

Winter Is Coming: Forecasting Electricity and Natural Gas Prices for Home Heating

Justin DePue, John Rooney, and Tony Jiang

2023-04-28

Contents

Introduction	1
Data	1
Analysis	2
Summary and Conclusions	17

Introduction

This study was born out of both curiosity and necessity. One of our teammates faced higher than expected energy bills this winter and was suddenly faced with the question of whether to shiver to save money or spend it on heating bills and be forced to eat Spaghetti-O's as sustenance. While spring has sprung and summer is on the horizon, we know that winter is coming and that we should prepare now in order that history not repeat itself.

As students recently armed with the tools to conduct time series analysis and forecasting, we realized we could complete our final project while aiding our teammate with knowledge for the future. Our study question thus emerged: would it be more cost-effective to heat an apartment in North Carolina using electricity or a natural-gas powered heater?

Full of optimism that we could complete a class requirement while doing some good for the world (for as Marvel taught us, when you help someone, you help everyone), we set out to find the data that would lead us to the answer we sought. Our journey led us to that great repository of energy knowledge, the US Energy Information Administration. There we found two data sets we felt confident would help us help our teammate: “North Carolina Price of Natural Gas Delivered to Residential Customers (Dollars Per Thousand Cubic Feet)”, which contained monthly data from January 1989 through January 2023, and “Average Retail Price of Electricity by State and Sector” which contained monthly data from January 2001 through January 2023.

Data

Both data sets were downloaded from the EIA as .csv files. The most significant data wrangling was done with the natural gas data as it was originally provided in dollars per thousand cubic feet and needed to be in dollars per kilowatt hour for comparison with electricity data. The electricity data was provided monthly for both regions and states by sector. The natural gas data was provided monthly for North Carolina accompanied by price.

Table 1: Summary Statistics

Variable	Observations	Min	Max	Mean Price	Std. Dev.	Median Price
Natural Gas (\$/Mcf)	410	5.54	30.43	13.78	5.64	12.54
Electricity(cents/kWh)	265	7.53	13.51	10.22	1.30	10.41

Table 2: Sample of Cleaned Natural Gas Data

year	price
Jan-1989	6.17
Feb-1989	6.30
Mar-1989	6.29
Apr-1989	6.80
May-1989	6.99
Jun-1989	8.02
Jul-1989	8.71
Aug-1989	8.97
Sep-1989	8.68
Oct-1989	7.44

Table 3: Sample of Cleaned Electricity Data

my_date	price_per_kWh
Jan.2001	7.53
Feb.2001	7.77
Mar.2001	8.02
Apr.2001	8.00
May.2001	8.22
Jun.2001	8.19
Jul.2001	8.31
Aug.2001	8.35
Sep.2001	8.35
Oct.2001	8.67

Analysis

Our analysis began by creating initial time series objects and plotting them, along with creating ACF and PACF plots to gain an initial sense of what the series looked like and what seasonality they may have.

We then decomposed the time series objects for further analysis to visualize the trend and seasonality of each series. The decomposition of the natural gas time series object shows clear seasonality in price along with a steep upward trend beginning in 2020. The decomposition of electricity prices shows a clear upward trend from the beginning of the series and shows a bimodal seasonality.

Figure 1: NC Residential Natural Gas Cost

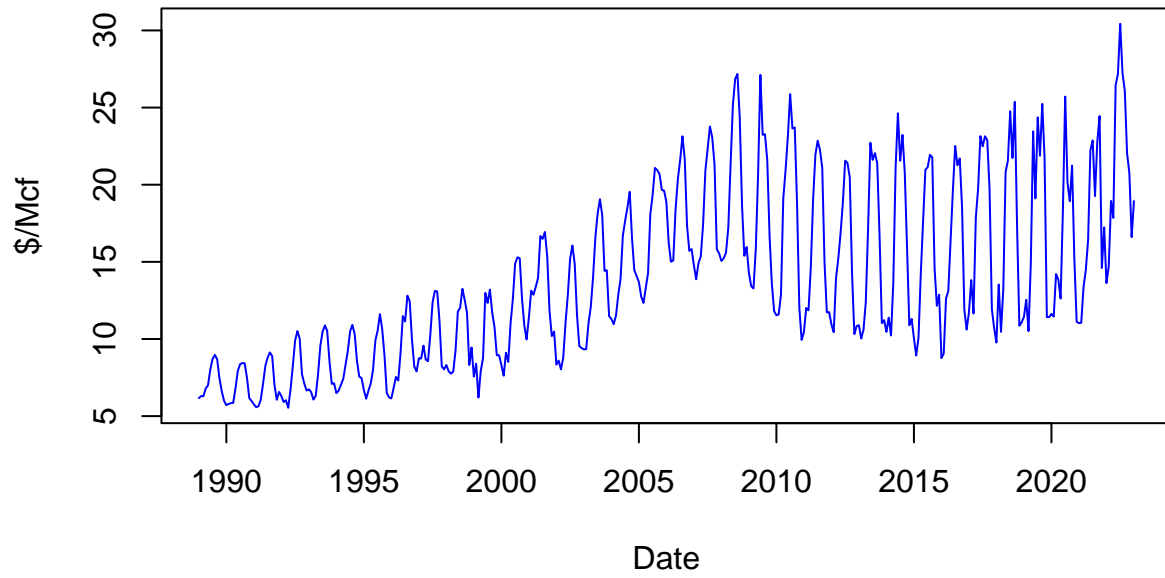


Figure 2: NC Residential Electricity Cost

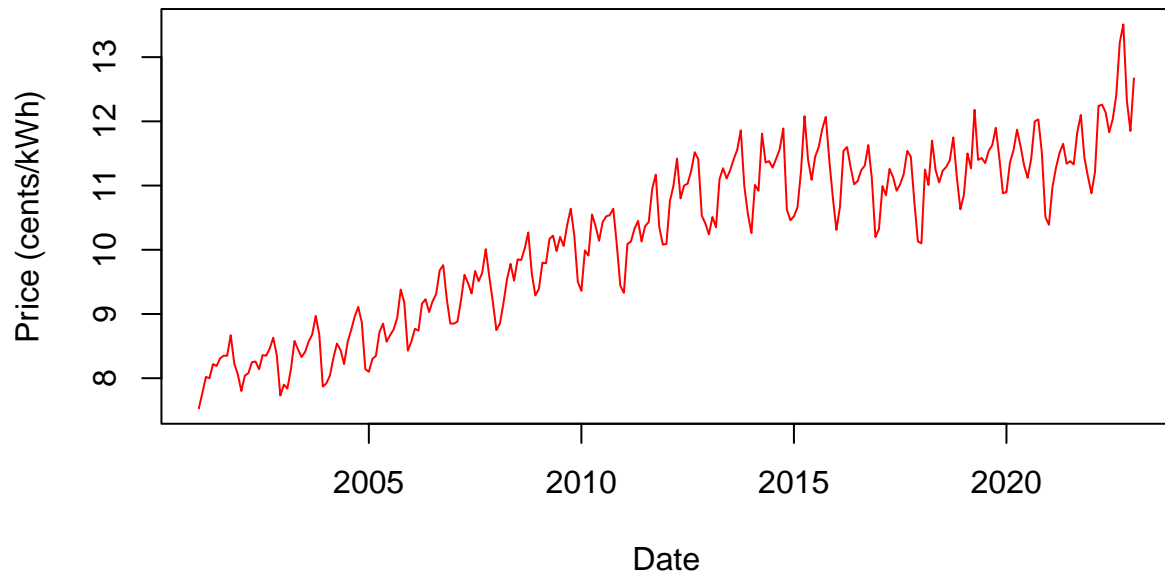


Figure 3: Comparison of Natural Gas and Electricity Costs

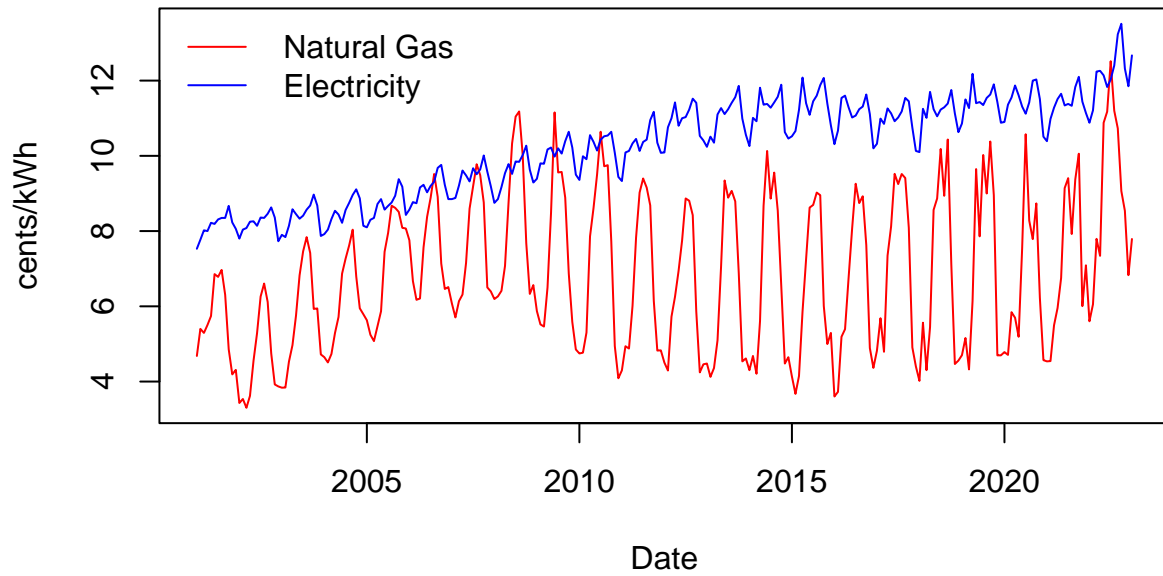


Figure 4: Decomposition of Natural Gas Prices Time Series

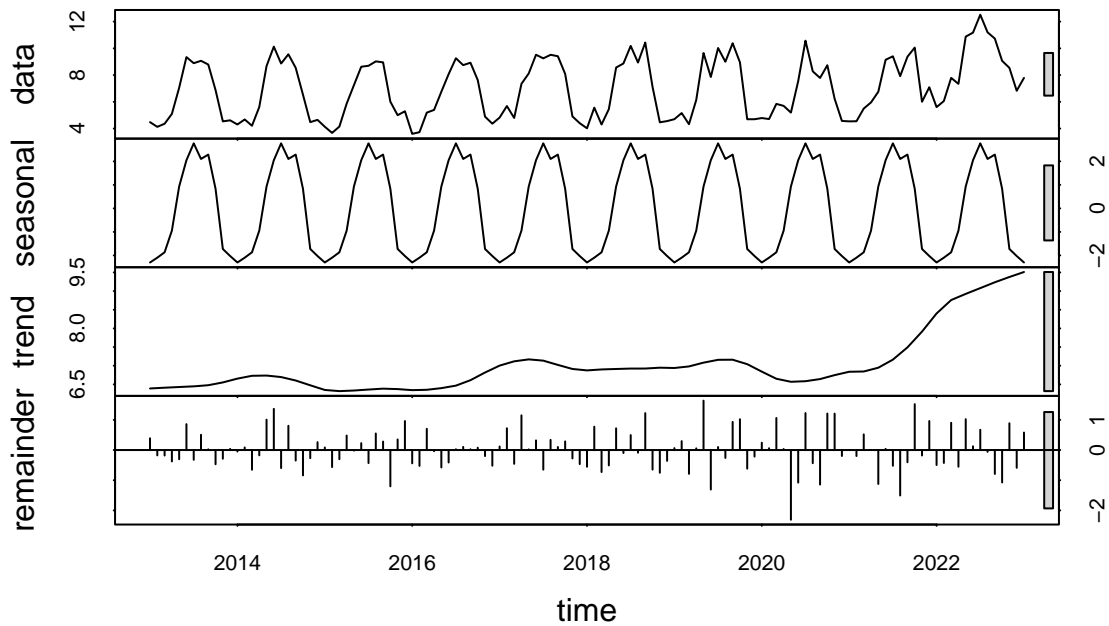


Figure 5: Decomposition of Electricity Prices Time Series

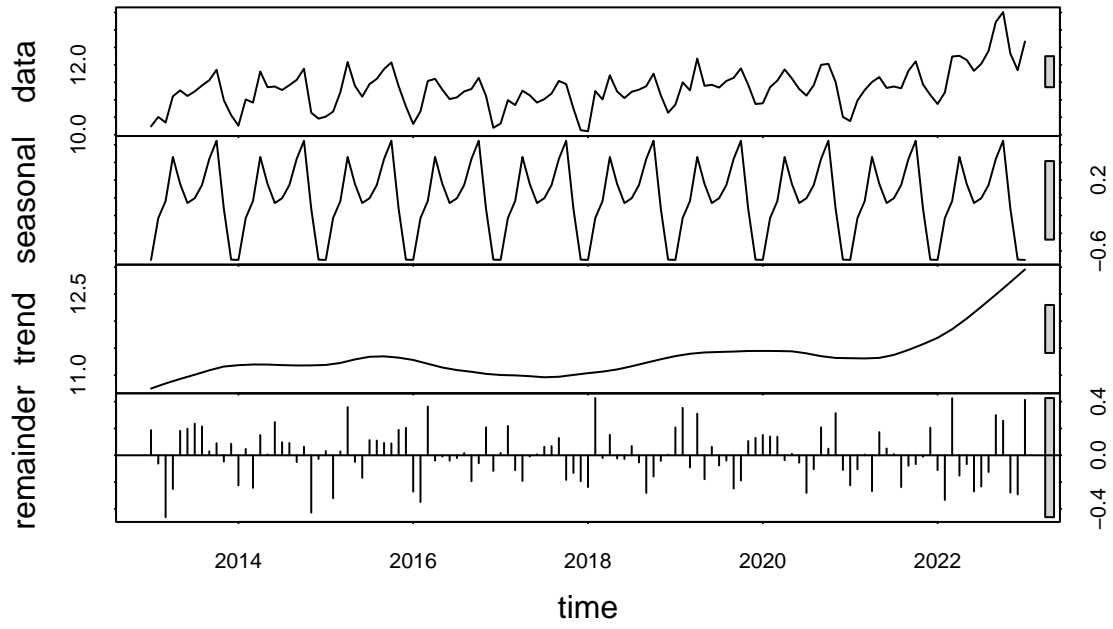


Fig. 6: ACF Electricity

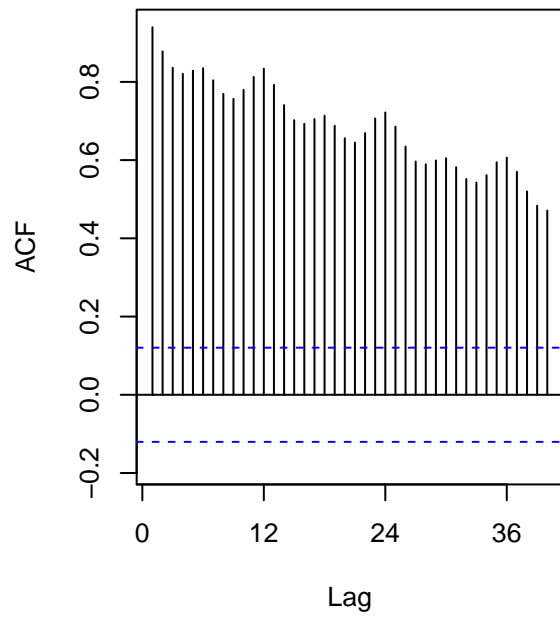


Fig. 7: ACF Non-Seasonal Electricity

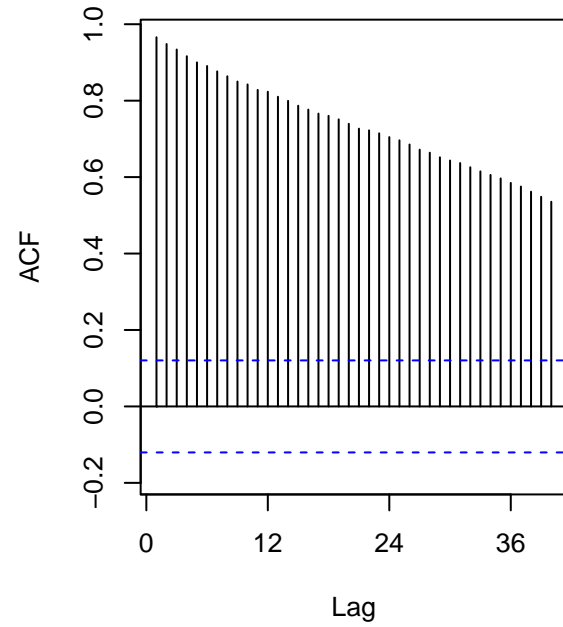


Fig. 8: PACF Electricity

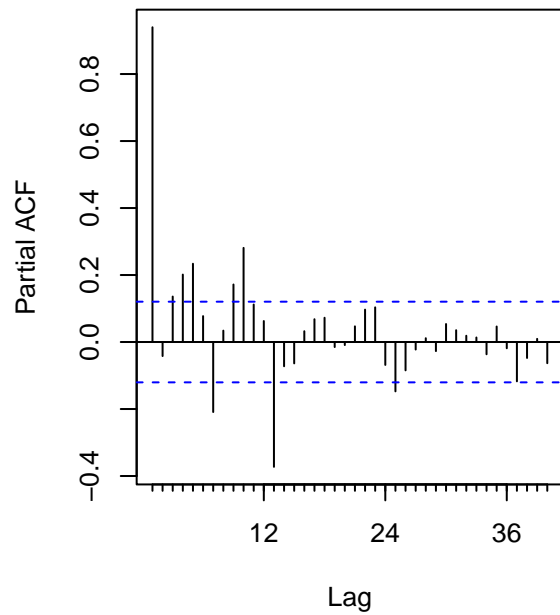


Fig. 9: PACF Non-Seasonal Electricity

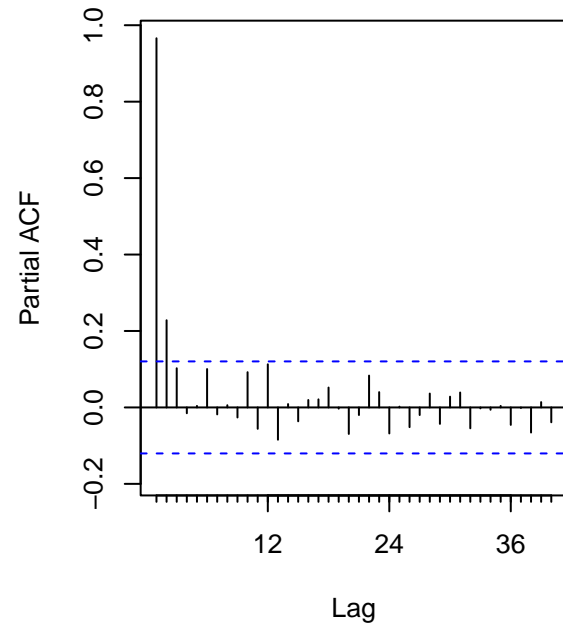


Fig. 10: ACF Natural Gas

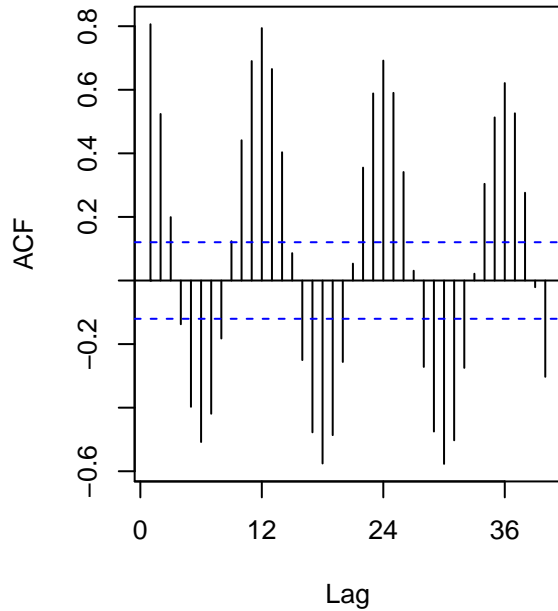


Fig. 11: ACF Non-Seasonal Gas

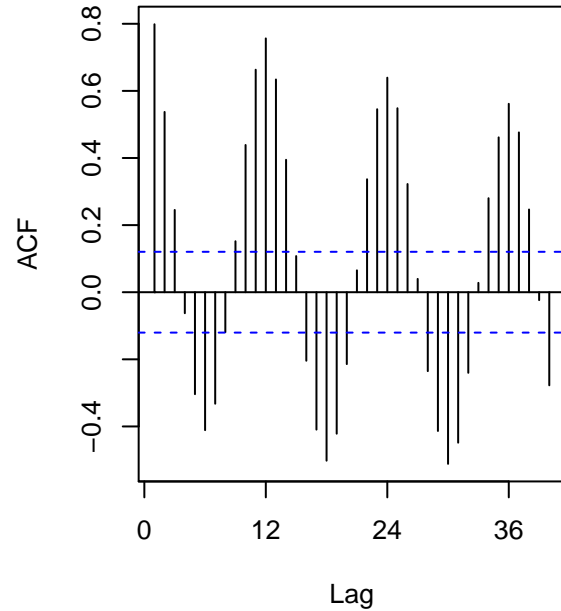


Fig. 12: PACF Natural Gas

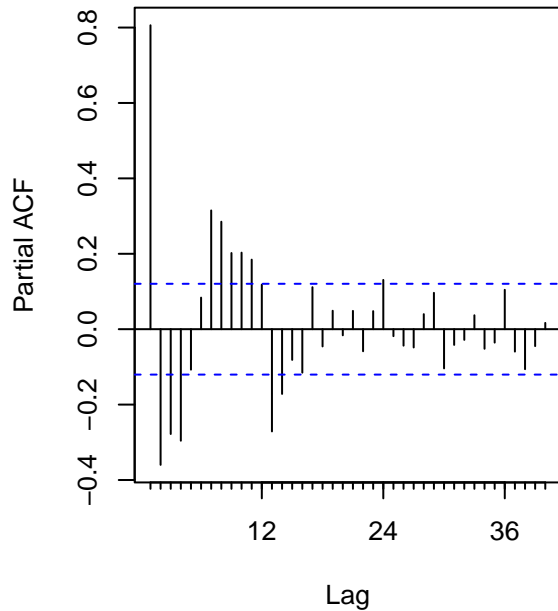
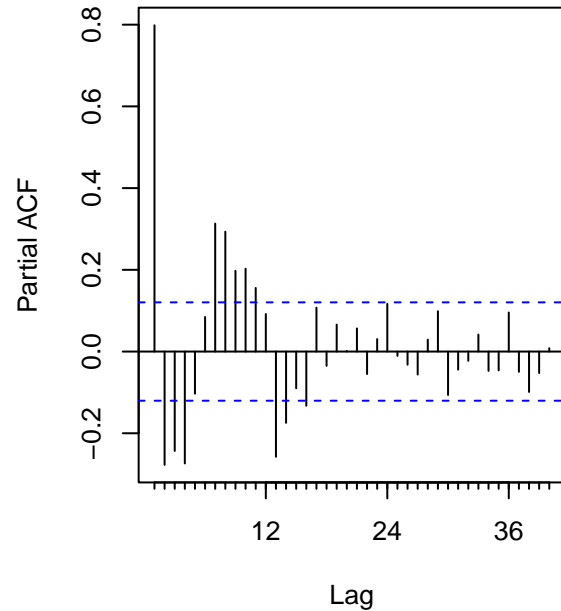


Fig. 13: PACF Non-Seasonal Gas



Several models were developed and tested to determine what would best fit the data we had and what may lead to the best forecast to determine whether heating via electricity or natural gas would be most cost-efficient in the upcoming winter. The five that were used were the Seasonal ARIMA, ARIMA with Fourier terms, Neural Networks, TBATS, and STL + EST.

Each of these tests used functions from the “forecast” library and required the use of time series data which

we created using the “ts” function from the “tseries” library.

While the Seasonal ARIMA was a logical first model to try, we quickly felt that the performance was not particularly strong and that developing and testing other models was warranted.

Figure 14: Electricity – SARIMA

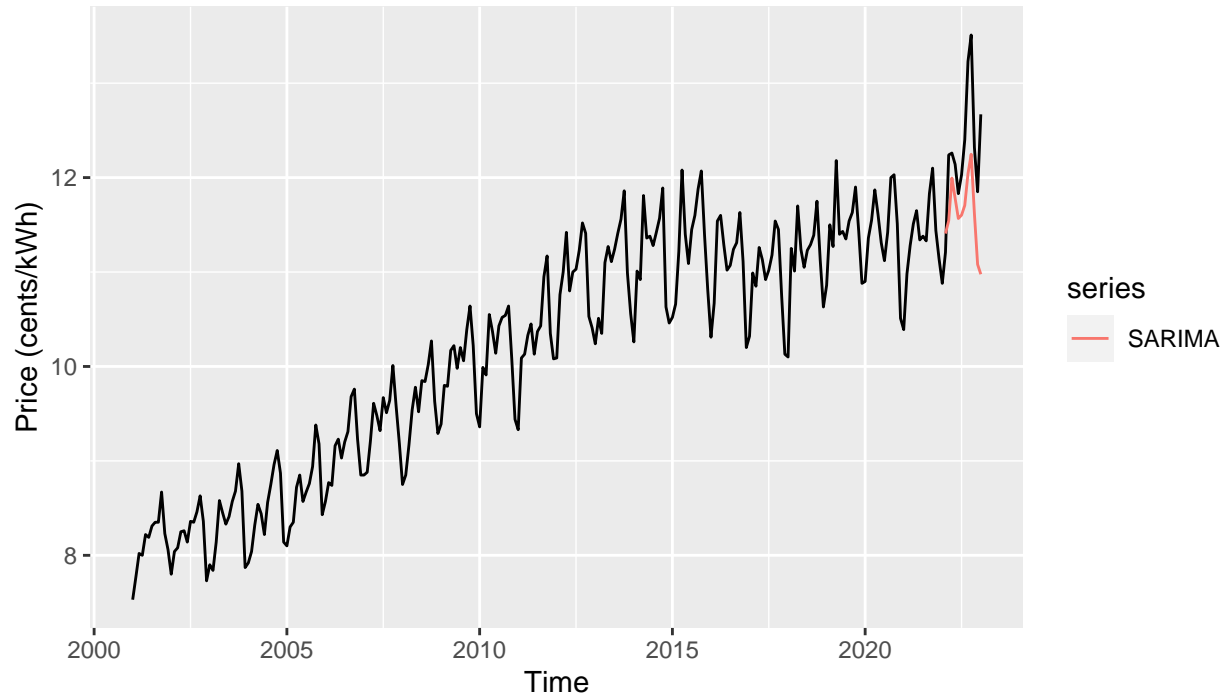
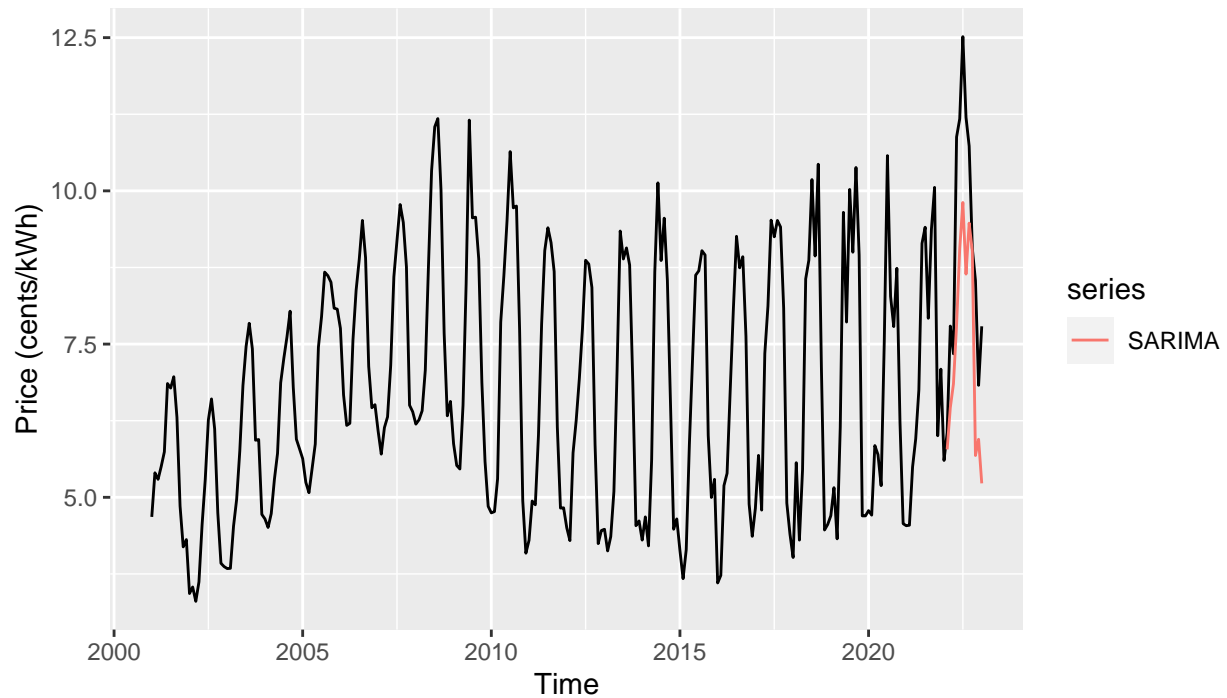


Figure 15: Natural Gas – SARIMA



We decided to explore more advanced models such as the ARIMA with Fourier Terms model, Neural Net-

works, TBATS, and STL + EST. ARIMA with Fourier terms is known as a dynamic harmonic regression model with an ARMA error structure, using the “fourier” function from package “forecast” to find terms that model seasonal components.

Figure 16: Electricity – ARIMA with Fourier Terms

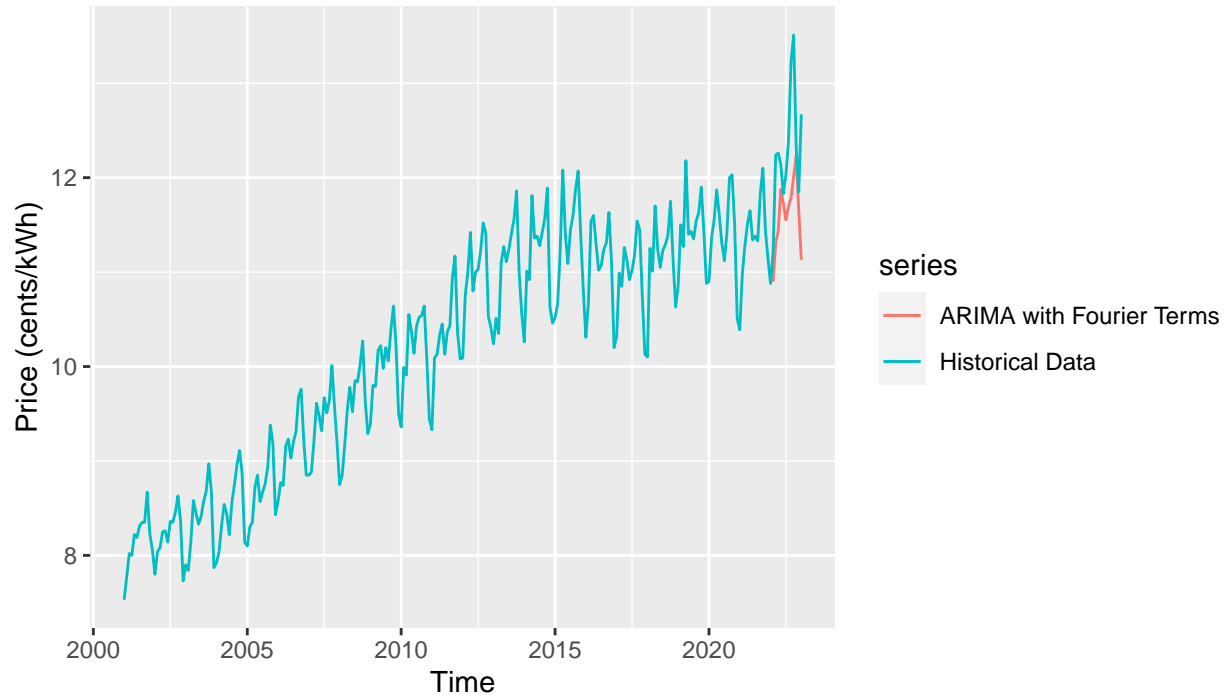
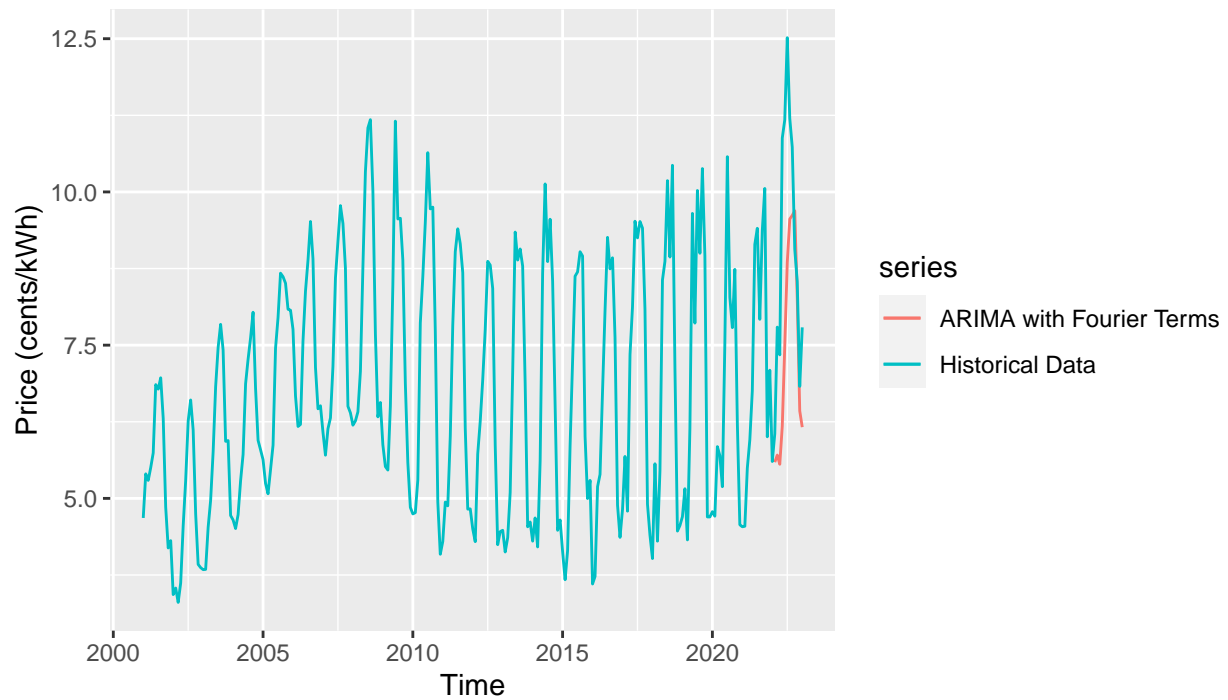


Figure 17: Natural Gas – ARIMA with Fourier Terms



We next developed and tested an STL model.

Figure 18: Electricity – STL + EST

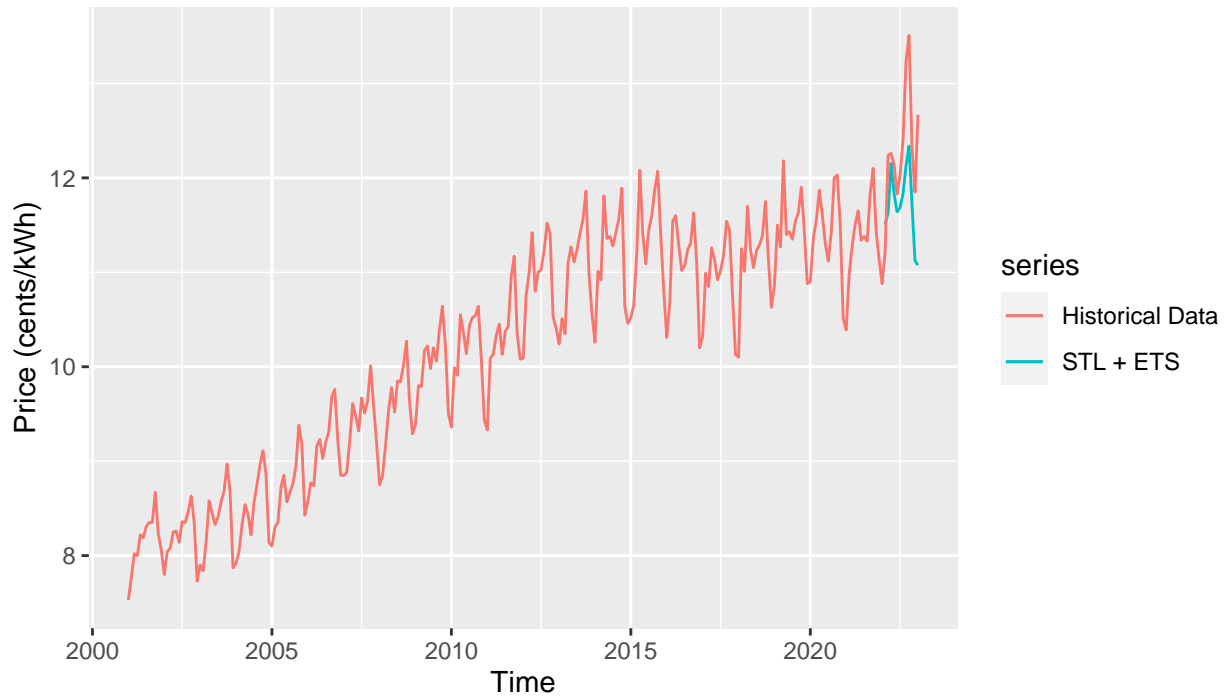
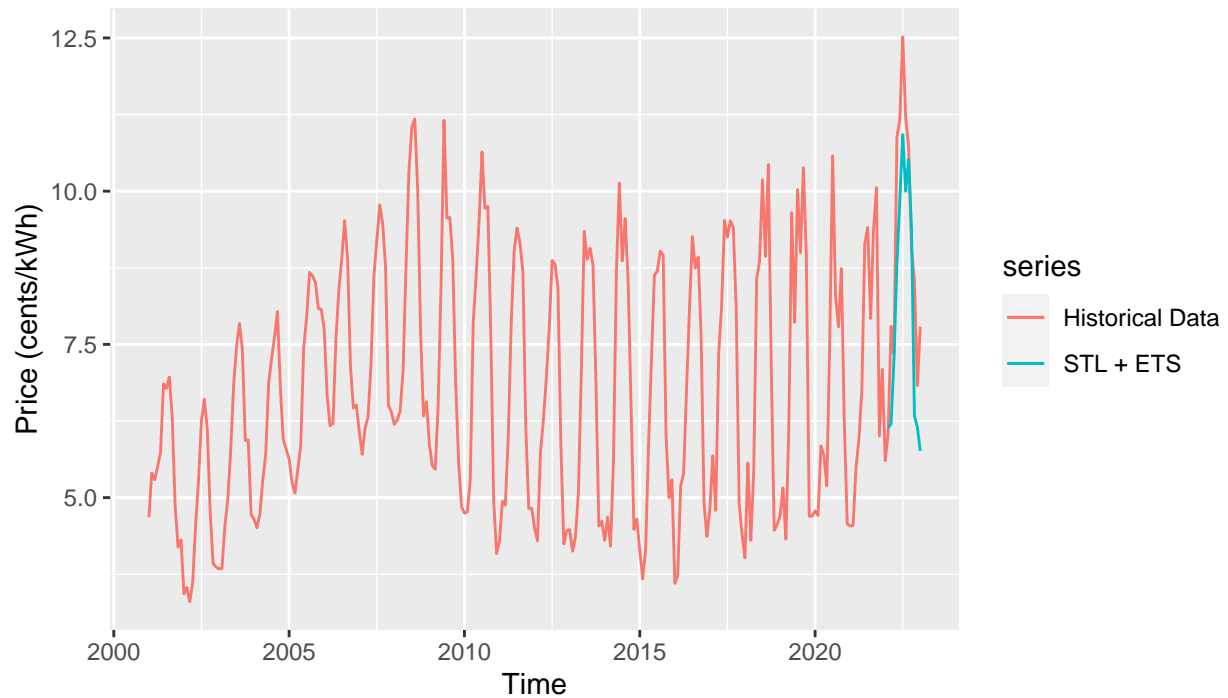


Figure 19: Natural Gas – STL + EST



We then developed and test a neural network model using the “`nnetar()`” function in the “`forecast`” package. We learned that the `p` and `P` arguments in `nnetar()` have significant impact on model performance. Therefore,

we worked to identify the optimal p and P combination through trial and error.

Table 4: Natural Gas Neural Net Forecast Accuracy using Various Seasonal Lag Inputs (p/P)

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
1/0	-1.17117	1.29507	1.17117	-10.50130	10.50130	0.28456	45.32660
1/1	-0.87546	0.98759	0.87546	-7.63058	7.63058	0.18585	4.70644
2/0	-1.03808	1.16810	1.03808	-9.19265	9.19265	0.29494	18.76692
2/1	-0.88192	0.99757	0.88192	-7.69509	7.69509	0.21209	4.72340
2/2	-0.87153	0.99922	0.87153	-7.60964	7.60964	0.23145	4.50574
1/2	-0.87155	0.99194	0.87155	-7.60113	7.60113	0.19811	4.59363
3/1	-0.89268	1.01001	0.89268	-7.79862	7.79862	0.22038	4.78321

Table 5: Natural Gas Neural Net Forecast Accuracy using Various Seasonal Lag Inputs (p/P)

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
1/0	-2.29026	3.00287	2.39525	-33.52695	34.93479	0.62861	18.47394
1/1	-1.78076	2.23338	1.96322	-25.37426	27.56868	0.28182	2.43279
2/0	-2.52278	3.28516	2.68509	-40.40244	42.48286	0.67312	11.98456
2/1	-1.89127	2.21942	1.96239	-27.37555	28.12476	0.24643	2.53647
2/2	-2.01413	2.34855	2.05988	-29.90887	30.39867	0.27090	2.68062
1/2	-1.91450	2.32140	1.98370	-27.79047	28.52038	0.27551	2.59953
3/1	-1.93431	2.26005	1.98593	-28.53177	29.08244	0.26445	2.60379
3/2	-1.89316	2.22117	1.94690	-27.73933	28.31184	0.23368	2.53577
33	-2.02489	2.33195	2.10907	-30.37214	31.25175	0.19048	2.45707

We found that combination 11 ($p = 1$ and $P = 1$) had the best modeling performance for our electricity series, and combination 32 ($p = 3$ and $P = 2$) had the best modeling performance for our natural gas series. We then used combination 11 to run a neural network model with fourier terms for our electricity series.

Figure 20: Electricity – Neural Network

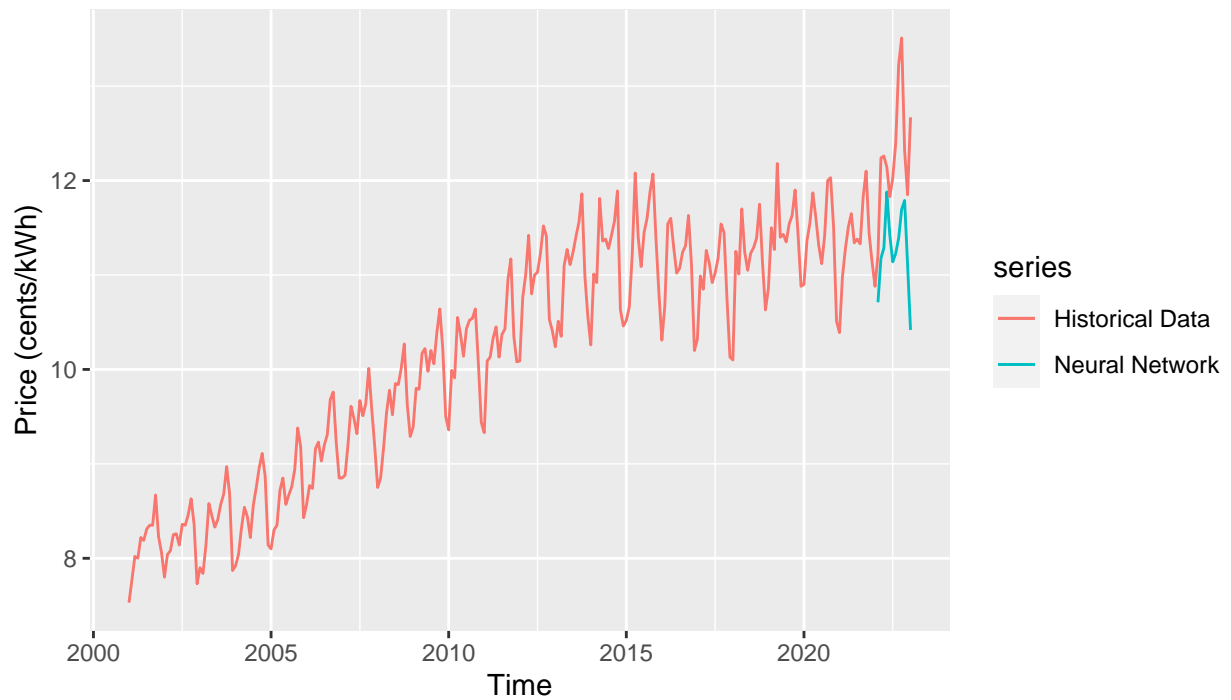
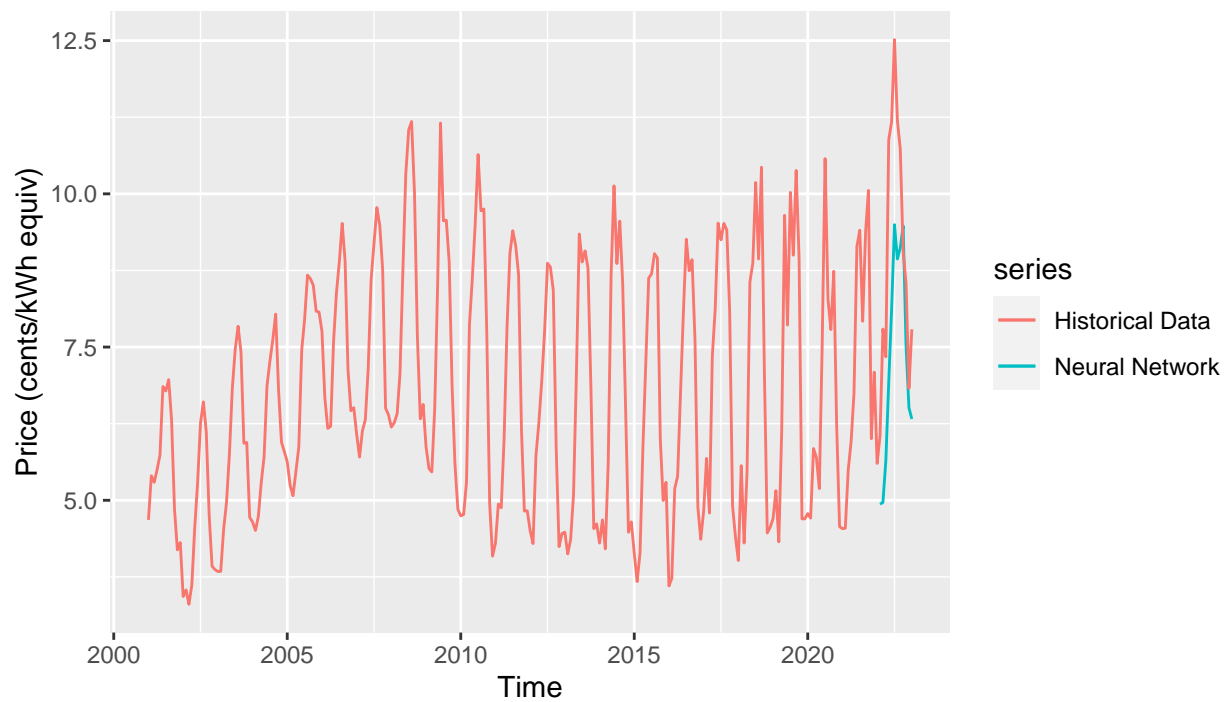


Figure 21: Natural Gas – Neural Network

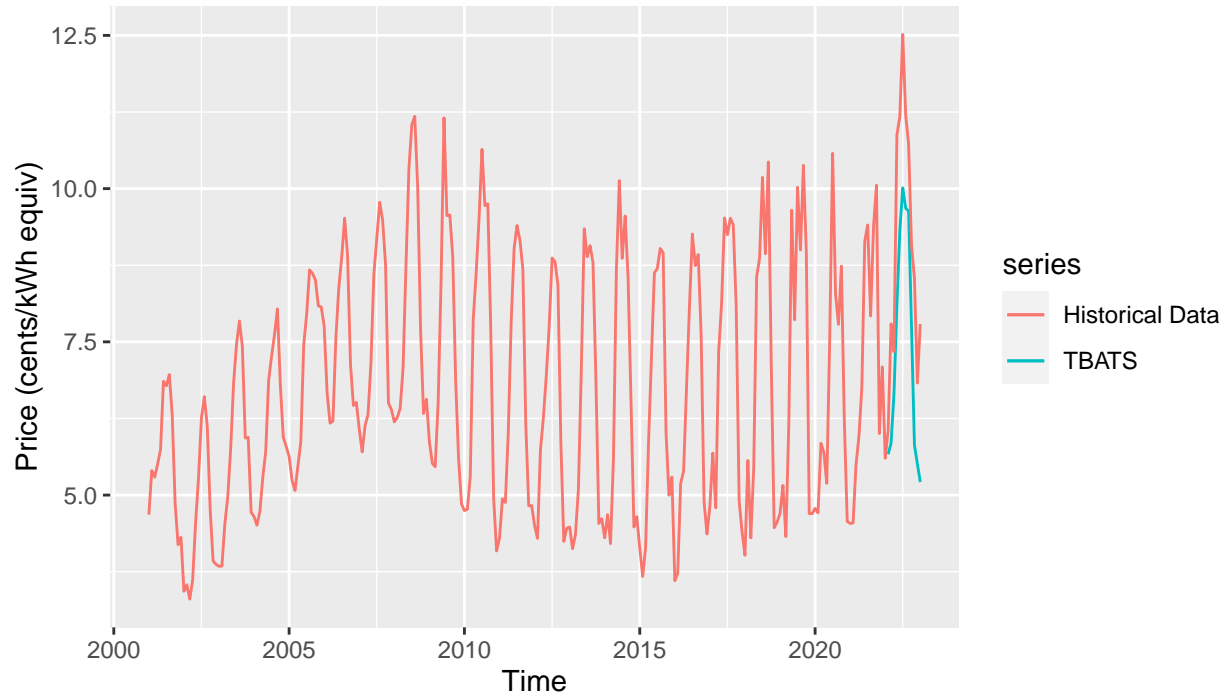


We then developed and tested a TBATS models for our electricity and natural gas series.

Figure 22: Electricity – TBATS



Figure 23: Natural Gas – TBATS



As we thought about other ways to model our time series, we realized that the Ukraine War has had a

significant impact on natural gas prices, and likely electricity price, too. We additionally realized temperature would be a good regressor to include since utility bills normally fluctuate in the same direction as temperature. Therefore, we created two covariates: *UKRWAR* and temperature. *UKRWAR* is an indicator variable with values of 0 and 1. Months before March 2022 have a value of 0, while months after and including March 2022 have a value of 1. The reason why we set the cutoff month at March 2022 despite the war starting in February of that year is because the impact of the war on monthly natural gas prices in February 2022 should be limited because of the war beginning that month. The temperature series is the monthly average temperature of the Raleigh area. This is largest geographic level of historical temperature data.

After creating all the covariates, we repeated our modeling but with covariates to improve the accuracy of the models. First we incorporated covariates into our neural network model.

Figure 24: Electricity – Neural Network

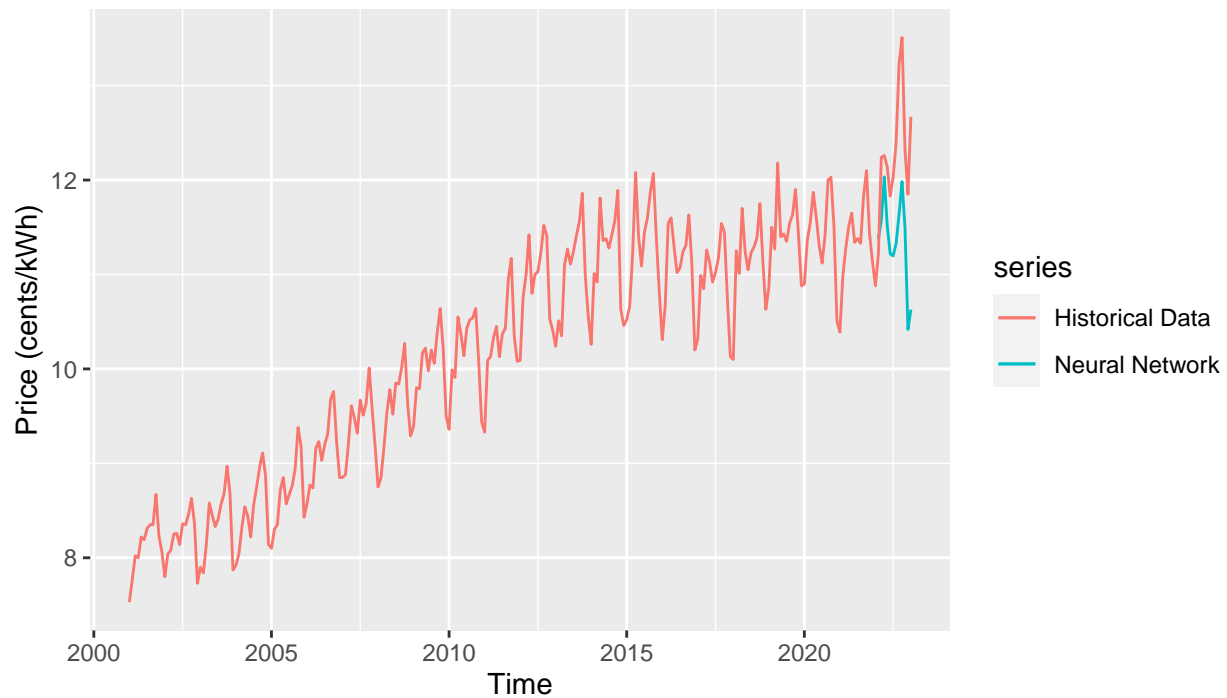
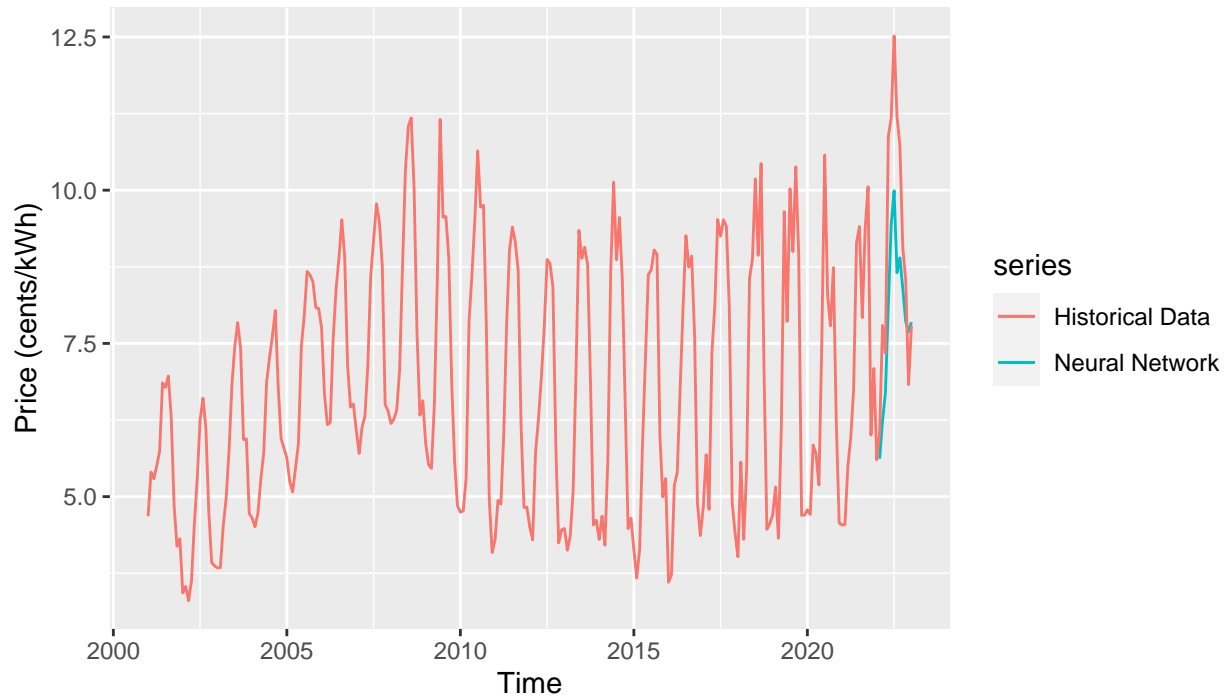


Figure 25: Natural Gas – Neural Network

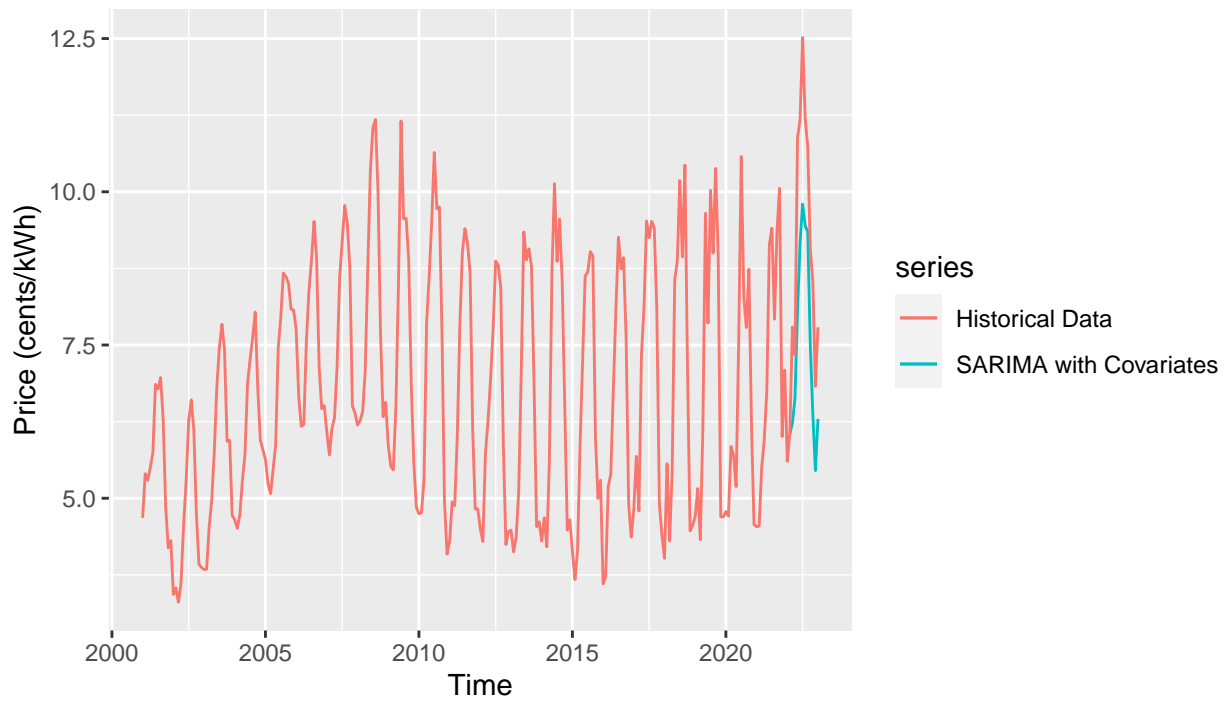


We then developed a seasonal arima model with temperature and fourier terms to model two series. UKR-WAR as a covariate was excluded because R reported no suitable ARIMA model when UKRWAR was included. The function used for this was “auto.arima(xreg)”.

Figure 26: Electricity – ARIMA with Covariates



Figure 27: Natural Gas – SARIMA with Covariates



Summary and Conclusions

When comparing the accuracy of each of the models that were tested, the STL model emerged as the best fit for both the price of electricity and the price of natural gas when comparing both the RMSE and MAPE scores. Both models missed the heights seen in 2022, however. There could be several reasons for this. One is the difficulty of modeling for the beginning of the war in Ukraine and the impact that has had on natural gas prices both regionally and around the world. Another is the difficulty of modeling for the influence of generationally high inflation. With both of these compounding each other, it is perhaps not surprising that all of the models failed to predict the rise in price for both electricity and natural gas.

Figure 28: Comparison of Modeling Methods – Electricity

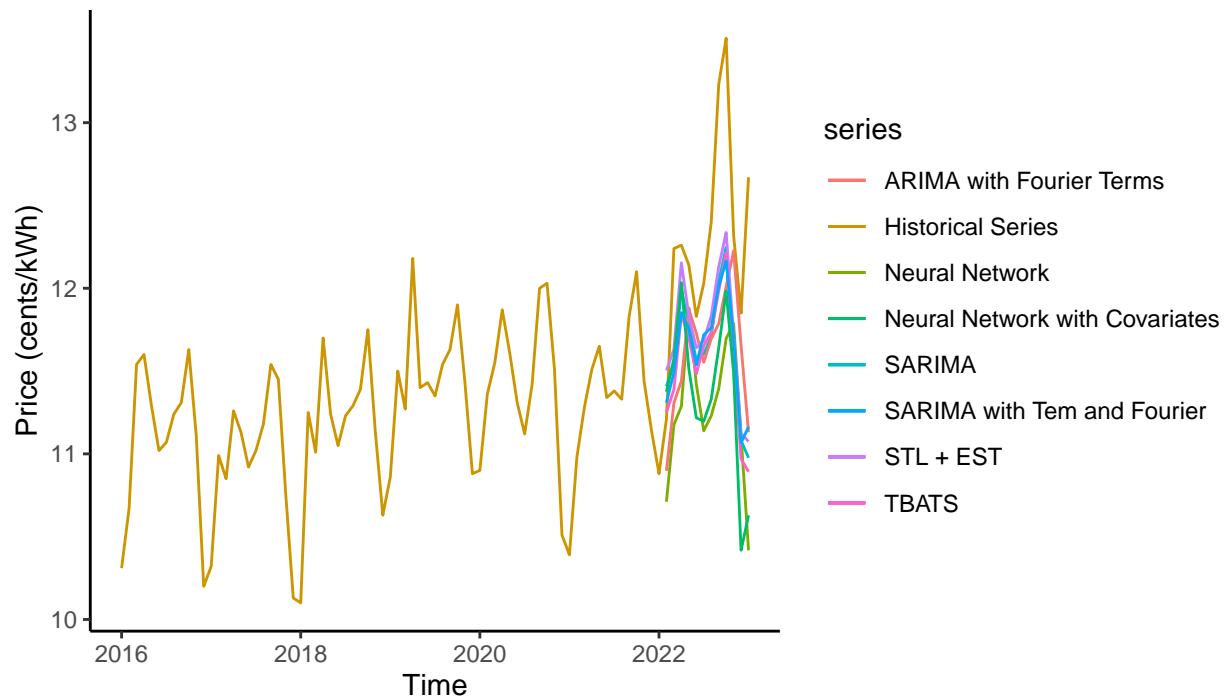
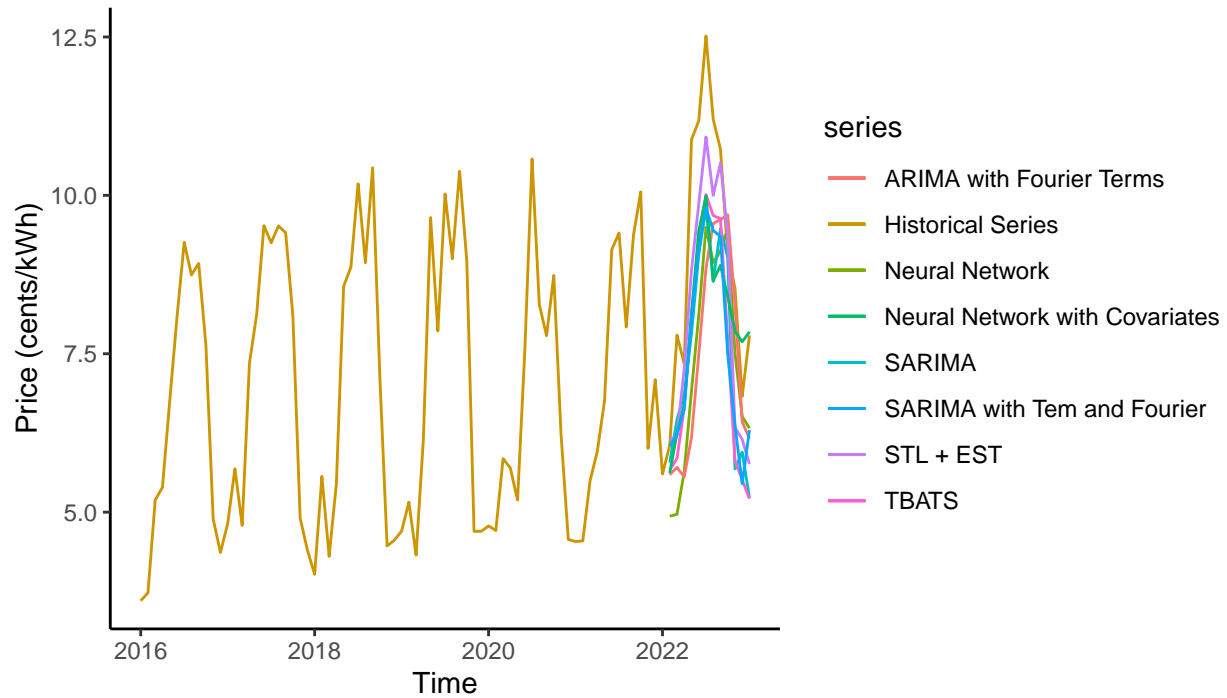


Figure 29: Comparison of Modeling Methods – Natural Gas



The best model for electricity by RMSE is: STL

Table 6: Forecast Accuracy for NC Residential Electricity Price

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
SARIMA	-0.67702	0.83764	0.70998	-5.83726	6.12612	0.25988	2.73983
ARIMA with Fourier	-0.69901	0.87542	0.69901	-6.04750	6.04750	0.08222	2.82856
STL	-0.58551	0.76708	0.63447	-5.01390	5.43949	0.25835	2.36056
Neural Network	-1.03677	1.20053	1.03677	-9.28452	9.28452	0.16347	2.99540
TBATS	-0.74971	0.89063	0.75658	-6.50522	6.56626	0.23462	2.86546
Neural Network with Covariates	-0.94054	1.11384	0.96798	-8.37641	8.61765	0.40584	2.65317
SARIMA with Tem and Fourier	-0.67930	0.81795	0.69585	-5.83257	5.97890	0.26220	2.92127

The best model for natural gas by RMSE is: STL

Our initial research question posed how our roommate could best plan for surviving, if not thriving, in the upcoming winter. Based on our analysis, the answer appears to be that he would be best off by buying a natural gas-powered heater. Given the vicissitudes of energy prices seen in recent months, and the uncertainty of global and domestic events that could further impact energy prices, this is hard to say with certainty as is the case with all forecasts. Revisiting these forecasts both ahead of the upcoming winter and annually thereafter with updated data would be well-advised.

Table 7: Forecast Accuracy for NC Residential Natural Gas Price

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
SARIMA	-1.67902	1.97784	1.67902	-23.35174	23.35174	-0.15535	1.81675
ARIMA with Fourier	-1.73517	2.31199	1.83980	-25.17030	26.24979	0.57850	3.12540
STL	-1.04788	1.35599	1.11314	-14.03428	14.80904	-0.33348	1.28849
Neural Network	-1.83150	2.19111	1.90008	-26.43336	27.15709	0.43102	2.70500
TBATS	-1.72858	1.89245	1.72858	-24.53539	24.53539	-0.32231	2.08610
Neural Network with Covariates	-1.20513	1.62012	1.35797	-14.54817	16.53368	0.45962	2.06239
SARIMA with Tem and Fourier	-1.63275	1.79495	1.63275	-21.52013	21.52013	-0.01071	1.95239

Figure 30: 12 Month STL Forecast of Electricity and Natural Gas Price

