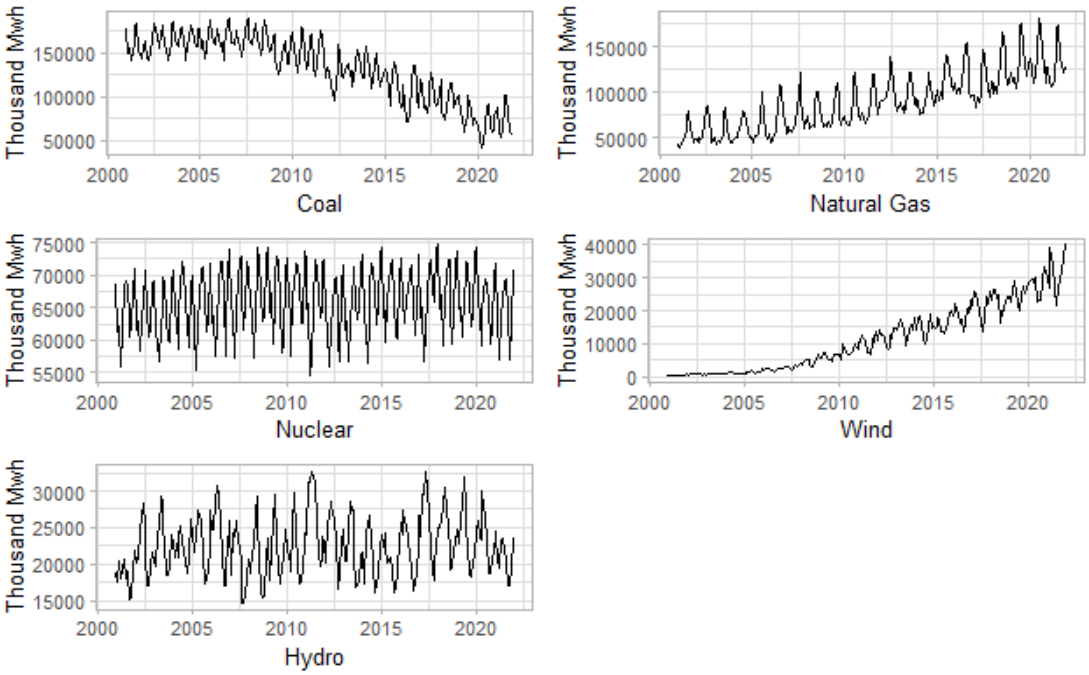


Figure 2: Time Series Plot of Data

Initial Plotting

Before we run any models on these data sets, we need to decompose each of them to determine any properties of the time series, such as seasonal trends or non-seasonal trends. The histograms are shown below (Figure 3):

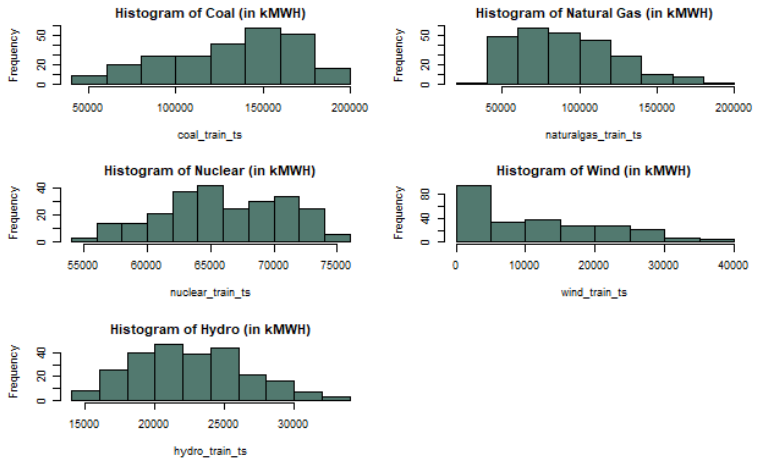


Figure 3: Histograms of Unprocessed Data Sets

We observe the trend graph for each data set, we can see that coal and natural gas exhibit a linear increase trend, wind exhibits a quadratic/ exponential trend, and nuclear and wind exhibit unclear trends. It is notable that there is a seasonal structure across all datasets, which could be attributed to the weather and climate of the United States throughout the year. Coal and natural gas have a strict increase in electricity production in the summer, caused by the increase in demand to cool homes as summer temperatures rise.

Conversely, renewable energy generates more electricity based on the environment. Wind plants are mostly distributed along the California coast and the central United States due to their highest amount of winds throughout the year (Bojek 2022). The movement of storms and pressure imbalances creates annual highs in winds during winter and spring (WRCC 2016), which is reflected in the time series plot (Figure 2). Similar patterns in seasonality occur for hydroelectric production, as more hydroelectric power plants (for example, in Oregon) have increased flow in spring and summer, which is also reflected in the data.

When observing our histograms, we immediately noticed that the distributions for Coal, Natural Gas, and Wind are skewed, and it is best advised to proceed with a Box-Cox transformation to ensure our data resembles a normal distribution and therefore has a stable variance, and further data tidying to ensure stationary series.

Data Transformations and Differencing

Prior to fitting data to time series models, we need to ensure that our data is both normal and stationary to ensure that our forecasts are stable and any predictions do not change over time. This way we can produce meaningful forecasts using our models. As we perform our changes to the data, we will

be using the Shapiro-Wilk (SW) test for the normality of the data sets after performing BoxCox transformations. The Shapiro-Wilks test statistic is:

$$W = \frac{(\sum_{i=1}^{n} a_i y_i)^2}{(\sum_{i=1}^{n} y_i - \mu)^2}$$

The test statistic is generated from getting the squared sum product of coefficients (produced from multiplying expected order statistics with covariance matrix) with sample y_i divided by the squared sum of samples difference from the mean. The hypotheses can be modeled as H_o: The data is normally distributed, and H_1: The data is not normally distributed. As we perform these changes to the data, we want to finish with these tests with p-values > 0.05.

To test for stationarity of data, we will run the Augmented Dickey-Fuller Test (ADF Test), which applies the data to a regression time series model to determine whether the root generate is a unit root. If there is a presence of a unit root, it would imply that the series is not stationary. More specifically the regression model the data is fitted to is:

$$Y_t = \mu + \beta t + \alpha Y_{t-1} + \sum_{j=1}^{p} \phi_j \Delta Y_{t-j} + \varepsilon_t$$

Where α is the coefficient of first lag (what we're testing if it is a unit root) on Y and Y_{t-1} is the first lag, and δY_{t-j} is the first difference at time $t - j$. In short, we want to test if H_0: the series is not stationary ($\alpha = 1$), and H_1: the series is stationary ($\alpha \neq 1$).

The Box-Cox transformation is a technique used to transform a set of data into a new one that resembles a normal distribution. The Box-Cox method estimates a parameter λ that dictates which transformation to apply to the data set. In Figure 3, λ is estimated by each graph's 95% confidence interval. For example, tables Natural Gas and Hydro capture $\lambda = 0$. For more abstract intervals between two λ values such as Coal, each transformation was tested, and whichever value yielded a more close resemblance to a normal distribution was ultimately selected. One discussion needed in particular is the Box-Cox transformation for the Wind data set. The 95% confidence interval for the wind data set yielded a significantly small window, which does not capture any formal λ value for a transformation, that is, it does not capture 0, 0.5, or 1. In this case, the same methodology was followed for coal as we continued to test possible λ values until we had a data set that closest resembled a normal distribution. What that alludes to is that our forecasts won't be as accurate as there wasn't a definite optimal transformation for the data, and it will likely be reflected when comparing forecasted values to our testing set.

In regard to differencing, we want to remove any seasonal and non-seasonal trends to ensure a stationary series prior to model fitting. Differencing refers to calculating the difference between observations of time series. The first differencing is to remove the seasonal trend – given that our data is monthly data, we will difference the data with lag = 12. Furthermore, we will perform an additional difference with lag = 1 to remove non-seasonal trends, the equations below are used to calculate the respective transformation:

Seasonal Differencing at lag = 12: $(1 - B^{12})Y_t = Y_t - Y_t$

Non-Seasonal Differencing at lag = 1: $(1 - B)Y_t = Y_t - Y_t$

After our transformation and differences, we have our data sets that satisfy the normal and stationary requirements, which means we can then proceed to determine a suitable model for each data set, as shown in Table 1 and Figure 4.

Data Set	Box-Cox Lambda	Box-Cox Transformation	Diff. at Lag = 1 SW Test	ADF Test
Coal	1	None	0.678	<0.01
Natural Gas	0.5	$\sqrt{y}$	0.2635	<0.01
Nuclear	1	None	0.842	<0.01
Wind	0	$log(y)$	0.5487	<0.01
Hydro	0	$log(y)$	0.1142	<0.01

Table 1: Data Sets with respective Transformation and Test Results

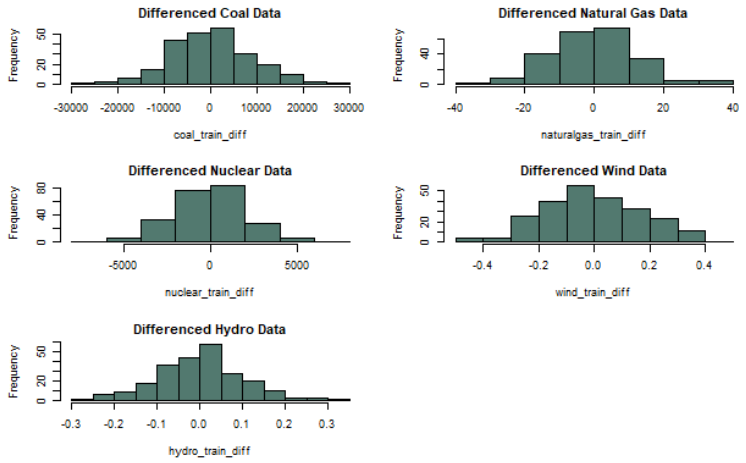


Figure 4: Histogram of all Data Sets after Transformation and Differencing

Time Series Modeling

This section is dedicated to the process of determining a valid model for each of our data set to be forecasted against our testing set. The steps to determine our best models are as follows:

1. **Model Parameter Selection:** Analysis of ACF and PACF graphs to determine suitable parameters for model fitting
2. **Model Fitting:** Runs through a selection of variables to determine the best 3 models using AICc value
3. **Model Validation:** Determines the absolute best SARIMA model by analyzing residuals.

For each data set model identification is performed using ACF and PACF plots. The ACF (autocorrelation function) and PACF(partial autocorrelation function) will help us estimate the order of autoregressive (AR) and moving average (MA) terms for a time series model if it exists. The ACF and PACF can be calculated as follows:

$$\text{(ACF)}\quad \hat{\rho}(k) = \frac{\sum_{t=k+1}^{T}(Y_t - \hat{\mu})(Y_{t-k} - \hat{\mu})}{\sum_{t=1}^{T}(Y_t - \hat{\mu})^2}$$

$$\text{(PACF)}\quad \phi_{k,k} = \frac{\rho(k) - \sum_{j=1}^{k-1}\phi_{k-1,j}\rho(k-j)}{1 - \sum_{j=1}^{k-1}\phi_{k-1,j}\rho(j)}$$

Where $\hat{\mu}$ is the sample mean, and Y_t denotes lag for time = t, and $\phi_{k,j} = \phi_{k-1,j} - \phi_{k,k}\phi_{k-1,k-j}$ for $j = 1, 2, .., k-1$.

Coal

We will begin with parameter estimates for the Coal data.

Model Estimation

Refer to the following ACF/PACF plots below:

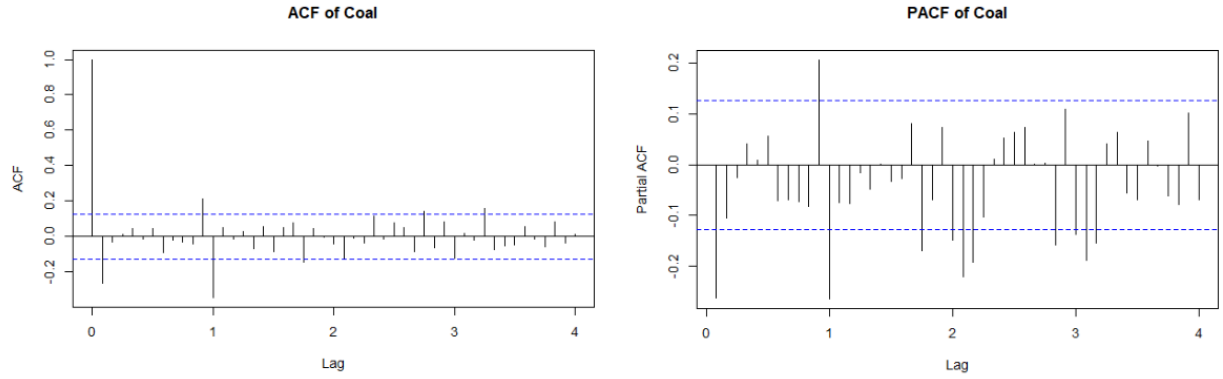


Figure 5: ACF/PACF of Differenced Coal Data

From a holistic view of both plots in Figure 5, we see both aspects of MA and AR terms in the ACF and PACF plots, thus we should consider finding terms for an ARMA model. Since we have dealt with seasonal and non-seasonal differencing, we will expect to extract seasonal and nonseasonal parameters from these plots. Therefore, to include both seasonal differencing and non-seasonal differencing, we will be finding parameters for a SARIMA model. Also written as SARIMA(p, d, q)x(P, D, Q)$_{S=n}$, the SARIMA model encompasses both AR(p) and MA(q) terms of a series as well as non-seasonal and seasonal differencing (d, Q) and seasonal AR(P) and MA(Q) terms, with the assumption that residuals are White Noise, that is constant variance and mean of 0. The SARIMA model can be written out as

$$\phi(B)\Phi(B^S)Y_t(1 - B)(1 - B^{12}) = \theta(B)\Theta(B^S)\epsilon_t$$

Where:

- $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$
- $\phi(B) = 1 - \phi_1 B^S - \phi_2 B^{2S} - \ldots - \phi_p B^{pS}$
- $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_p B^p$
- $\Theta(B) = 1 + \Theta_1 B^S + \Theta_2 B^{2S} + \ldots + \Theta_p B^{pS}$

We have also included terms $(1 - B)(1 - B^{12})$ to account for the differencing (lag 1 and lag 12, respectively). From the ACF plot, we will be estimating possible values for the MA term of the time series model. We can see that there are significant spikes at lag 1, 11, and 12. It must be noted that "Lag" in these plots refers to the seasonal lags, as we have differenced the data at lag = 12. The rule of thumb will be followed to select both non-seasonal and seasonal parameters for the time series models:

- **Non-Seasonal:** Observe spikes within the first season (in this case, first 12 lags
- **Seasonal:** Observe spikes within lags 12, 24, ... as you want to focus on the lags between the end of each season (intervals of 12)

Therefore, we will select candidates q = 0, 1, 11, and Q = 1. 0 was chosen as a candidate for q as the past from lag 0 to lag 1 was significant and crosses 0. From the PACF plot, we select p = 1 and 11, and P = 1, 2, 3 using the same methodology. Additionally, since we differenced our data once each for seasonal and non-seasonal, we have a d=D=1, and an S value of 12, because we have a seasonal period of 12 months. Every combination of these parameters will be fitted to a SARIMA model with our transformed data (as the SARIMA model will include differencing) and will determine the best models using Akaike Information Criteria that have been adjusted for small sample sizes (AICc).

$$AICc = (-2log(L) + 2k) + 2k(k+1))/(T - k - 1)$$

Where L is the likelihood of the model, k is the number of parameters in the model, and T is the number of samples. We adjust the information criteria using k and T in order to account for the different model sizes. The top three models we have concluded to run further tests on are:

- **Model 1:** SARIMA(1, 1, 0)x(3, 1, 1)$_{S=12}$ (AICc = 4906.069)
- **Model 2:** SARIMA(1, 1, 1)x(3, 1, 1)$_{S=12}$ (AICc = 4906.069)
- **Model 3:** SARIMA(1, 1, 0)x(2, 1, 1)$_{S=12}$ (AICc = 4912.311)

For each of these candidate models we will analyze their residuals to confirm that they are 1) normal and 2) stationary and 3) resemble white noise. The first model will be used as an example. In Figure 6, the histogram weakly resembles a uniform distribution, so we can assume normality. Additionally, the ACF plot seems to yield fairly good stationarity, whereas PACF yields weak stationarity. To test for stationarity, we perform a few portmanteau tests such as Box-Ljung and Box-Pierce. Box-Pierce is a statistical test to determine whether H_0: autocorrelations are independently distributed (stationary) or H_1: The autocorrelations are not independently distributed (not stationary). The test statistic is computed using

$$Q_{BP} = T \sum_{k-1}^{K} \hat{\rho}_{\epsilon}^2(k)$$

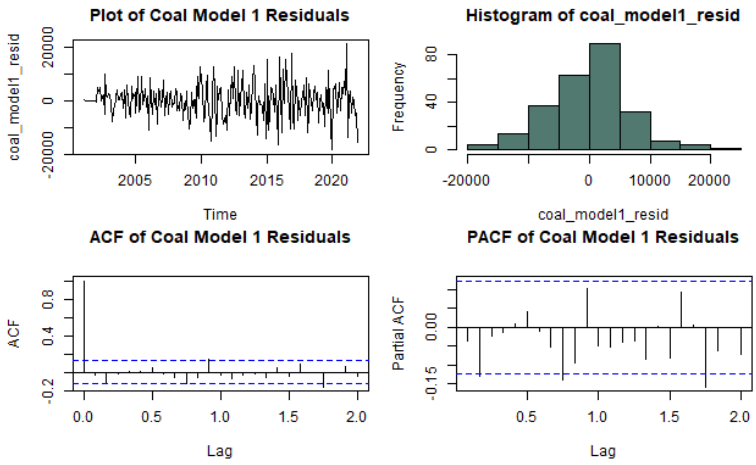


Figure 6: Residuals for Coal Model 1

Under the χ^2_h distribution where K is the number of degrees of freedom (lag subtracted by number of model parameters) and T is the total length of the sequence after differencing. We aim to yield a p-value greater than 0.05 to show that our residuals are uncorrelated. Additionally, Box-Ljung is an identical test to that of Box-Pierce (same hypotheses), but more accurate as it accounts for sample size. The test statistic is computed using

$$Q_{LB} = T(T+2)\sum_{k-1}^{K}\frac{\hat{\rho}_\epsilon^2(k)}{T-k}$$

Under a χ^2_k distribution, where T is the total length of the sequence after differencing, $\hat{\rho}_\epsilon^2(k)$ is the squared sample autocorrelation at lag k. Similarly, we also want to yield a p-value greater than 0.05. Lastly, is the Yule-Walker test, which tests whether a series is an autoregressive (AR) process – if the test yields an AR(p) model where p >0, then the roots of the AR(p) model must be checked to be outside of the unit disc to confirm the model is still stationary. The process of the test includes calculating the autocorrelations up to order p of the AR(p) model, and if any significant terms are identified designates the presence of an AR model at that lag p. After performing the following tests, we have failed both portmanteau tests with p-values under 0.05. We then considered the 4th best model (denoted as New Model 1) with parameters p = 1, q = 1, P = 2, Q = 1, and this newly fitted model passed all tests. The final test to be performed is the unit root test, to confirm that the roots of each component of the model (AR, SAR, MA, SMA) lie outside of the unit disc to show that the series is stationary, plots for all models for Coal can be viewed in Appendix A1. As all tests have passed, we conclude that our first top model is a SARIMA model, written as

$$(1 - 0.2135B)(1 - 0.1566B^{12} + 0.1415B^{24})(1 - B^{12})(1 - B)(Y_t)$$
$$= (1 - 0.5354B)(1 - 0.8720B^{12})\epsilon_t$$

For the second-best model, the histogram appears to have a better distribution of residuals than the previous models, as well as the ACF and PACF including less significant spikes. Additionally, all diagnostic tests of the model have passed including the unit circle tests. The written model for this equation is

$$(1 - 0.2013B)(1 + 0.0543B^{12} + 0.1609B^{24} + 0.2154B^{36})(1 - B^{12})(1 - B)(Y_t)$$
$$= (1 - 0.5023B)(1 - 0.7880B^{12})\epsilon_t$$

Lastly, our third best model yields similar results to New Model 1, but still passed all diagnostic tests and unit circle checks. The final equation for the model was:

$$(1 - 0.2220B)(1 - 0.1773B^{12})(1 - B^{12})(1 - B)(Y_t)$$
$$= (1 - 0.5245B)(1 - 0.9173B^{12})\epsilon_t$$

Thus we have compiled our top 3 performing models, and their graphical results can be viewed in Figure 7 and Table 2.

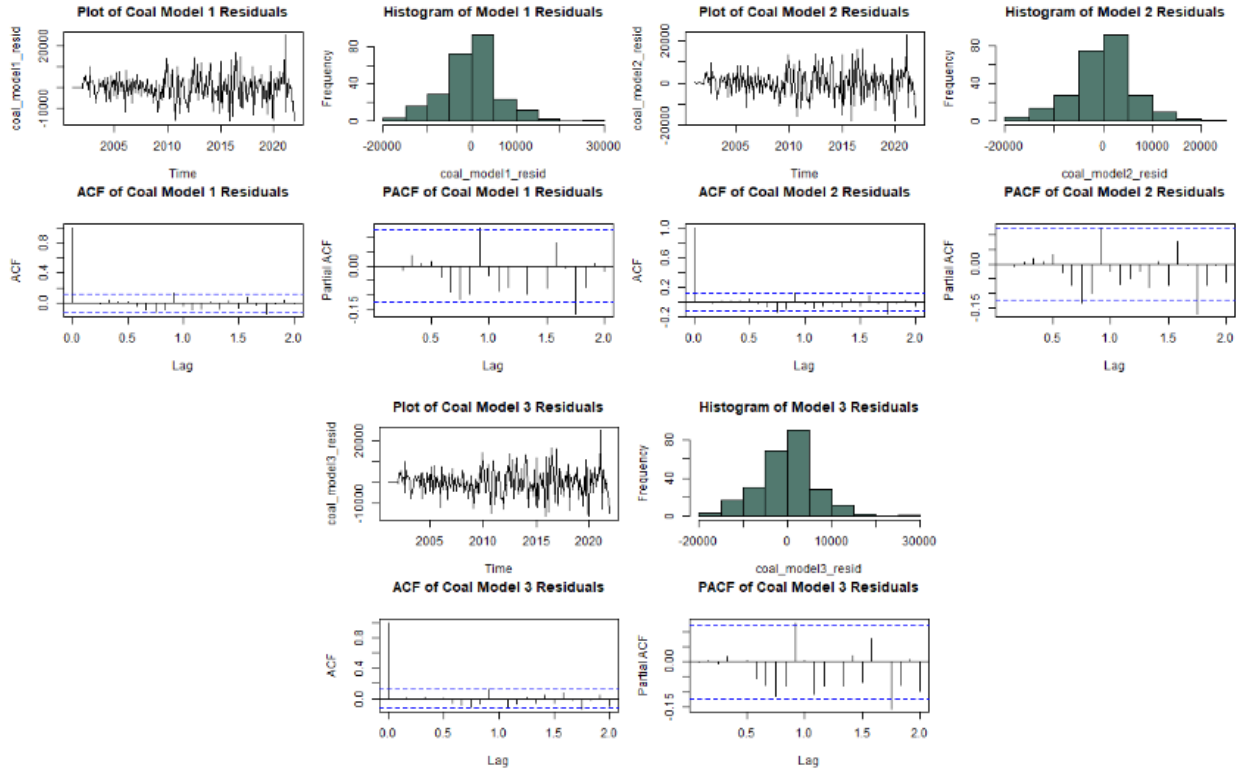


Figure 7: Residual Graphs for Coal Data

Model #	Model	Box-Pierce (Residuals)	Box-Ljung (Residuals)	Yule-Walker (Residuals)	Unit Circle (Model)	AICc
Model 1	SARIMA(1, 1, 0)x(3, 1, 1)$_{S=12}$	0.04306	0.03347	AR(2), Pass	N/A	4906.069
New Model 1	SARIMA(1, 1, 1)x(2, 1, 1)$_{S=12}$	0.08686	0.06754	AR(0)	Pass	4912.311
Model 2	SARIMA(1, 1, 1)x(3, 1, 1)$_{S=12}$	0.1034	0.08289	AR(0)	Pass	4906.069
Model 3	SARIMA(1, 1,0)x(1, 1, 1)$_{S=12}$	0.1547	0.1246	AR(0)	Pass	4912.640

Table 2: Summary of Best Models and Results for Coal Data Set

Model Forecasting

When we forecast our models, we will be comparing our predicted values to our test set, which is the 12 observations of 2022 and we will be measuring model performances between other models of the same data set. Figure 8 shows the projected forecasts for 2022 (red) are plotted against the actual values (blue) for each model, with grey bounds being confidence intervals..

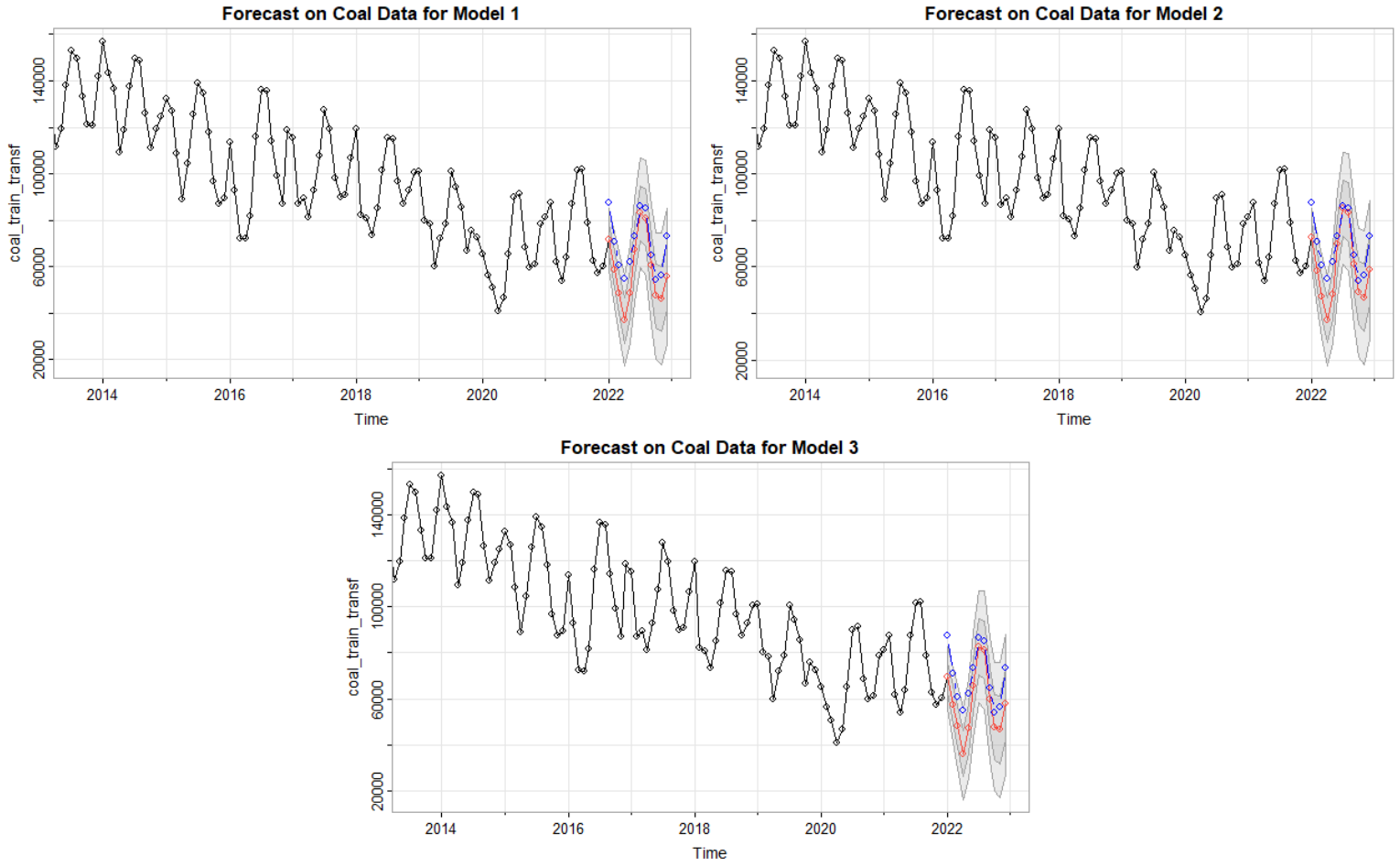


Figure 8: Forecasts for Best 3 Coal Models (Red = Forecasted Values, Blue = True Values)

From our forecasts, we can immediately see that our predicted values were more extreme when predicting dropping values than the actual values, this could likely be due to the massive drop in electricity production around 2020, likely caused by COVID-19 – as social restrictions to mitigate the spread of virus caused a decrease in production. The model likely adopted this drastic drop into its forecasts. When comparing between models, we will use the Forecast Mean Squared Error Metric (FMSE) which is computed as

$$FMSE_h(T) = \frac{1}{h}\sum_{j=1}^{h}((y_T + j - \hat{y}_{T+j})^2$$

Where T is the given time and h is the step, it is noted that for any future models that were transformed, predicted values were back-transformed to compare with unmodified test data.

Model #	Model	FMSE
Model 1	SARIMA(1, 1, 1)x(2, 1, 1)$_{S=12}$	1.32e08
Model 2	SARIMA(1, 1, 1)x(3, 1, 1)$_{S=12}$	1.19e08
Model 3	SARIMA(1, 1,0)x(1, 1, 1)$_{S=12}$	1.46e08

Table 3: Coal Model Forecast Results

These FMSE values are exceptionally large, and as said earlier, the large dip in 2020 could attribute to its inaccurate forecasted values. We conclude that Model 2 (SARIMA(1, 1, 1)x(3, 1, 1)$_{S=12}$) is the best performing forecast model.

Natural Gas

Model Estimation

We then proceed to the next data set, where the same methodology applies. The ACF/PACF plots in Figure 9 shows that we similarly identify aspects of both MA and AR terms, and we can identify candidate parameters for our SARIMA model. We will select possible values q = 1 or 2, Q = 1, p = 1, 2, or 11, and P = 1, 2, or 3.

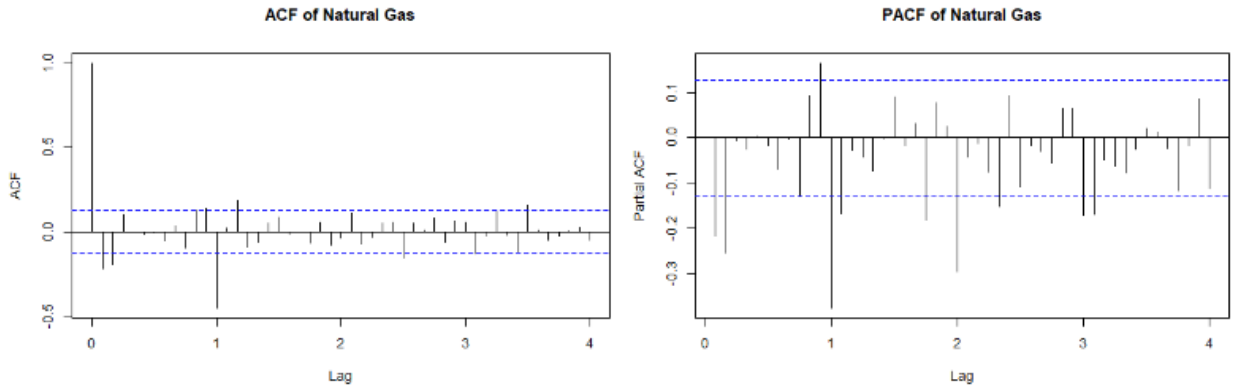


Figure 9: ACF/PACF of Differenced Natural Gas Data

We then proceed to fit all possible models and extract the three models with the lowest AICc values. It is noted that the average AICc values for this data set are significantly lower than that of Coal (around 4900) this is most likely due to the noise in Coal data, making the model more difficult to fit. The top 3 selected models were

- **Model 1:** SARIMA(2, 1, 2)x(1, 1, 1)$_{S=12}$ (AICc = 1721.499)
- **Model 2:** SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$ (AICc = 1722.211)
- **Model 3:** SARIMA(2, 1, 1)x(1, 1, 1)$_{S=12}$ (AICc = 1722.537)

Model 1 achieved the lowest AICc score. This model passed all diagnostic tests as well as unit circle tests. The written equation for the model is

$$(1 + 0.2360B - 0.6982B^2)(1 + 0.02830B^{12})(1 - B^{12})(1 - B)(Y_t)^{1/2} = (1 - 0.0901B - 0.9098B^2)(1 - 0.8617B^{12})\epsilon_t$$

Model 2 was the second best model that passed all diagnostic tests, and distributions from the histogram seem are identical to that of Model 1. Time series plots loosely resemble white noise in that mean is close to 0 but becomes noisier as time increases, which reflects the original data. What this means is that this increase in noise will reflect in the errors in the forecast, which will be explored at the end of this section. ACF/PACF graphs seem to resemble white noise well, and unit circle tests passed as well. The written equation for this model is

$$(1 - 0.7927B)(1 - 0.014B^{12})(1 - B^{12})(1 - B)(Y_t)^{1/2} = (1 - 1.1732B)(1 - 0.8456B^{12})\epsilon_t$$

Lastly is Model 3, which yielded very similar plots to Model 2, therefore we expect similar forecast performance to Model 2. The fitted model passed all diagnostic tests and portmanteau tests as well as successfully passing the unit circle test. The equation for this model is

$$(1 - 0.6277B - 0.1120B^2)(1 - 0.0061B^{12})(1 - B^{12})(1 - B)(Y_t)^{1/2} = (1 - 0.9999B)(1 - 0.8460B^{12})\epsilon_t$$

Additionally, all unit circle tests can be found in Appendix A2.
Figure 10 and Table 4 display the summarized test results for the Natural Gas data set.

Model #	Model	Box-Pierce (Residuals)	Box-Ljung (Residuals)	Yule-Walker (Residuals)	Unit Circle (Model)	AICc
Model 1	SARIMA(2, 1, 2)x(1, 1, 1)$_{S=12}$	0.1783	0.1527	AR(0)	Pass	1721.499
Model 2	SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$	0.07831	0.06084	AR(0)	Pass	1722.211
Model 3	SARIMA(2, 1,1)x(1, 1, 1)$_{S=12}$	0.07106	0.05491	AR(0)	Pass	1722.537

Table 4: Summary of Best Models and Results for Natural Gas Data Set

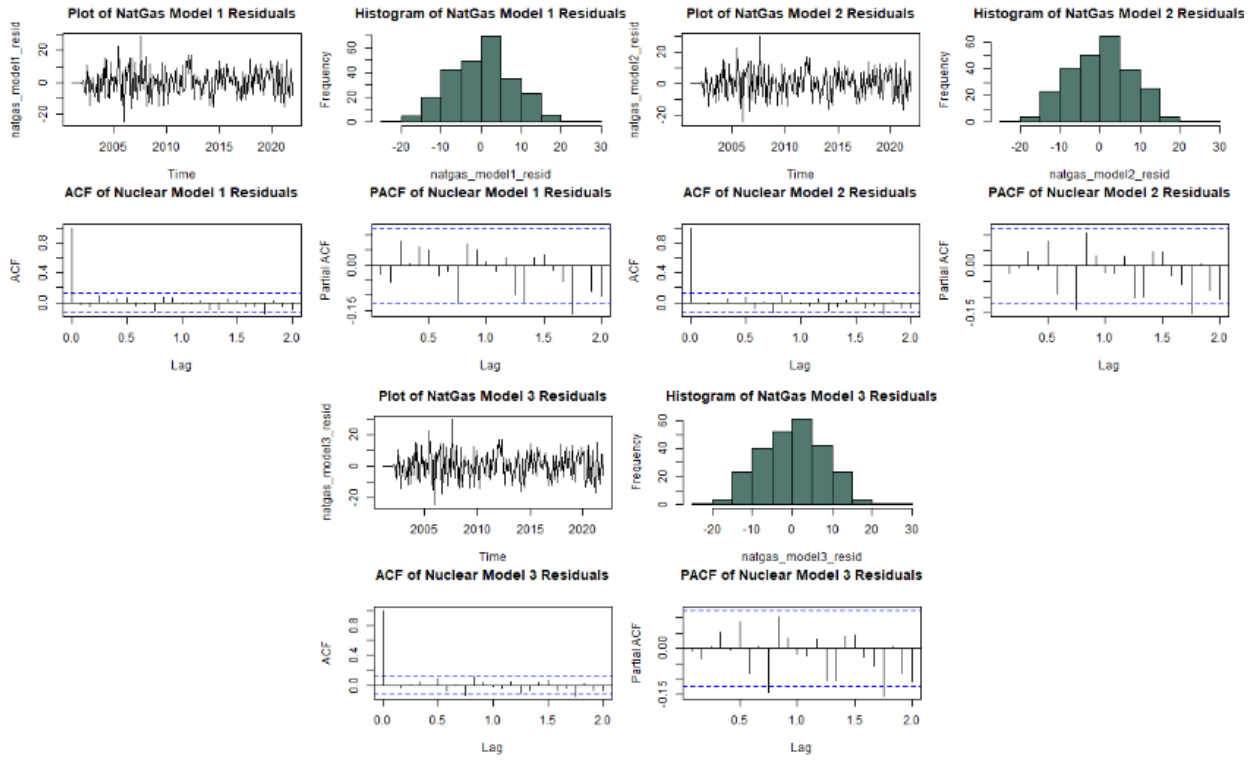


Figure 10: Residual Graphs for Natural Gas Data

Model Forecasting

We then proceed to forecast these models to assess their performance. From Figure 11, Models, 2 and 3 performed relatively well, but were conservative in their forecasts, this was most likely due to the drop in production in 2021, which disrupted the increasing trend in the previous years and predicted the same increase for 2022, but in actuality production of energy of 2022 stabilized as if the drop never occurred. Model 1 is slightly worse compared to the other models, as it predicted even more conservatively. From our three models, we select Model 2 (SARIMA(1, 1, 2)x$(1, 1, 1)_{S=12}$) as our best-performing model, as summarized in Table 5

Model #	Model	FMSE
Model 1	SARIMA(11, 1, 1)x$(1, 1, 1)_{S=12}$	5.24e07
Model 2	SARIMA(1, 1, 2)x$(1, 1, 1)_{S=12}$	4.82e07
Model 3	SARIMA(2, 1,1)x$(1, 1, 1)_{S=12}$	4.83e07

Table 5: Natural Gas Model Forecast Results

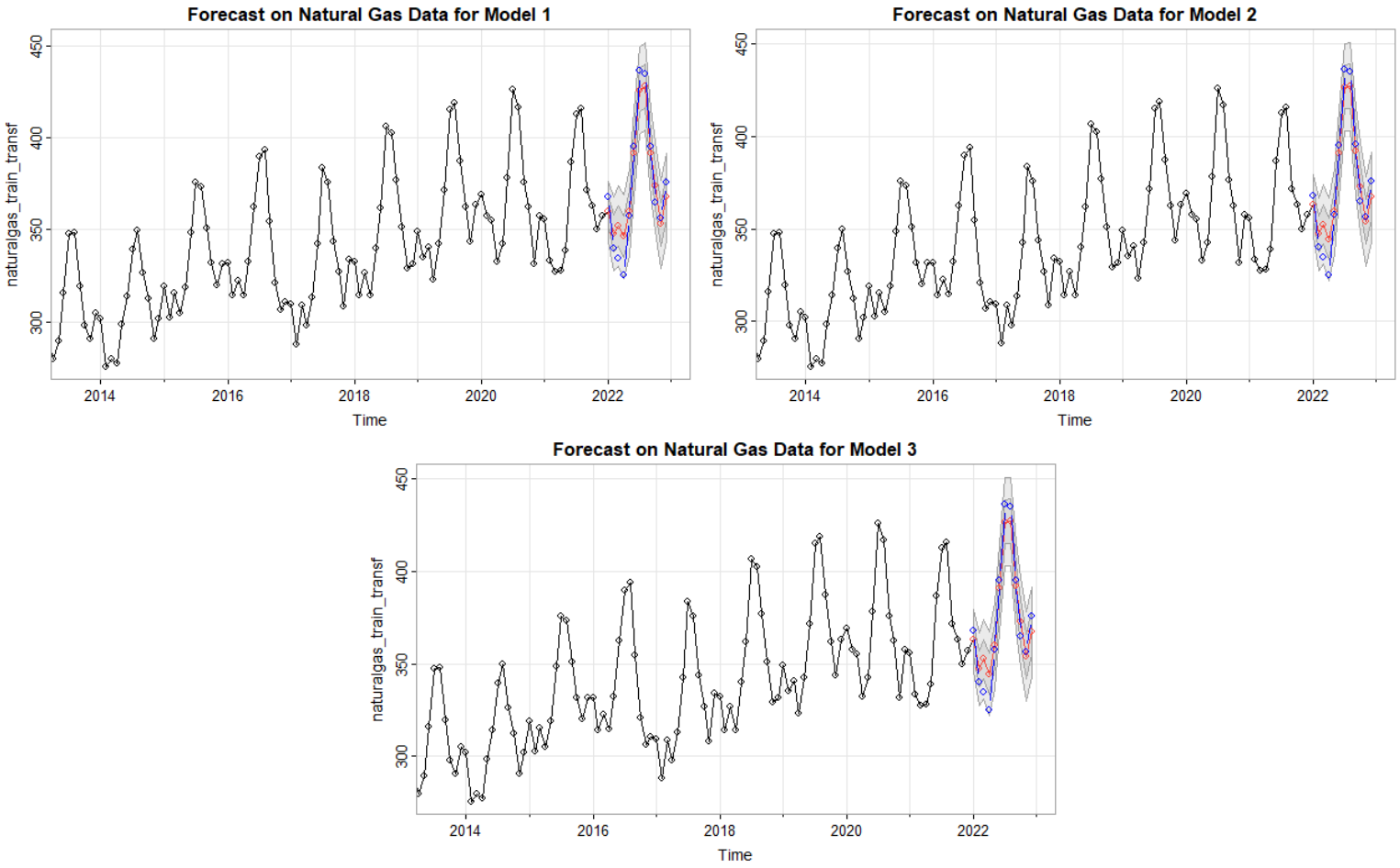


Figure 11: Forecasts for Best 3 Natural Gas Models (Red = Forecasted Values, Blue = True Values)

Nuclear

Model Estimation

The next data set, Nuclear, includes ACF/PACF plots (Figure 12) that have more possible parameters compared to the previous two data sets. We will select possible values q = 1, 2, 6, or 10, Q = 1, p = 1, 2, 6, or 10, and P = 1, 2, or 3. We tested all possible models and concluded with the top three models with the lowest AICc values to run diagnostic tests on:

- **Model 1:** SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$ (AICc = 4192.163)
- **Model 2:** SARIMA(1, 1, 1)x(1, 1, 2)$_{S=12}$ (AICc = 4192.941)
- **Model 3:** SARIMA(1 1, 2)x(1, 1, 1)$_{S=12}$ (AICc = 4194.207)

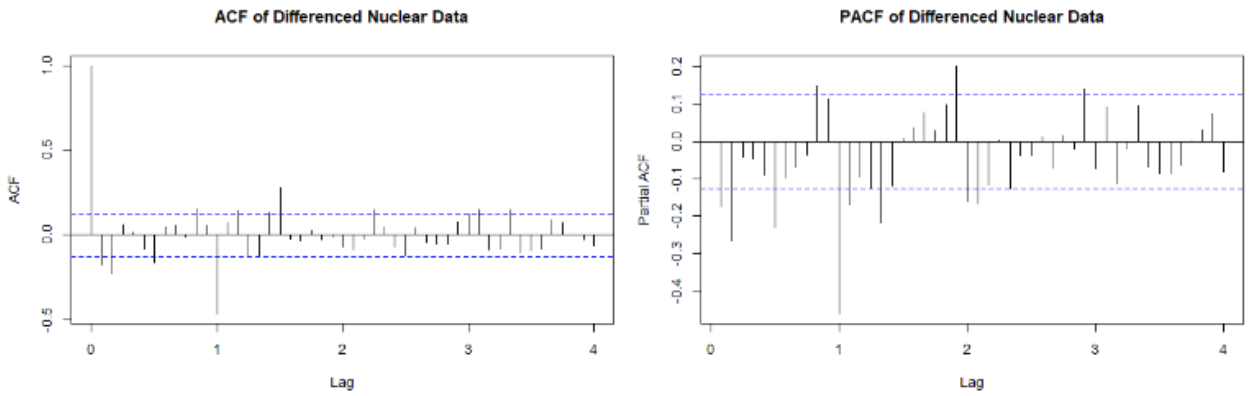


Figure 12: ACF/PACF of Differenced Nuclear Data

Model 1 achieved the lowest AICc, with a small skewed distribution of residuals. This skew would reflect in our errors that will be shown later. The model passed the portmanteau tests as well as the unit circle checks. The finalized equation for this model is

$$(1 - 0.5241B)(1 + 0.0077B^{12})(1 - B^{12})(1 - B)(Y_t)$$
$$= (1 - 0.9281B)(1 - 0.9023B^{12})\epsilon_t$$

Model 2 is the next best model with the second lowest AICc value, with identical performance as Model 1, the distribution of the residuals includes an outlier, and we expect the model to perform about the same as that of Model 1. The finalized equation for this model is:

$$(1 - 0.5239B)(1 + 0.6933B^{12})(1 - B^{12})(1 - B)(Y_t)$$
$$= (1 - 0.9281B)(1 - 0.9023B^{12})\epsilon_t$$

Model 3 yielded a histogram of residuals that more closely resembles Model 1 than Model 2, but ACF/PACF plots are all similar to one another. The finalized equation for this model is;

$$(1 - 0.4912B)(1 - 0.0080B^{12})(1 - B^{12})(1 - B)(Y_t)$$
$$= (1 - 0.8858B - 0.0357B^{2})(1 - 0.9026B^{12})\epsilon_t$$

Figure 13 and Table 6 displays the summarized test results for the Natural Gas data set. As all models have passed, we proceed onward to forecasting. (Unit circle tests found in Appendix A3)

Model #	Model	Box-Pierce (Residuals)	Box-Ljung (Residuals)	Yule-Walker (Residuals)	Unit Circle (Model)	AICc
Model 1	SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$	0.6894	0.6635	AR(0)	Pass	4192.183
Model 2	SARIMA(1, 1, 1)x(1, 1, 2)$_{S=12}$	0.5389	0.5097	AR(0)	Pass	4193.825
Model 3	SARIMA(1 1, 2)x(1, 1, 1)$_{S=12}$	0.5998	0.572	AR(0)	Pass	4194.207

Table 5: Summary of Best Models and Results for Nuclear Data Set

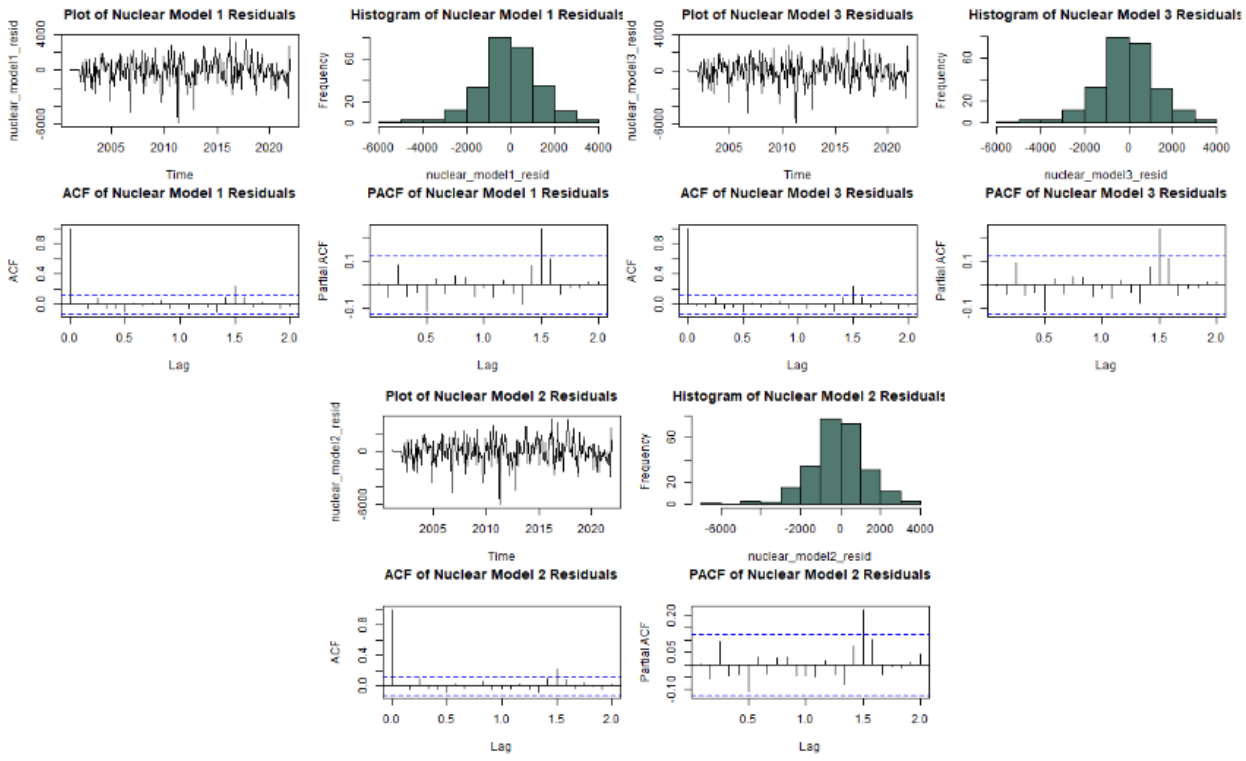


Figure 13: Residual Graphs for Nuclear Data

Model Forecasting

Forecasting these models yielded fairly accurate results (Figure 14). What is present in all three forecasts is that they are all forecasted above the actual value of electricity production in 2022. This is actually due to a sudden decrease in nuclear energy production in 2022, as studies display a sudden increase in decommissioning rates and nuclear plants are financially struggling (Wesoff 2022). As this is an abrupt change that only occurs in the test set, this is a reason why there is this error. Overall, we elect Model 1 to be the best forecasting model for the Nuclear data set (as shown in Table 6).

Model #	Model	FMSE
Model 1	SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$	1.44e06
Model 2	SARIMA(1, 1, 1)x(1, 1, 2)$_{S=12}$	1.54e06
Model 3	SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$	1.73e06

Table 6: Nuclear Model Forecast Results

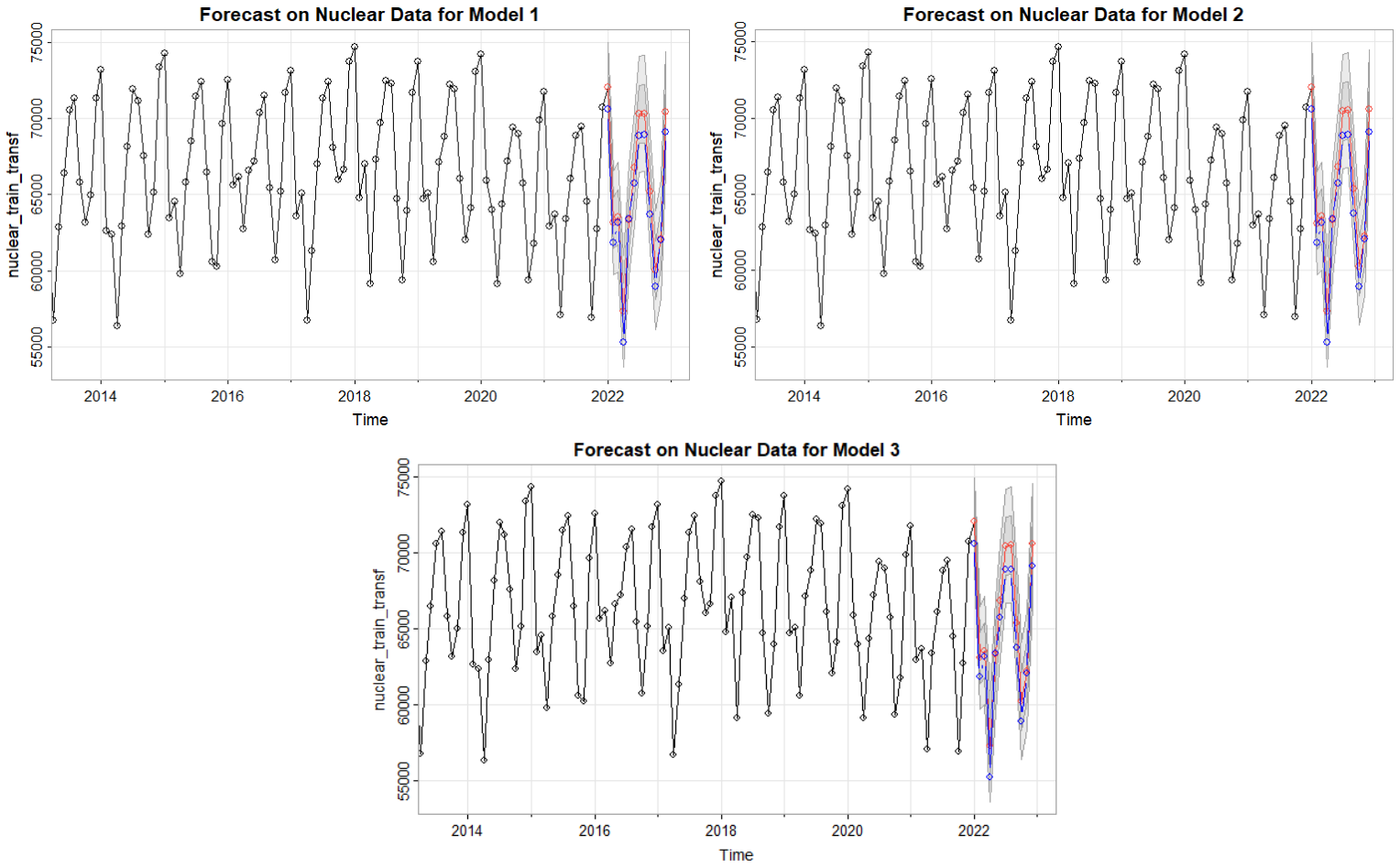


Figure 14: Forecasts for Best 3 Nuclear Models (Red = Forecasted Values, Blue = True Values)

Wind

Model Estimation

Recalling that there wasn't a precise transformation for the wind data, it is reflected in the ACF/PACF plots (Figure 15). There is a larger range of possible parameters to be tested in the SARIMA model.

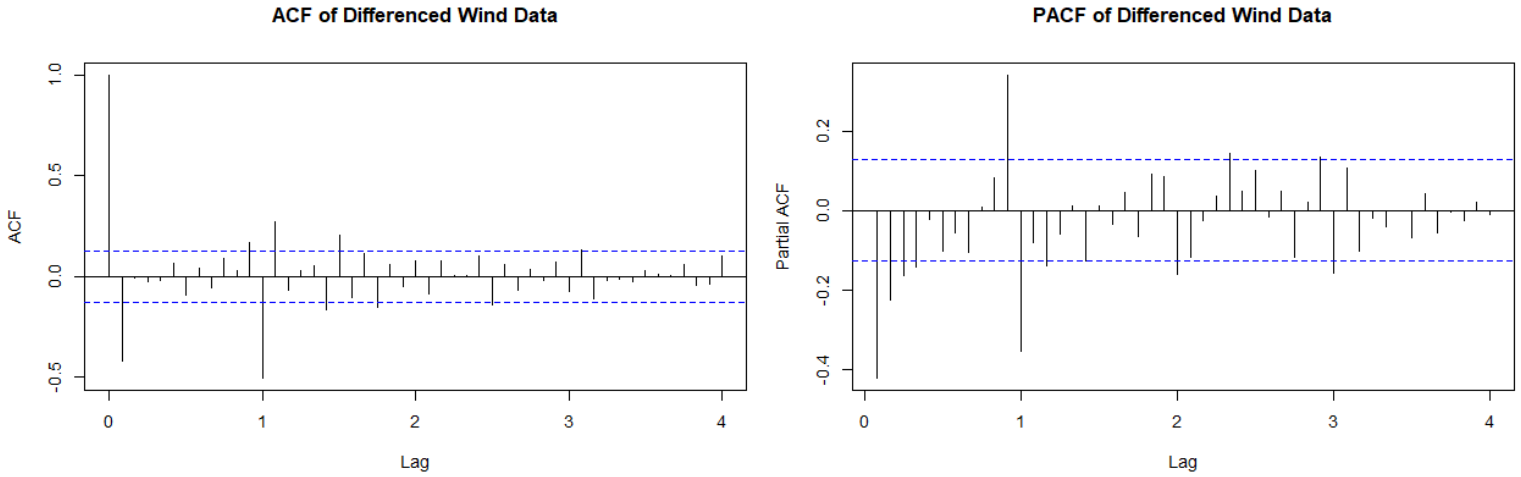


Figure 15: ACF/PACF of Differenced Nuclear Data

For the ACF, we selected significant values of 1 and 11, and for seasonal lags, we selected 1 (q = 1 or 11, Q = 1). As for the PACF, we identified significant values of 1, 2, 3, 4, and 11, and seasonal lags 1, 2, and 3 (p = 1, 2, 3, 4, or 11, P = 1, 2, or 3). After testing all possible models, we have the following results

- **Model 1:** SARIMA(2, 1, 1)x(1, 1, 1)$_{S=12}$ (AICc = -323.4258)
- **Model 2:** SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$ (AICc = -322.5606)
- **Model 3:** SARIMA(3, 1, 1)x(1, 1, 1)$_{S=12}$ (AICc = -321.3492)

Model 1 well resembles a normal distribution from its residuals, and its ACF/PACF plots well resemble white noise. After running diagnostic tests, all tests have passed including the unit circle checks. The final equation for this model is

$$(1 - 0.2624B - 0.1354B^2)(1 + 0.1569B^{12})(1 - B^{12})(1 - B)log(Y_t)$$
$$= (1 - 0.8386B)(1 - 0.6280B^{12})\epsilon_t$$

As for the second model, residuals also represent a normal distribution, yet the ACF/PACF plots have slightly more significant peaks, this means that the lags are a little more correlated, and will most likely have a higher error when forecasting. This model also passes the unit circle check. The model for this equation is

$$(1 - 0.2066B)(1 + 0.1597B^{12})(1 - B^{12})(1 - B)log(Y_t)$$
$$= (1 - 0.7656B)(1 - 0.6142B^{12})\epsilon_t$$

Lastly, Model 3 is identical to Model 2 in terms of the resemblance of normal distribution and to white noise. All diagnostic tests have passed including the unit circle test. All results for unit circle tests can be viewed in Appendix A4. The equation is as follows:

$$(1 - 0.2657B - 0.1354B^2 - 0.0155B^3)(1 + 0.1580B^{12})(1 - B^{12})(1 - B)log(Y_t)$$
$$= (1 - 0.8442B)(1 - 0.6252B^{12})\epsilon_t$$

Figure 16 and Table 7 displays the summarized test results for the Natural Gas data set.

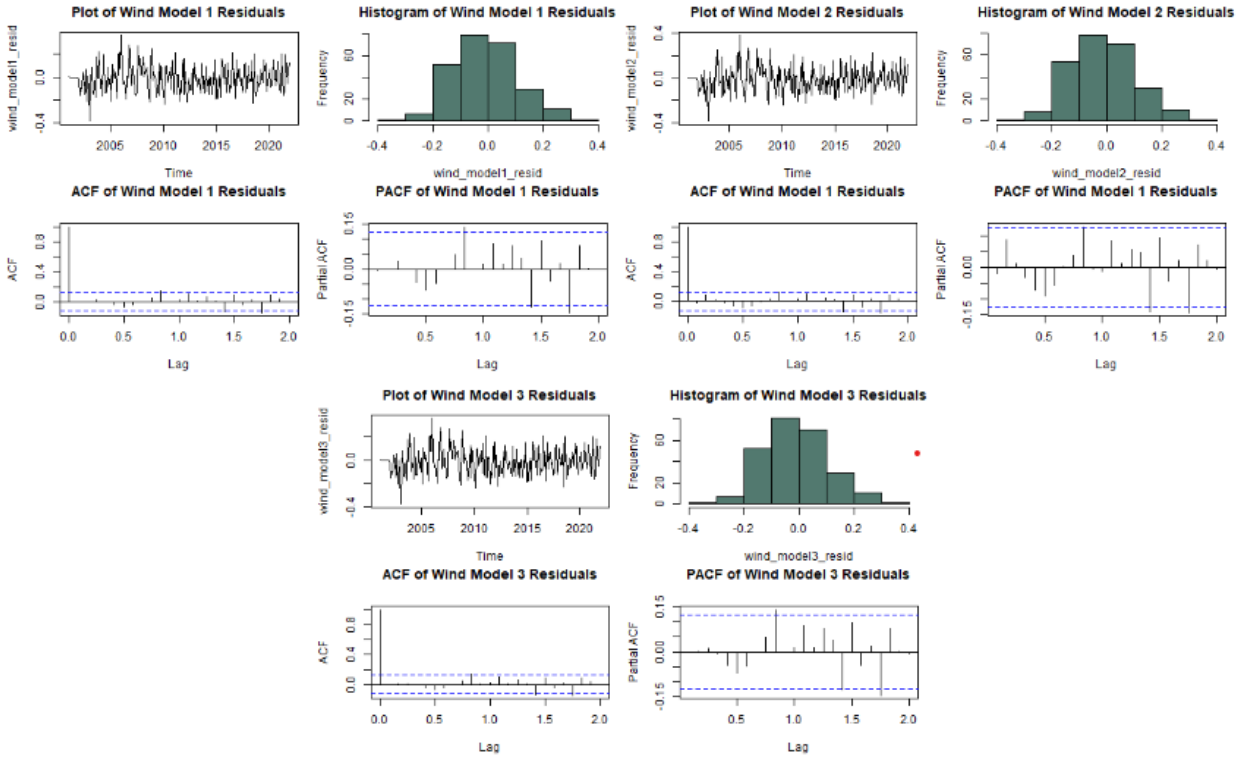


Figure 16: Residual Graphs for Wind Data

Model #	Model	Box-Pierce (Residuals)	Box-Ljung (Residuals)	Yule-Walker (Residuals)	Unit Circle (Model)	AICc
Model 1	SARIMA(2, 1, 1)x(1, 1, 1)$_{S=12}$	0.3282	0.2879	AR(0)	Pass	4192.183
Model 2	SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$	0.2318	0.2008	AR(0)	Pass	4193.825
Model 3	SARIMA(3, 1, 1)x(1, 1, 1)$_{S=12}$	0.2579	0.2229	AR(0)	Pass	4194.207

Table 7: Summary of Best Models and Results for Wind Data Set

Model Forecasting

For wind, our forecasts were not as accurate as in comparison to the Nuclear data, as once again our forecasts were underestimating the peaks. This is likely due to how erratic wind electricity becomes as it becomes more developed, as seen in the trend in Figure 2, the data becomes more turbulent after 2014. This is reflected in the forecast, as the model fails to account for this increasing turbulence. There is also a

a rounded peak instead of the usual peak in 2020 in the production of electricity from wind power which would also deter accurate forecasts. This disruption could be alluded to the COVID-19 pandemic, causing the nationwide shutdown of businesses and construction of possible wind plants and the lockdown decreased the availability of maintenance for wind turbines which may contribute to a rounded peak in electricity production. Out of our three models, we elect Model 3 (SARIMA(3, 1, 1)x(1, 1, 1)$_{S=12}$) as our best-performing forecasting model. Model forecasts along with results can be viewed in Figure 17 and Table 8

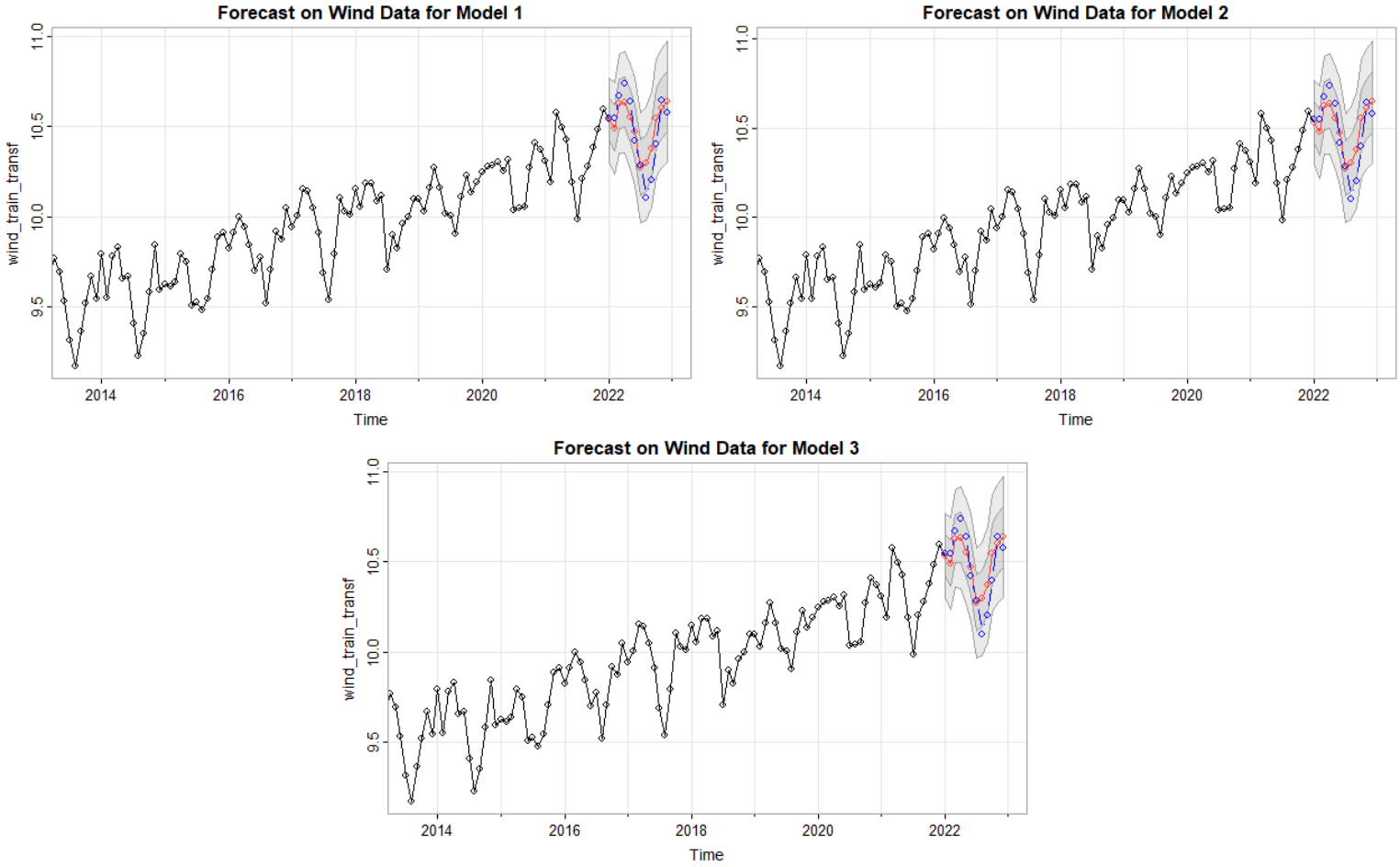


Figure 17: Forecasts for Best 3 Wind Models (Red = Forecasted Values, Blue = True Values)

Model #	Model	FMSE
Model 1	SARIMA(2, 1, 1)x(1, 1, 1)$_{S=12}$	1.078e06
Model 2	SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$	1.078e06
Model 3	SARIMA(3, 1, 1)x(1, 1, 1)$_{S=12}$	1.075e06

Table 7: Wind Model Forecast Results

Hydro

Model Estimation

Hydroelectric, similar to Nuclear has no clear trend and remains relatively stable throughout the years. We will further examine the ACF/PACF to further determine possible candidate parameters for our SARIMA model. From Figure 18, we select significant values of 1 and 2 and significant values at lags 1 and 2 for seasonal terms (q = 1 or 2, Q = 1). For PACF, we selected significant values at 1, 2, and 11, and seasonal terms 1, 2, 3, and 4 (p = 1, 2, or 11, P = 1, 2, 3, or 4) After fitting all possible models we conclude with the following top 3 models:

- **Model 1:** SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$ (AICc = -532.2892)
- **Model 2:** SARIMA(1, 1, 1)x(1, 1, 2)$_{S=12}$ (AICc = -530.3733)
- **Model 3:** SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$ (AICc = -530.1744)

Model 1's residuals well slightly resemble a normal distribution, with data slightly left skewed. Since there is a slight skew in the histogram, it will reflect in the accuracy of the forecast. Both ACF and PACF resemble white noise. All diagnostic tests have passed including unit circle checks, with the equation as

$$(1 - 0.7453B)(1 - 0.1140B^{12})(1 - B^{12})(1 - B)log(Y_t)$$
$$= (1 - 0.9999B)(1 - 0.9999B^{12})\epsilon_t$$

Model 2's residuals were identical to that of Model 1, with no significant spikes in ACF/PACFs and a solid normal distribution from the histogram. All diagnostic tests were performed and all passed, including the unit circle check. Model 2 has concluded to be written as

$$(1 - 0.7469B)(1 + 0.2903B^{12})(1 - B^{12})(1 - B)log(Y_t)$$
$$= (1 - 0.9999B)(1 - 1.1739B^{12} + 0.1734B^{24})\epsilon_t$$

Lastly Model 3 yields similar results to both Models 1 and 2, with good ACF/PACFs and similar normal distribution of residuals. This model as well passed all diagnostic tests and unit circle checks. The equation for Model 3 is written as

$$(1 - 0.6421B)(1 - 0.1058B^{12})(1 - B^{12})(1 - B)log(Y_t)$$
$$= (1 - 0.8696B - 0.0514B^{2})(1 - 0.9999B^{12})\epsilon_t$$

Unit circle checks for all models can be found in Appendix A5. The summary of model test results can be found in Table 8.

Model #	Model	Box-Pierce (Residuals)	Box-Ljung (Residuals)	Yule-Walker (Residuals)	Unit Circle (Model)	AICc
Model 1	SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$	0.6464	0.6162	AR(0)	Pass	-532.2892
Model 2	SARIMA(1, 1, 1)x(1, 1, 2)$_{S=12}$	0.5568	0.5255	AR(0)	Pass	-530.3733
Model 3	SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$	0.5389	0.5055	AR(0)	Pass	-530.1744

Table 8: Summary of Best Models and Results for Wind Data Set

Model Forecasting

When forecasting our data (Figure 18) we have noticed a large error in forecasting. This is due to an unforeseen anomaly in 2022, where there was a record low in hydroelectric power produced. Despite the model continuing seasonal patterns, it is clear that it would not be able to detect this abrupt change. This is due to the fact that there was a record drought in 2022 for the United States, with 1.59 inches below average in rainfall (Bateman 2023). Considering factors such as rainfall have a possible impact on hydroelectricity production, further discussion on extraneous factors will be in the subsequent section.

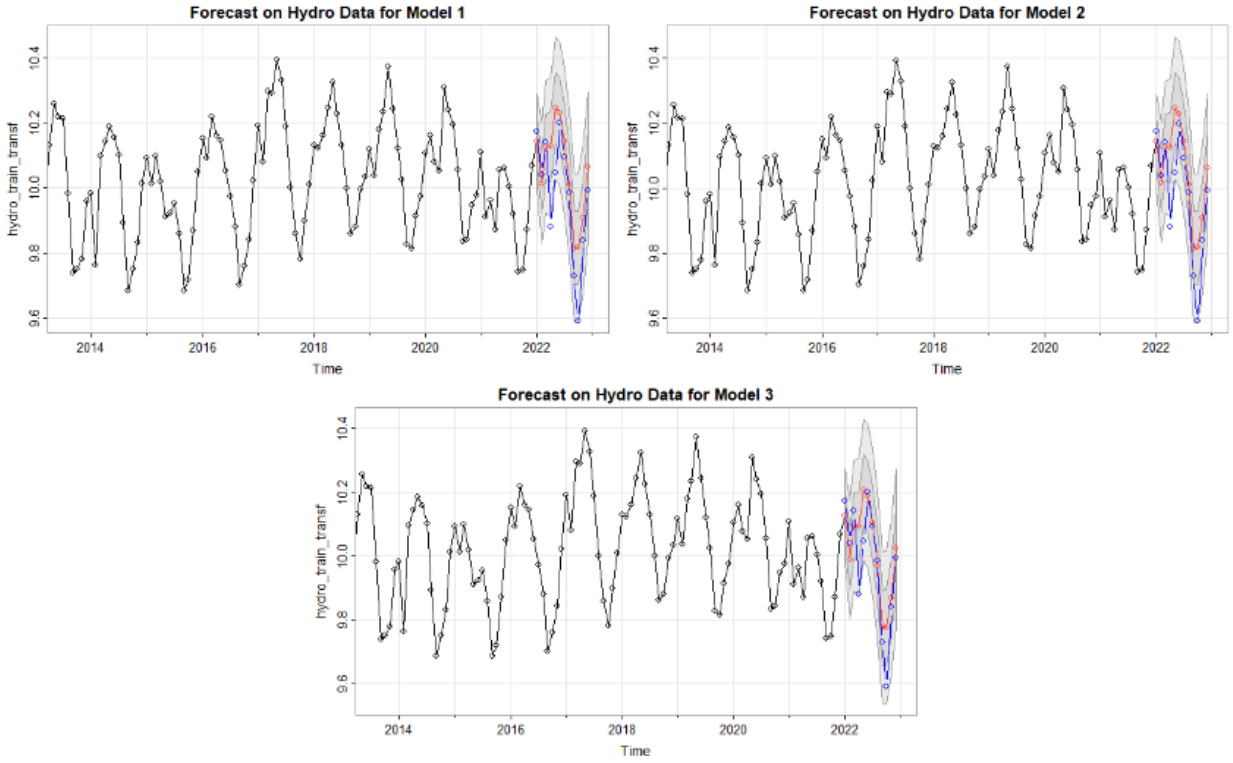


Figure 18: Forecasts for Best 3 Hydro Models (Red = Forecasted Values, Blue = True Values)

Model #	Model	FMSE
Model 1	SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$	6.577e06
Model 2	SARIMA(1, 1, 1)x(1, 1, 2)$_{S=12}$	6.575e06
Model 3	SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$	4.298e06

Table 8: Wind Model Forecast Results

Discussion/Conclusion

The forecasted models on electricity production from different sources (fossil fuels, renewable energy, nuclear) have performed with varying degrees of success, the final table with best performing models can be viewed on Figure 19.

Data Set	Model
Coal	SARIMA(1, 1, 1)x(3, 1, 1)$_{S=12}$
Natural Gas	SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$
Nuclear	SARIMA(1, 1, 1)x(1, 1, 1)$_{S=12}$
Wind	SARIMA(3, 1, 1)x(1, 1, 1)$_{S=12}$
Hydro	SARIMA(1, 1, 2)x(1, 1, 1)$_{S=12}$

Figure 19: Best Models for All Data Sets

The majority of the models have experienced a loss in accuracy to sudden changes in trend due to extraneous events such as COVID-19 or climate change. For example, COVID-19 caused a massive halt in the production and maintenance of power plants, causing an abrupt drop in electricity production and altering trends of the time series models.

One model, in particular, performed fairly well in comparison to the other models, which was the Nuclear model. This is likely due to the fact that Nuclear has produced consistent amounts of electricity in the past for the past 20 years (Figure 2), and any excess production of nuclear power plants risks nuclear radiation to inhabitants within range of it (National Cancer Institute, 2011). Not only that, the cost of building a new nuclear power plant in the United States is incredibly expensive, costing billions of dollars and taking over 10 years to construct.

On the discussion of extraneous factors, the previously stated example was the forecast on the Hydro data set from the record drought occurring in the United States, dropping average precipitation and thus showing evidence of low electricity production in 2022. This potential correlation could allude to the rest of the models – where inaccurate forecasts highlight the pitfalls to just univariate time series models, as all of the models depend on multiple extraneous factors, with climate setting just being one of them. Despite this, each model was fairly able to perpetuate any seasonal trends in their respective forecasts but failed to predict any trends, with the exception of nuclear data. In the context of whether the models are suitable, we believe that these models are not suitable for real-world forecasting, as we require more complex models such as multivariate models that would factor in said extraneous factors, such as weather conditions or cost of resources to produce energy.

In regards to if there is a negative correlation between fossil fuels and renewable energy, it is actually a much more complex question that can be answered purely on electricity production rates – as there may be a general decrease in trend for coal, but an overall increase in electricity production for natural gas, which in itself shows a contradiction in trend among fossil fuels. To further explore this question, we would need additional research into both the pros and cons of additional power plants. Furthermore, we do not believe that we can determine when fossil fuels will reach zero, as our models cannot accurately forecast purely based on their own past values and would require additional information. One point of interest is that electricity produced from natural gas has been and will most

likely continue to increase given the consistent linear increase in trend over the past 20 years. Although more energy is produced this will reflect poorly on the environment, as an increase in the burning of natural gas increases the number of pollutants in the atmosphere, damaging the environment.

In conclusion, the efficacy of our estimated time series models for the different electricity production methods is unsuitable for public use, and should rather consider study into more complex models.

References

AEP (2022). *AEP Generation*. [online] www.aep.com. Available at: https://www.aep.com/about/businesses/generation [Accessed 11 Mar. 2023].

Bateman, J. (2023). *Record drought gripped much of the U.S. in 2022*. [online] www.noaa.gov. Available at: https://www.noaa.gov/news/record-drought-gripped-much-of-us-in-2022.

Bojek, P. (2022). *Wind Electricity – Analysis*. [online] IEA. Available at: https://www.iea.org/reports/wind-electricity.

EIA (2016). *Electricity Data Browser*. [online] Eia.gov. Available at: https://www.eia.gov/electricity/data/browser/.

First Energy (2022). *Generation System*. [online] www.firstenergycorp.com. Available at: https://www.firstenergycorp.com/about/generation_system.html.

Hyndman, R.J. and Athanasopoulos, G. (2016). *Forecasting: Principles and Practice*. [online] Otexts.com. Available at: https://otexts.com/fpp2/.

Kartal, M.T., Samour, A., Adebayo, T.S. and Kılıç Depren, S. (2023). Do nuclear energy and renewable energy surge environmental quality in the United States? New insights from novel bootstrap Fourier Granger causality in quantiles approach. *Progress in Nuclear Energy*, 155, p.104509. doi:https://doi.org/10.1016/j.pnucene.2022.104509.

National Cancer Institute (2011). *Accidents at Nuclear Power Plants and Cancer Risk - National Cancer Institute*. [online] www.cancer.gov. Available at: https://www.cancer.gov/about-cancer/causes-prevention/risk/radiation/nuclear-accidents-fact-sheet#:~:text=At%20high%20doses%2C%20ionizing%20radiation.

NREL (2020). *100% Clean Electricity by 2035 Study*. [online] www.nrel.gov. Available at: https://www.nrel.gov/analysis/100-percent-clean-electricity-by-2035-study.html.

PG&E (2022). *Clean energy solutions*. [online] www.pge.com. Available at: https://www.pge.com/en_US/about-pge/environment/what-we-are-doing/clean-energy-solutions/clean-energy-solutions.page?WT.mc_id=Vanity_cleanenergy.

Southern Company (2021). *Southern Company energy source generation 2021*. [online] Statista. Available at: https://www.statista.com/statistics/646624/share-of-energy-supply-of-southern-company-by-source/#:~:text=Gas%20was%20the%20major%20source [Accessed 11 Mar. 2023].

State Impact Center (2020). *Climate Change and Public Health*. [online] stateimpactcenter.org. Available at: https://stateimpactcenter.org/insights/climate-and-health#healthEffectsOfBurningFossilFuels.

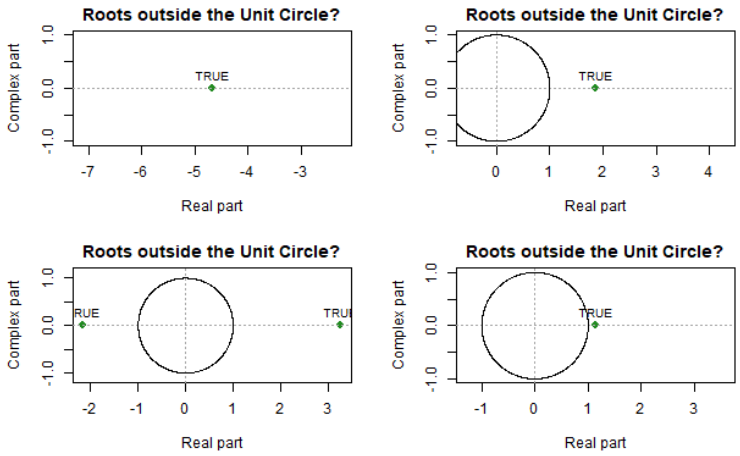
Wesoff, E. (2022). *How did the US nuclear industry fare in 2022?* [online] Canary Media. Available at: https://www.canarymedia.com/articles/nuclear/how-did-the-us-nuclear-industry-fare-in-2022.

WRCC (2016). *Western Regional Climate Center*. [online] Dri.edu. Available at: https://wrcc.dri.edu/Climate/narrative_ca.php.

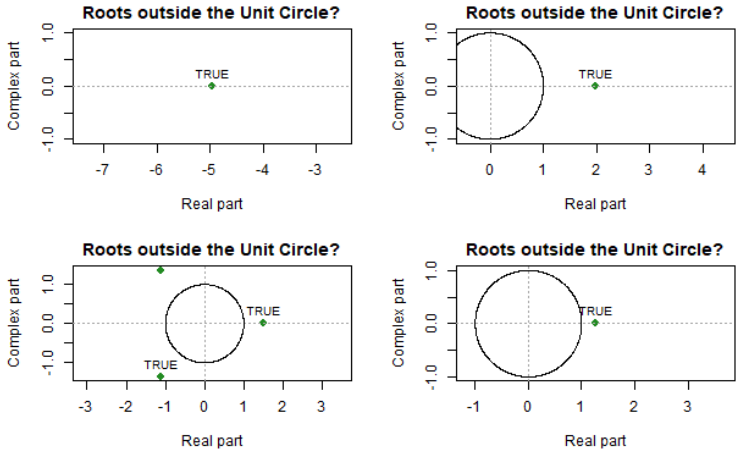
Appendix

[A1] Unit Circle Checks for Coal Data Set

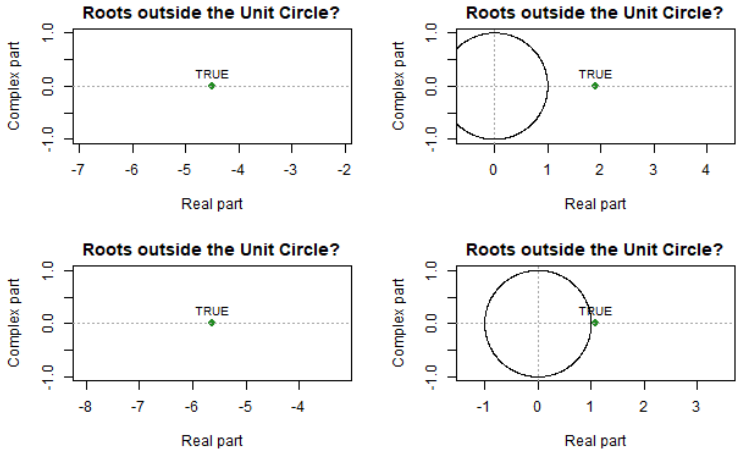
Model 1 (p = 1, q = 1, P = 2, Q = 1) Unit Circle Check



Model 2 (p = 1, q = 1, P = 3, Q = 1) Unit Circle Check

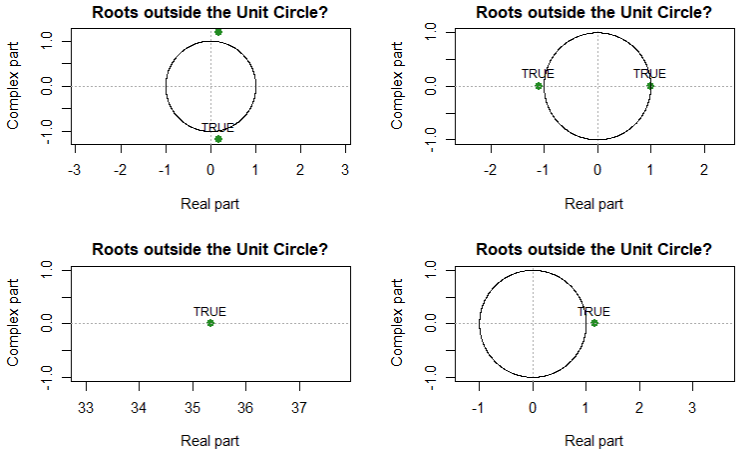


Model 3 (p = 1, q = 1, P = 1, Q = 1) Unit Circle Check

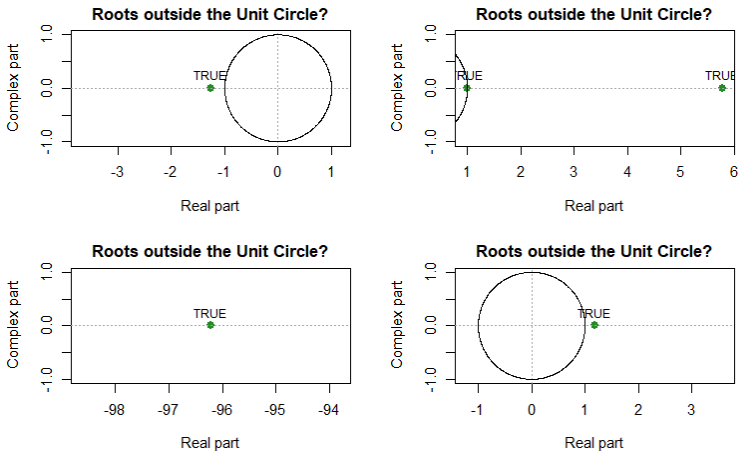


[A2] Unit Circle Checks for Natural Gas Data Set

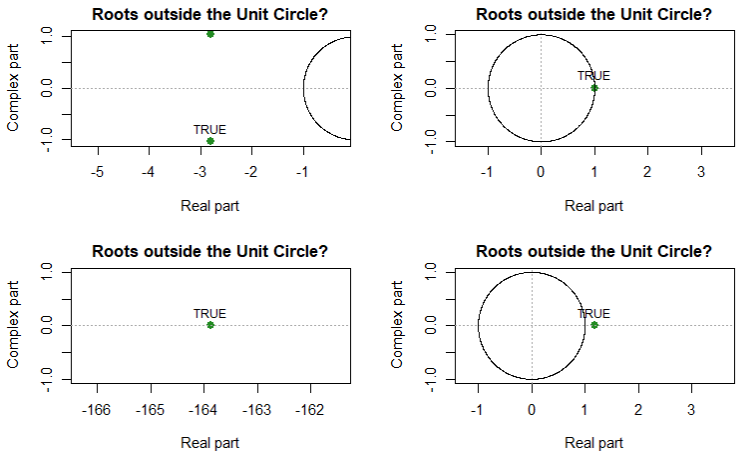
Model 1 (p = 2, q = 2, P = 1, Q = 1) Unit Circle Check



Model 2 (p = 1, q = 2, P = 1, Q = 1) Unit Circle Check

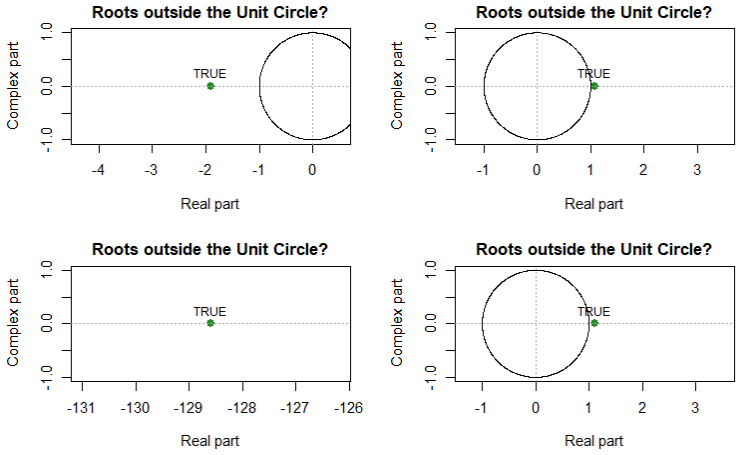


Model 3 (p = 2, q = 1, P = 1, Q = 1) Unit Circle Check

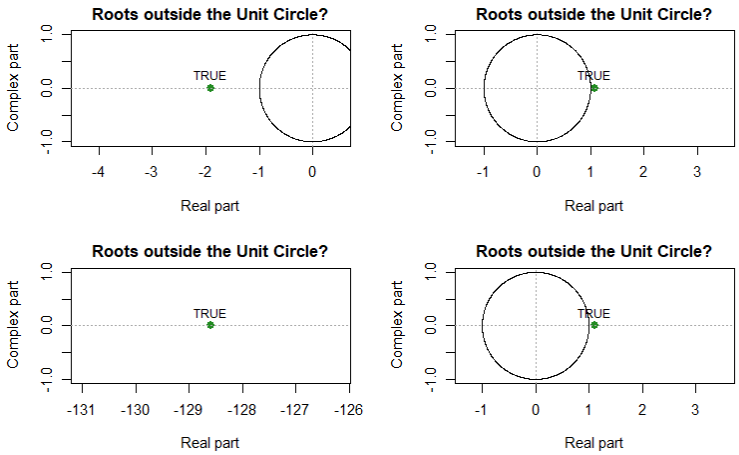


[A3] Unit Circle Checks for Nuclear Data Set

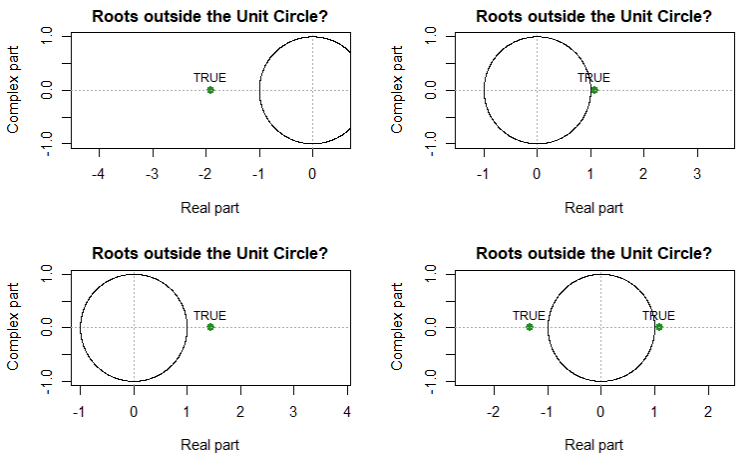
Model 1 (p = 1, q = 1, P = 1, Q = 1) Unit Circle Check



Model 2 (p = 1, q = 1, P = 1, Q = 2) Unit Circle Check

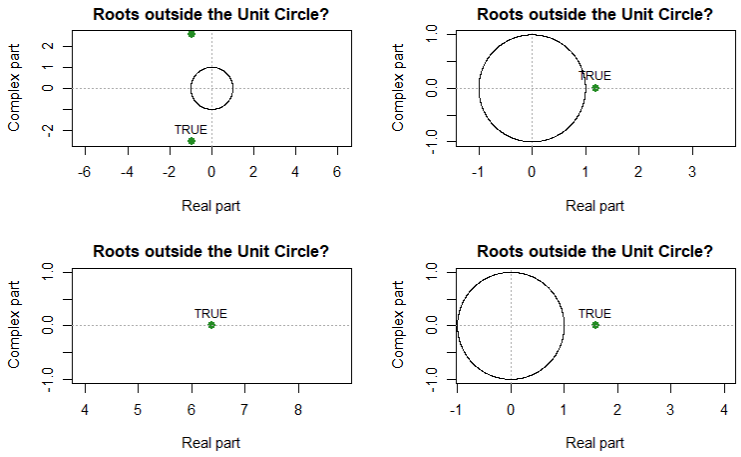


Model 3 (p = 1, q = 2, P = 1, Q = 1) Unit Circle Check

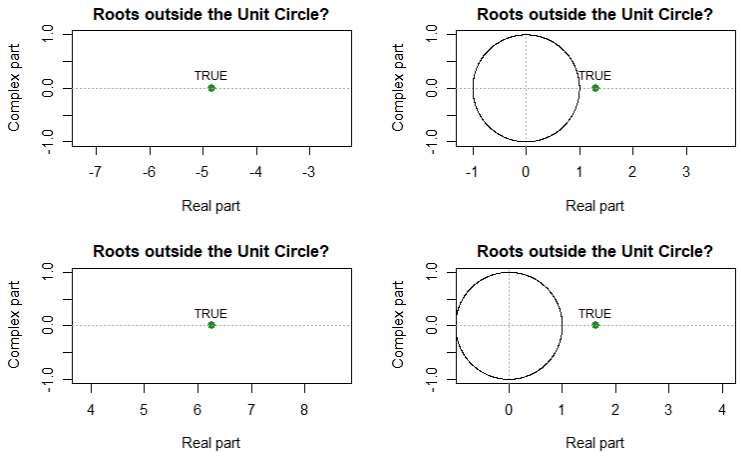


[A4] Unit Circle Checks for Wind Data Set

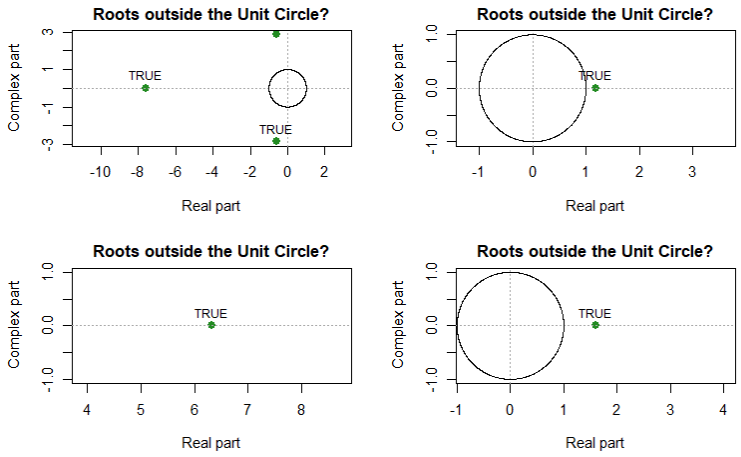
Model 1 (p = 2, q = 1, P = 1, Q = 1) Unit Circle Check



Model 2 (p = 1, q = 1, P = 1, Q = 1) Unit Circle Check

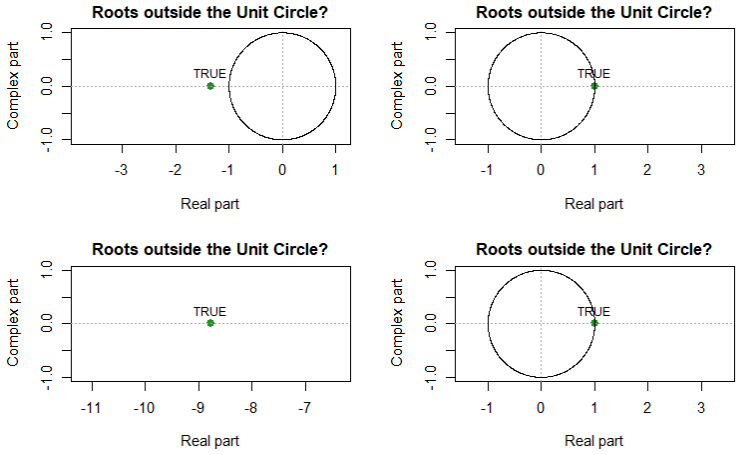


Model 3 (p = 3, q = 1, P = 1, Q = 1) Unit Circle Check

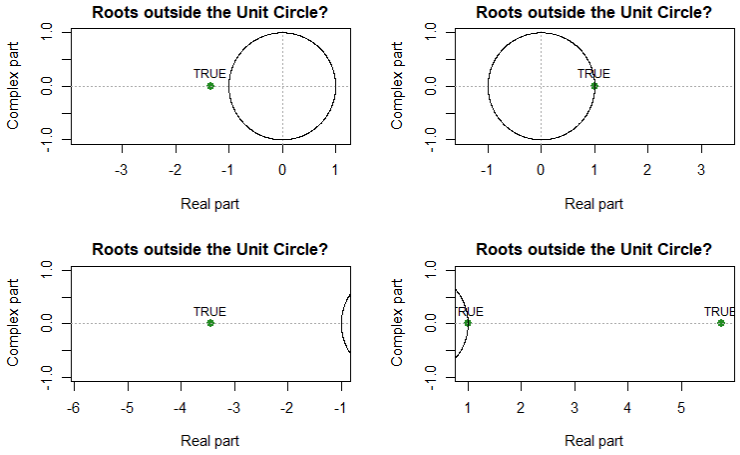


[A5] Unit Circle Checks for Hydro Data Set

Model 1 (p = 1, q = 1, P = 1, Q = 1) Unit Circle Check



Model 2 (p = 1, q = 1, P = 1, Q = 2) Unit Circle Check



Model 3 (p = 1, q = 2, P = 1, Q = 1) Unit Circle Check