Abstract— This article introduces the forecasting of natural gas consumption by using forecasting models such as ARIMA, ETS, TBATS, NN and PROPHET. R-Studio, which uses pre-processed data after outlier analysis and unit root control, is used in the study. After achieving necessary actions, forecast models are established and their performance on both training and test data is compared according to various evaluation criteria.

Time Series Analysis of Natural Gas Consumption

Bilgehan Aydoğdu

2428886

I.INTRODUCTION

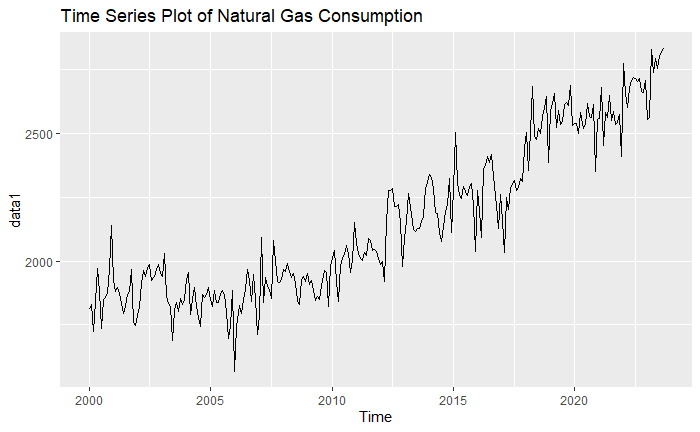
The Natural Gas Consumption gas is a reflection of how much natural gas is being used by both households and industries during a specific period of time. This consumption data is crucial for understanding energy trends and economic activities. For this purpose, natural gas consumption data from January 2000 to September 2023 is analyzed and explored.

The aim of this study is to understand the behavior of the Natural Gas Consumption values by considering the past trends effects during given time frame. This study provides an in-depth look at time series data on natural gas use, improving our understanding of consumption patterns and providing valuable information on long-term trends and dynamics in the use of natural gas resources.

In addition to the exploration of the data, the Natural Gas Consumption values are forecasted by using forecast models such as ARIMA, ETS, TBATS, Neural Network and Prophet models. Then, forecast performances were compared with each other to decide which forecast method was better.

II.DATA DESCRITPION & PREPROCESSING

The data set is taken from website of The Federal Reserve Bank of St. Louis, which is <https://fred.stlouisfed.org/>. The data set contains 285 observations that includes the natural gas consumption values from January 2000 to September 2023.



***Graph 1:*** Time Series Plot of Data Set

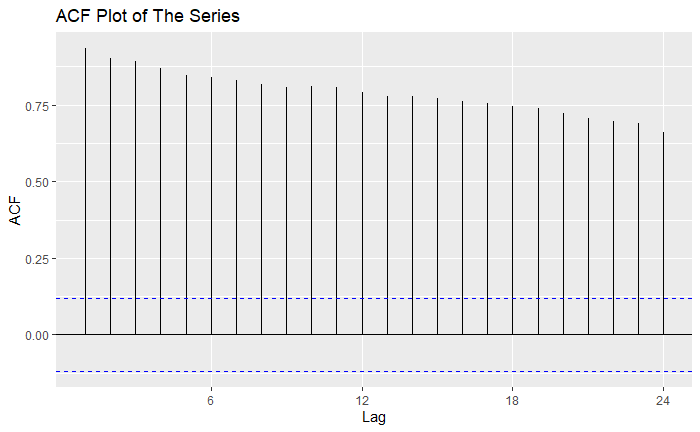
By looking at the time series plot given in Graph 1, it can be said that the series has increasing trend. So, it means the series is not stationary in mean. However, the line does not seem to a straight line. So, it might be said that the series may have stochastic trend, but we need to check.

At the beginning of the analysis, the data set is divided into two parts which are train set and test set. Test set is selected as las one year, which is the last 12 observations were selected for test set.

A graph with numbers and dots

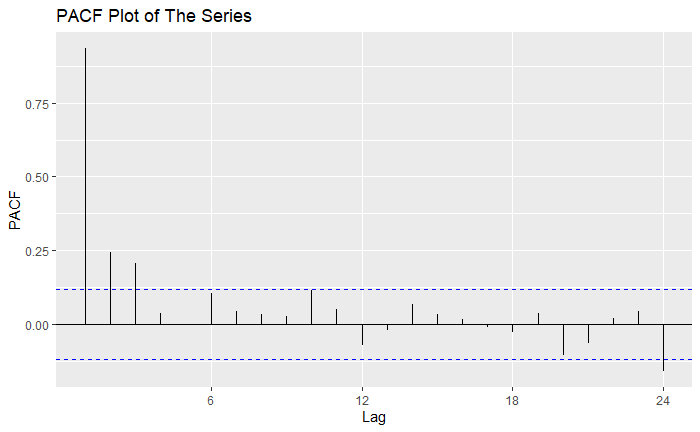
Description automatically generated***Graph 2:*** *Anomaly Detection Plot*

Then, the anomalies in the data set were checked and as it might be seen in Graph 2, the series has no anomalies.



***Graph 3:*** *ACF Plot of The Series*

When the ACF plot of the series is examined, given in Graph 3, It might be seen that there is a slow linear decay. This output validates the assessment stated in Graph 1. As a result, it might be said that the process is non-stationary.



***Graph 4:*** *PACF Plot of The Series*

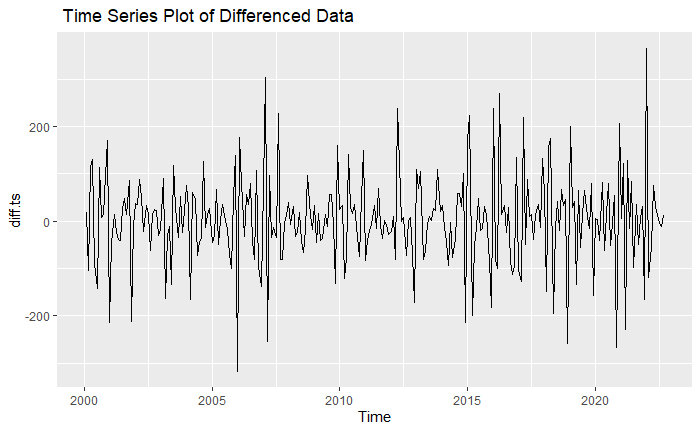
When the PACF plot of the series is examined, it cuts off after lag 1. So, the plots shows the series is not stationary, which we have already mentioned that the process is non-stationary at the previous graph. So, our aim is to remove this problem.

Before taking any step, it might be reasonable to check that series need any transformation. Lambda value is generated for our dataset. However, since the lambda value equals to 1 it might be said that the series does not need any transformation.

After checking this step, to make the dataset stable and get a stationary mean, tests are used. Non-stationary and trend level of the series is checked by KPSS, ADF, HEGY and Canova-Hensen tests. According to the KPSS level test, the series is not stationary (p <0.05). The result of ADF test is similar to the KPSS test and shows that the series is not stationary (p>0.05). KPSS and ADF test gave mixed results for trend of the series. KPSS shows that the series has stochastic trend, but ADF suggest deterministic trend. However, when time series plot of the series is examined, it is obvious that the series has a stochastic trend. So, we may ignore this part of the result of the ADF test.

HEGY and Canova-Hensen tests are used to check if the series has regular and seasonal unit roots. As a result of these tests, both tests show that the series has a regular unit root, but it does not have seasonal unit root. By looking at these test, the next step for our analysis taking difference to make the series stationary.

At this step, first order regular differencing is taken for the series. ADF and HEGY tests are conducted to check if the regular differencing solved the stationary problem or not. According to the results of ADF and HEGY tests, the series became stationary after taking first order regular differencing. The differenced series does not have regular or seasonal unit roots.

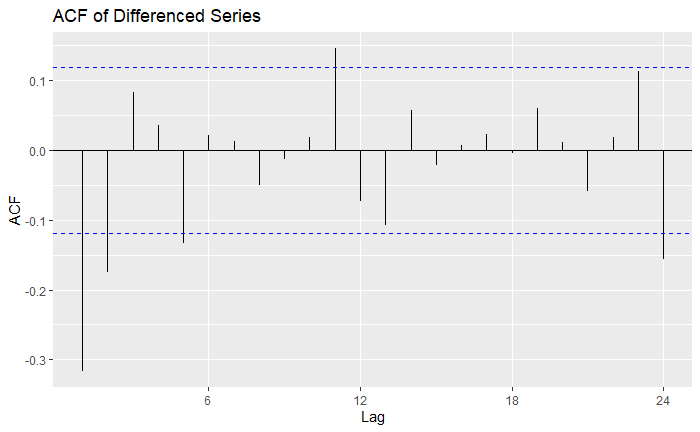


***Graph 5:*** *Time Series Plot of Differenced Data*

The trend is removed, as shown in Graph 5. The following step is suggesting a model for the series.

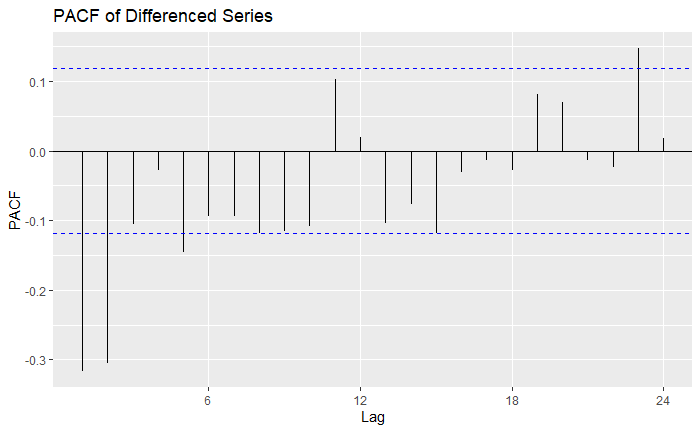
III. MODEL SUGGESTION

Stationary of the series satisfied, we will suggest a model using the ACF and PACF plots of the series. Moreover, we also consider several other techniques for the series.



***Graph 6:*** *ACF Plot of Differenced Series*

By looking at ACF plot of differenced series, given in Graph 6, the process cuts off after lag 2 but it might be said that lag 1 as well. Therefore, we may suggest MA(2) or MA(1) orders for the process.



***Graph 7:*** *PACF Plot of Differenced Series*

When the PACF plot is examined, shown at Graph 7, it is clearly seen that the process cuts off after lag 2. So, AR(2) process might be suggested for the process. Therefore, according to the both ACF and PACF plots, ARIMA(2,1,1) and ARIMA(2,1,2) models may be suggested for the series.

Furthermore, The Extended Sample Autocorrelation Function (ESACF) method is used to determine the model of the series, but it does not give healthy results. Then, auto.arima function is used and the function suggests ARIMA(1,1,1)(2,0,0)[12] and ARIMA(1,1,2)(2,0,0)[12] models.

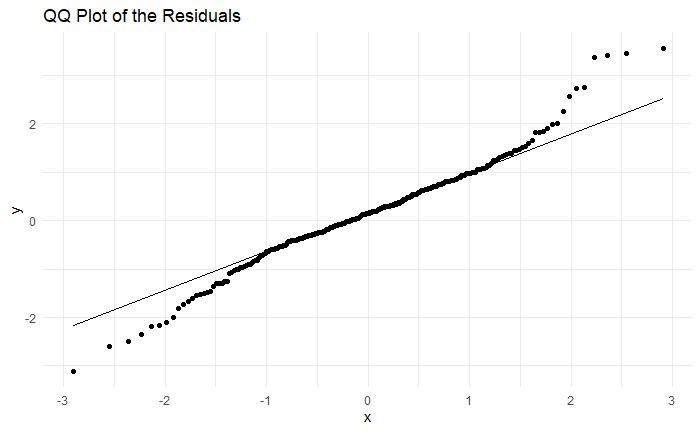
We have four models that may fit the series. To decide which one is best, significance and performances of the models are compared. After checking the significance of the models, ARIMA(2,1,1), ARIMA(2,1,2) and ARIMA(1,1,2)(2,0,0)[12] models are satisfied the significance. AIC values will be compared to choosing the best one.

***Table 1:*** *Summary of Model*

|  |
| --- |
| 1. Series: ts\_train |
| 1. ARIMA(1,1,1)(2,0,0)[12] |
|  |
| 1. Coefficients: |
| 1. ar1 ma1 sar1 sar2 |
| 1. 0.2602 -0.7653 -0.0713 -0.2170 |
| 1. s.e. 0.0925 0.0616 0.0629 0.0644 |
|  |
| 1. sigma^2 = 6996: log likelihood = -1588.8 |
| 1. AIC=3187.61 AICc=3187.83 BIC=3205.64 |

When the AIC values are compared, the model ARIMA(1,1,2)(2,0,0)[12] has the lowest AIC values among them, given in Table 1. We will proceed with the model's diagnostic checks since we have selected the best model.

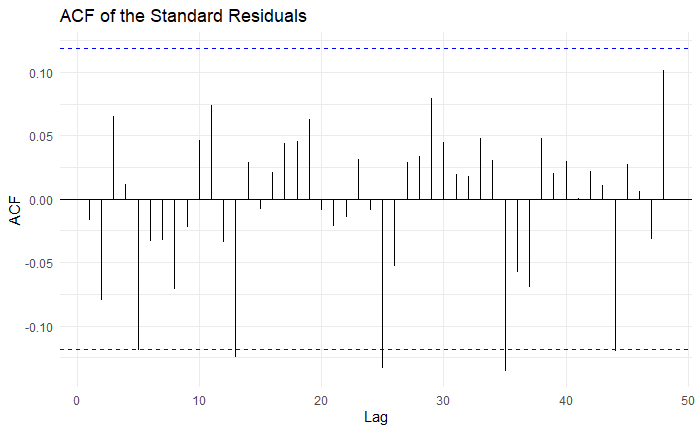
This stage should involve verifying the model's quality of fit and the accuracy of the underlying assumptions. Firstly, Q-Q plot is examined and using formal tests like the Shapiro-Wilk and Jarque-Bera allow us to verify the normality assumption.



***Graph 7:*** *Normal-QQ plot of the residuals*

QQ-Plot of the residuals, shown in Graph 7, shows S shape and we may conclude that the residuals do not satisfy the normality. To be sure, Shapiro-Wilk and Jarque-Bera tests are conducted. Unfortunately, both tests do not satisfy the normality. To remove this problem, other significant models are tested but none of them satisfied the normality. So, we move on our analysis assuming that the residuals distributed normally.

The residuals in a perfect forecasting model should be uncorrelated. To check that both visual and formal tests are performed.



***Graph 8:*** *ACF plot of the standard residuals*

According to the ACF plot of standard residuals, given in Graph 8, almost all spikes are in the WN band, which satisfies our assumption. To be sure Breusch-Godfrey test is conducted to check possible autocorrelation in residual series. The result of this test shows that the residuals of the model are uncorrelated. Moreover, Portmanteau lack of fit test is performed, and the test shows that the residuals of the model are uncorrelated.

To satisfy the last assumption, which is heteroscedasticity, both visual and formal tests are performed.

A graph of a line graph

Description automatically generated with medium confidence

***Graph 9:*** *ACF and PACF of the squared* *residuals*

For the visual representation, as it seen at Graph 9, ACF and PACF plots shows that almost all spikes are in of the white noise bands. This is evidence of homoscedasticity. To be sure about this judgement, Engle’s ARCH test is performed.

As a result of that, the test supports the visual way since p-value is greater than 0.05. Thus, we do not have heteroscedasticity problem and the ARCH impacts are not present.

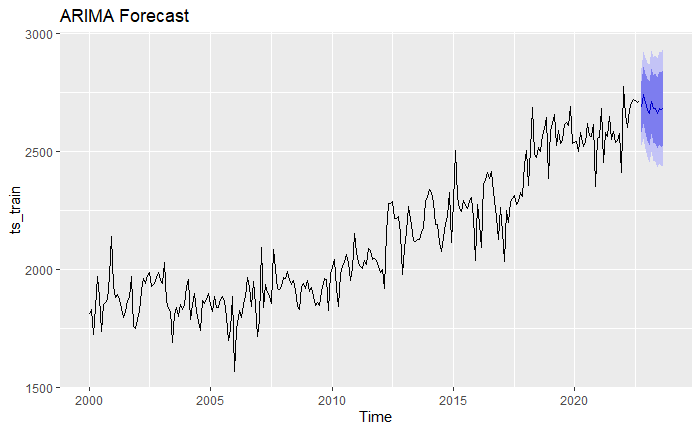
V. FORECASTING

For the first step of the forecasting, for the Minimum MSE forecasting, ARIMA was performed. Secondly, the “ets” function is used to find the optimal exponential smoothing model. The output below, Table 2, shows the optimal exponential smoothing model for the series.

***Table 2:*** *Summary of ETS Model*

|  |
| --- |
| ETS(A,N,N) |
|  |
| Call: |
| ets(y = ts\_train, model = "ZZZ") |
|  |
| Smoothing parameters: |
| alpha = 0.4303 |
|  |
| Initial states: |
| l = 1812.6721 |
|  |
| sigma: 86.7216 |
|  |
| AIC AICc BIC |
| 3972.014 3972.104 3982.843 |

By looking at Table 2, it might be seen that we have Simple Exponential Smoothing with an additive error. Then, the Shapiro-Wilk test is conducted to check the normality. As a result, the residuals of the exponential smoothing model do not follow a normal distribution (p<0.05). The visual representation of ARIMA forecast is represented below.



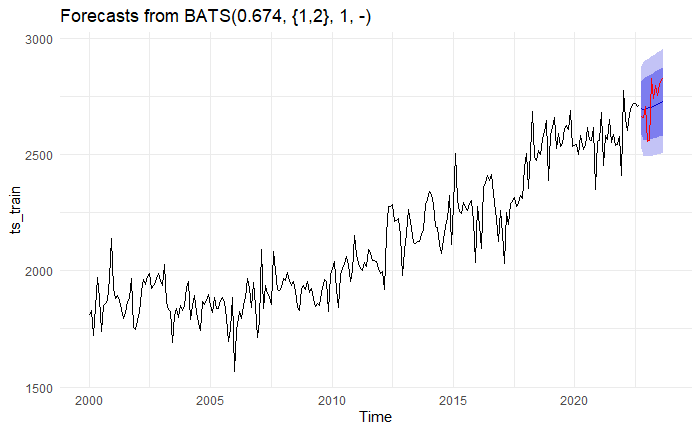
***Graph 10*** *: ARIMA Forecast Plot*

After exponential smoothing model, for the next step, BATS model was conducted. The summary of the TBATS model is given below (Table 3).

***Table 3:*** *Summary of TBATS Model*

|  |
| --- |
| BATS(1, {0,0}, -, -) |
|  |
| Call: tbats(y = ts\_train) |
|  |
| Parameters |
| Alpha: 0.4294992 |
| Seed States: |
| [,1] |
| [1,] 1859.019 |
|  |
| Sigma: 86.47049 |
| AIC: 3970.438 |

Then, to check the normality of the residuals, the Shapiro-Wilk test was conducted and the results shows that the residuals of the TBATS model do not follow the normal distribution, since the p-value is less than 0.05. The forecast plot for the TBATS model is given below.



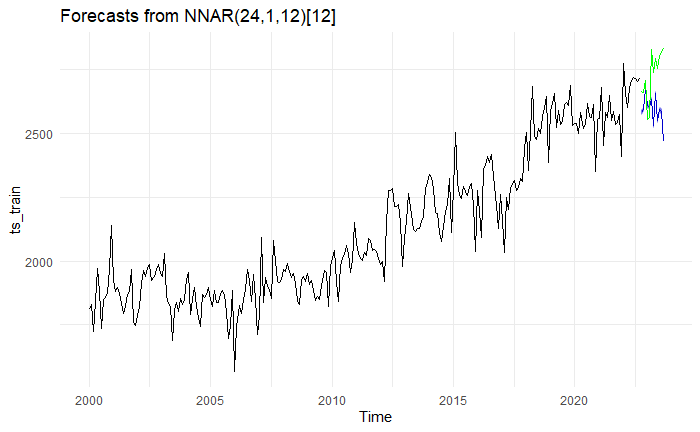
***Graph 11:*** *TBATS Forecast Plot*

As the fourth model, Neural Network model was performed. The following Table 5 represents the summary of the Neural Network model.

***Table 4:*** *Summary of NNETAR Model*

|  |
| --- |
| model |
| ## Series: ts\_train |
| ## Model: NNAR(24,1,12)[12] |
| ## Call: nnetar(y = ts\_train) |
| ## Average of 20 networks, each of which is |
| ## a 24-12-1 network with 313 weights |
| ## options were - linear output units |
| ## sigma^2 estimated as 25.48 |

Normality of the residuals of Neural Network model was checked by using Shapiro-Wilk test and the result shows that the residuals do not follow normal distribution since p-value is less than 0.05. The following plot shows the forecast plot of neural network.



***Graph 12****: NNETAR Forecast Plot*

Finally, the prophet model was conducted. The visual representation of the prophet model is given below.

A graph showing a graph of a trend

Description automatically generated with medium confidence

***Graph 13****: Prophet Prediction Plot*

When the normality of the residuals is checked by using the Shapiro-Wilk test, the residuals do not satisfy the normality.

After completing the models, the forecast values according to each model were performed. Furthermore, their accuracy values are performed and represented in the following tables.

***Table 5:*** *The train accuracy of models*

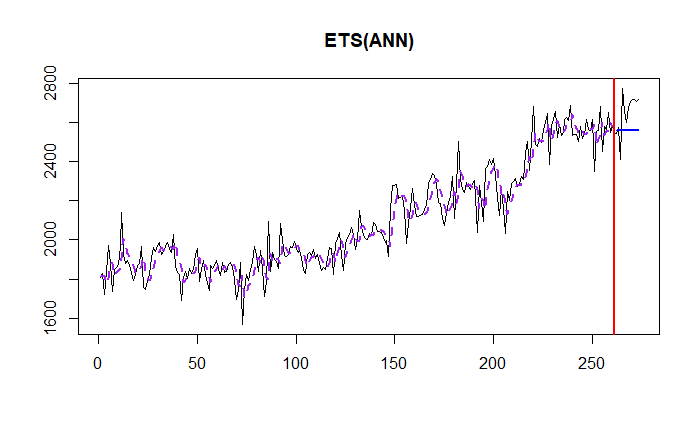
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | ME | RMSE | MAE | MAPE | MASE | ACF1 |
| ARIMA | 12.59486 | 82.87002 | 61.2298 | 2.8969 | 0.5523 | -0.0167 |
| ETS | 7.6355 | 86.403 | 65.110 | 3.0928 | 0.5873 | 0.0839 |
| TBATS | -3.557 | 83.446 | 61.733 | 2.9308 | 0.5568 | -0.0608 |
| NNETAR | **0.0337** | **5.0482** | **3.2945** | **0.1568** | **0.0297** | **-0.0156** |
| PROPHET | 10.219 | 84.174 | 69.867 | 2.583 |  | 0.2969 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | ME | RMSE | MAE | MAPE | MASE | ACF1 |
| ARIMA | 38.992 | 102.48 | 92.892 | 3.392 | 0.837 | 0.563 |
| ETS | 17.3095 | 96.169 | 84.047 | 3.0958 | 0.758 | 0.4111 |
| TBATS | 19.086 | 88.655 | 77.132 | 2.840 | 0.695 | 0.330 |
| NNETAR | 136.3096 | 180.501 | 152.526 | 5.5007 | 1.3758 | 0.559 |
| PROPHET | 10.219 | 84.174 | 69.867 | 2.583 |  | 0.2969 |

***Table 6:*** *The forecasting performance of models*

The Table 5 and Table 6 gives us a mixed result. In Table 5, it can be seen that the optimal train accuracy model is neural network model. However, in Table 6, it is hard to choose an optimal model. So, for the forecast performances, it might be Prophet model that has the optimal test accuracy.

The forecasting performance of the models are represented in the following plot.



V. DISCUSSION AND CONCLUSION

In this research, as a first step, data is separated into two parts as train set and test set. After that, the anomalies in the train data were checked and stated that there is no anomalies in the train set. Then, the stationarity condition was checked by using tests such as KPSS, ADF, HEGY and Canova-Hensen. According to the results of these tests, the necessary steps such as differencing taken to make the series stationary. After making the series stationary, suggested models that may fit the series were checked. As a result, the appropriate model was selected from these models by looking at their significance and AIC-BIC values.

On residuals, diagnostic checks are applied after the optimal model has been fitted. We faced with the problem of non-normality, and other models were tried to satisfy the condition, but this method did not give any healthy solution. Then, the tests such as Breusch-Godfrey and portmanteau lack of fit test was conducted to show that the residuals are uncorrelated. As the last assumption, heteroscedasticity was checked both visual and formal way. As a result of the Engle’s ARCH test, we conclude that there is no ARCH effects presence. For the forecasting part, five different forecast methods were used such as ARIMA, ETS, TBATS, NNETAR and Prophet. Each forecast method was considered both visual and formal way. For the each model, accuracy for train and test set was evaluated and represented at the Table 5 and Table 6. All in all, the results gave us mixed results. For train accuracy neural network method was the best we cannot say that for the test set. It is hard to decide the best model for test set, but it might be prophet since it has lower RMSE and MAPE values.

Finally, we learned a lot of information, some healthy and some unhealthy, about the time series analysis of natural gas consumption from this project. As a researcher, we gained in-depth knowledge in how to analyze and interpret the time series analysis and forecast part of this project.

VI. REFERENCES

- Lecture Notes

- Recitation Notes

-U.S. Bureau of Transportation Statistics, Natural Gas Consumption [NATURALGASD11], retrieved from FRED, Federal Reserve Bank of St. Louis;

https://fred.stlouisfed.org/series/NATURALGASD11, January 20, 2024.