Time Series Analysis of Henry Hub Natural Gas Spot Price

*Abstract*— This paper forecasts monthly Henry Hub natural gas spot prices using various forecasting models such as ARIMA, ETS, and NN. R-Studio is used. Preprocessed data is used before making projections. Outlier detection, anomaly cleaning, and stationarity checking are all performed. Various forecast models are fitted, and their performances on both train and test data are compared using a variety of comparison criteria.

Keywords—Forecast, natural gas spot price, ets, etc.

# INTRODUCTION

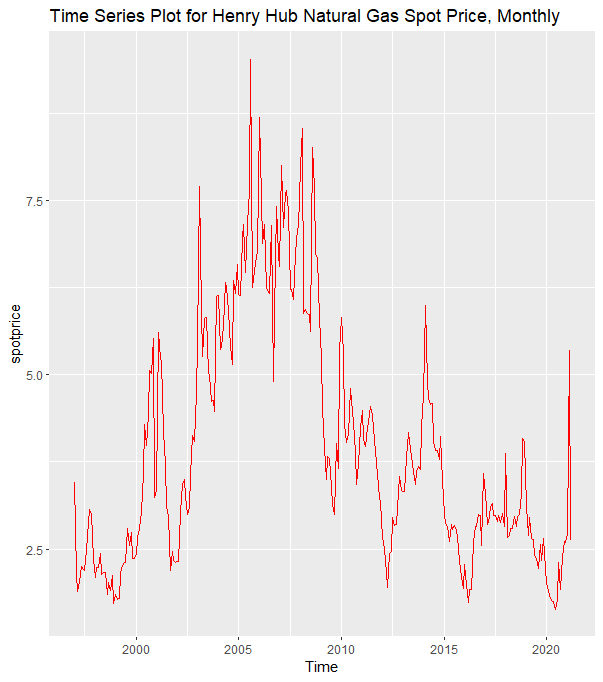
Henry Hub is a natural gas pipeline in Erath, Louisiana, that serves as the official delivery location for New York Mercantile Exchange futures contracts (NYMEX).The Henry Hub Natural Gas Spot Price is determined in US dollars per million BTU.

The primary goal of this research is to comprehend Henry Hub's behavior. A monthly natural gas pricing dataset from 1997 to 2021 is researched and explored for this aim.

ARIMA, NN and ETS models are used while doing forecast. Then, their performance was examined using various comparison criteria such as root mean squared error (RMSE), mean absolute error (MAE), and so on. R Studio 2022.12.0 was used for all analyses during this investigation.

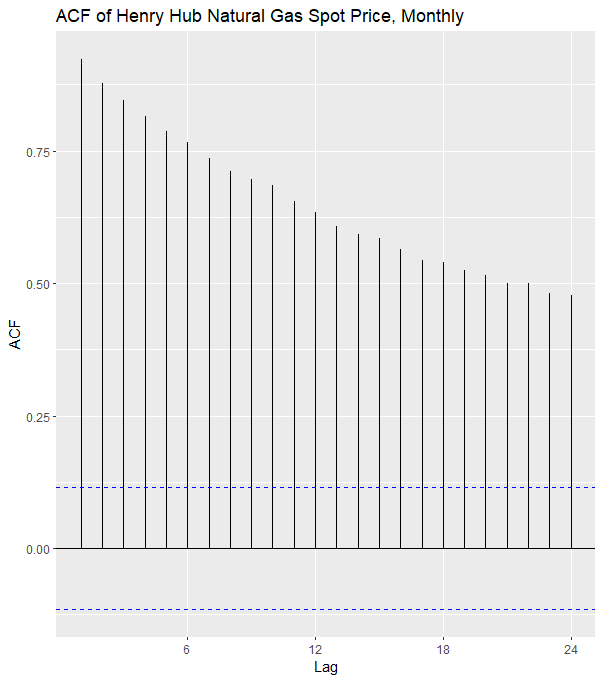
# DATA DESCRIPTION AND PREPROCESSING

The data set is taken from https://github.com/rishabh89007/Time\_Series\_Datasets/blob/main/HH%20Spot%20Price.csv. The data set contains 291 observations recorded from 1997 to 2021 monthly.



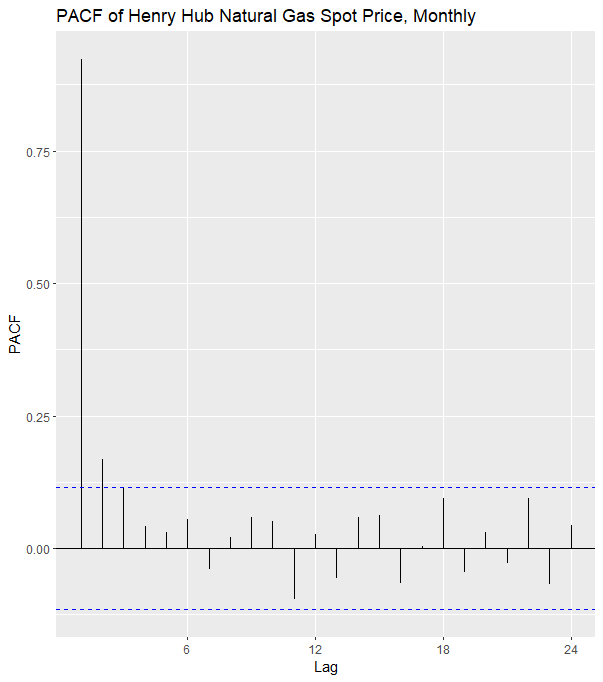
***Graph 1****: Time Series Plot of Henry Hub Natural Gas Spot Price*

The plot seems non-stationary. The mean term is not constant. It is changed over time there is an increasing trend. It also displays some up’s and down’s which are the indication of stochastic trend. However, we cannot reach a particular judge from that plot.



***Graph 2****: ACF Plot of Henry Hub Natural Gas Spot Price*

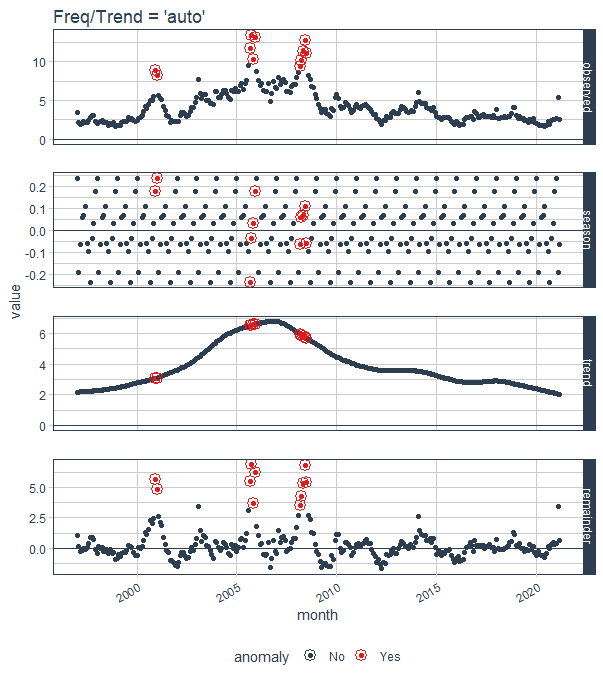
The ACF plot shows an exponential decay, which supports the result indicated in the first plot. Hence, it is said that we have a non-stationary process.



***Graph 3****: PACF Plot of Data Set*

PACF is observed to cut out after the second lag. However, since the non-stationary process has already been determined by inspecting the ACF and time series plots of the data, there is no need to interpret the PACF plot of the data set.

Before the analysis, the data set is formatted into a tible format to detect anomalies in the data. Anomalies are checked using STL decomposition, and it shows that the series has abnormalities.



***Graph 4****: Anomaly Detection Plot*

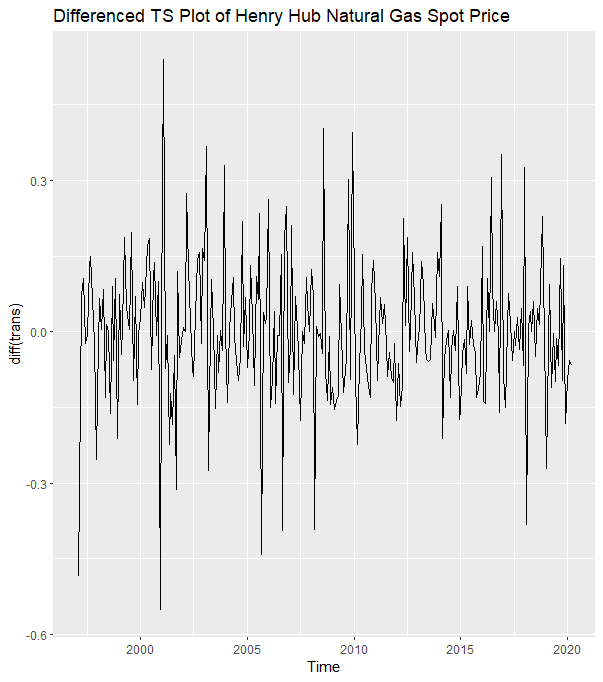
The anomalies in the data are removed and replaced by interpolated values by the clean\_anomalies function. After cleaning irregularities, the data set is divided into a test set and a train set. While doing this, the last 12 observations are kept as a test set since monthly data is used.

When working with time series, covariance-stationarity is a necessary assumption of time series modeling process. Therefore, it is reasonable to look for a variance-stabilizing transformation that will make the data closer to fulfilling this assumption. When we generate, a lambda value for our dataset and Boxcox transformation is the most appropriate transformation for the data set to stabilize the variance.

After satisfying stationary in variance, the required tests are applied to achieve stationary in mean. Since the monthly dataset is used in this study, the non-stationary is checked and analyzed by both KPSS and ADF tests.

The first level of the KPSS test shows that the data is not stationary. (p<0.05) Then, the second level of this test represents that the data suffer from a stochastic trend. (p<0.05) The ADF test also represents this result. (p>0.05). Also, the existence of unit roots is detected from the ADF test. Seasonality check is done by HEGY.test function and concluded that data don’t have any seasonality problem.

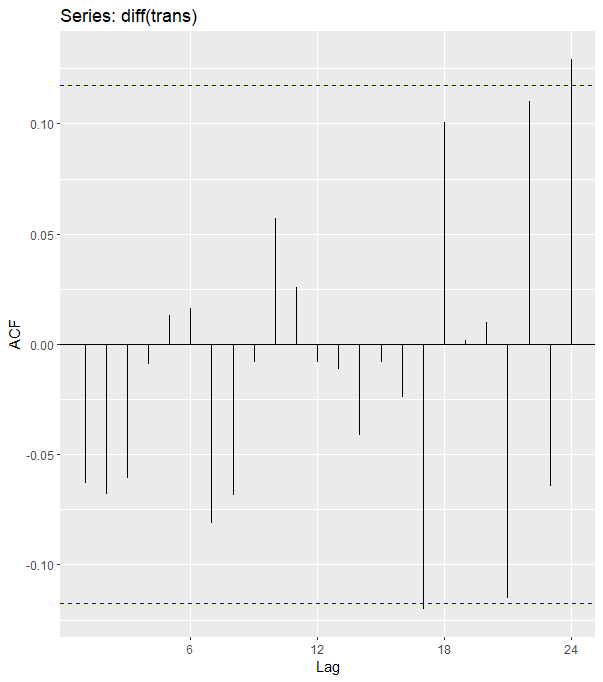
To remove unit root, non-stationarity, and stochastic trend, differencing should be applied. The number of the regular and seasonal differences are calculated by ndiffs and nsdiffs functions, respectively, and it concluded that data need only one regular difference. After taking the first-order regular difference of the series ADF test, the KPSS test and visual tools(autoplot, ACF, and PACF) were examined. The result of ADF shows that the data does not have a unit root (p<0.05), the KPSS test shows that the data is stationary, and the trend is removed (p<0.05). The visual tools also supported the formal test results.



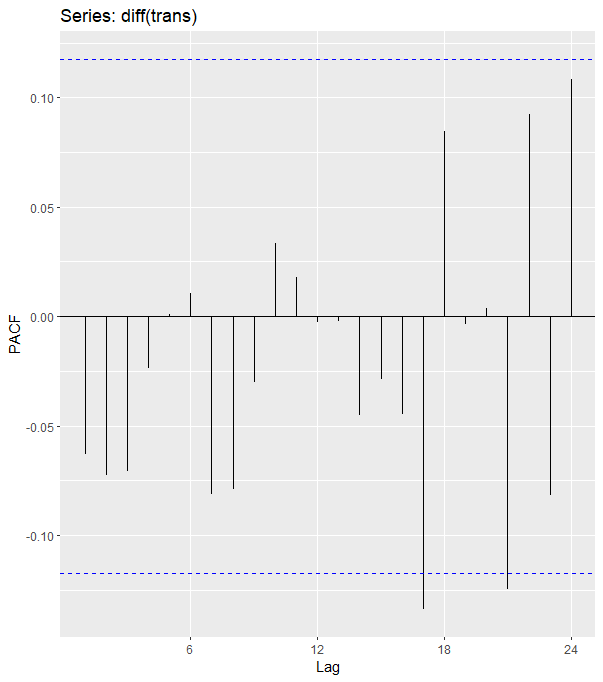
***Graph 5****: Time Series Plot of Differenced Data Se*

# MODEL SUGGESTION

After obtaining the stationary series, ACF and PACF plots of the series used to suggest a model.

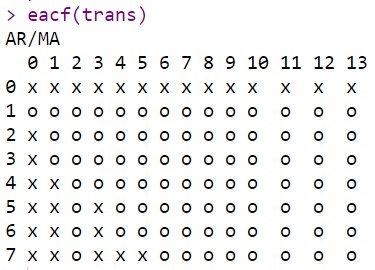


***Graph 6****: ACF Plot of Stationary Data Set*



***Graph 7****: PACF Plot of Stationary Data Set*

It is seen that ACF and PACF show White Noise behavior, i.e., all lags are inside the WN band. The Extended Sample Autocorrelation Function (EACF) method used to detect ARMA process parameters.



Main aim in EACF method is to select the model with fewer parameters by drawing a triangle consisting of "o" terms. Hence, the table suggests the ARIMA (1,1,0) model for the data set.

AIC table is the last method used to identify the model. To select the best one, the AIC values are considered. The order having the smallest AIC value shows the most appropriate model for the data set. Also, if the ratio between these estimates and their standard errors (s.e) are greater than +2 or less than -2 , we can say that these parameters are significant and the model is also significant. Using this information, it is said that ARIMA(1,1,1)(0,0,0)12, the model with a small AIC value, is selected as the best model.

***Table 1:*** *AIC Table*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| p | d | q | P | D | Q | M | AIC |
| 0 | 1 | 0 | 1 | 0 | 1 | 12 | -297.51 |
| 0 | 1 | 0 | 0 | 0 | 0 | 12 | -300.33 |
| 1 | 1 | 0 | 1 | 0 | 0 | 12 | -297.5 |
| 0 | 1 | 1 | 0 | 0 | 0 | 12 | -299.7 |
| 1 | 1 | 1 | 0 | 0 | 0 | 12 | -302.76 |
| 0 | 1 | 2 | 0 | 0 | 0 | 12 | -299.59 |
| 1 | 1 | 2 | 1 | 0 | 0 | 12 | -298.77 |

# MODELLING AND DIAGNOSTIC CHECKING

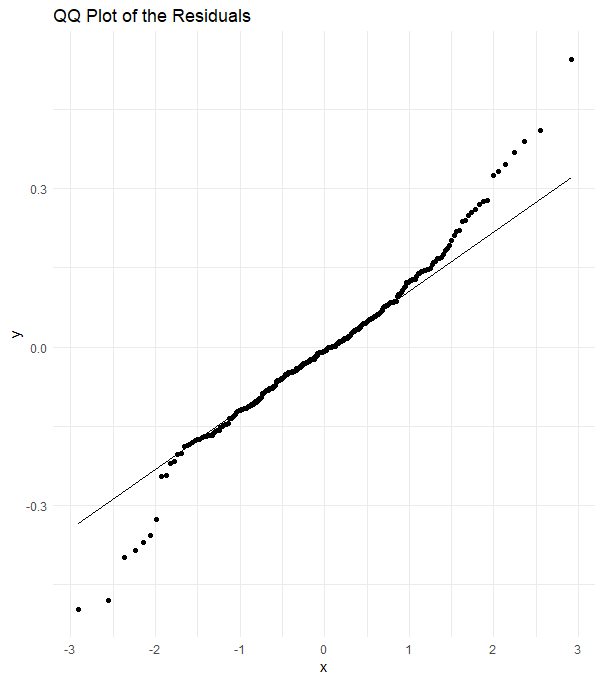
After choosing the best model, diagnostic checks of the model is done.

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

|  |
| --- |
| 1. ARIMA(1,1,1) |
| 1. Coefficients ar1 ma1 |
| 1. 0.8432 -0.9280 |
| 1. s.e 0.0759 0.0517 |
| 1. sigma^2 estimated as 0.01941 log likelihood=154.38 2. AIC= -302.76 AICc=-302.68 BIC=-291.88 |

Model checking in time series analysis is similar to the traditional regression analysis and depends on the residual analysis.

Normality assumptions are checked with the visual inspection tool Q-Q plot, and the formal tests are Shapiro-Wilk and Jarque-Bera.

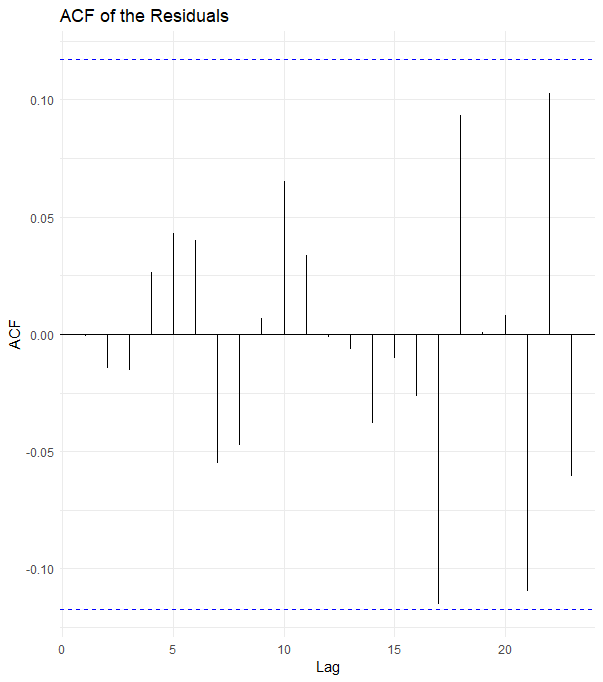


***Graph 7****: QQ plot of the standard residuals*

Since the Q-Q plot shows that most of the model's residuals lie on a 45-degree straight line, this indicates residuals may be normally distributed. To be sure about non-normality, Shapiro-Wilk and Jarque-Bera tests should be applied. Unfortunately, both tests show that errors do generally not distribute (p<0.05). We have non-normal residuals, but we assume normality and continue the analysis by assuming normality.

Now, the serial autocorrelation is checked. The residual plot as visual, Ljung-Box, and Box-Pierce test methods are applied to detect serial correlation.

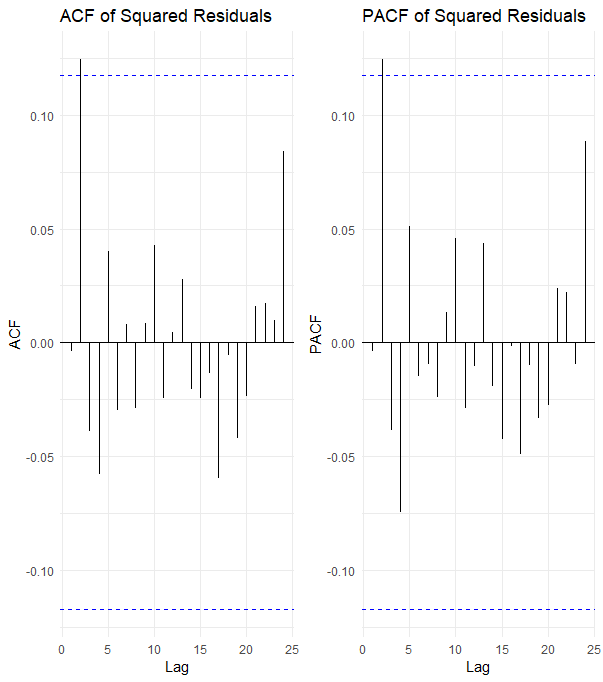
The first way of checking the serial correlation is the ACF plot of residuals. All spikes should be in the White Noise band because we do not have any correlation problem.



***Graph 8****: ACF Plot of the residuals*

As seen, all spikes are in the White Noise. Thus, our residuals are uncorrelated. After checking the ACF plot of residuals, we should apply the formal tests mentioned above. To do so, both Box-Ljung and Box-Pierce are applied, showing that residuals belonging to the best model are uncorrelated. (p>0.05)

The last assumption checked is the heteroscedasticity of the residuals. We can look at the ACF and PACF plot of squared residuals to test this assumption and apply White's and Bresuch-Pagan's tests.



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

It is seen that most of the squared residuals are in the 95% White Noise Band. Therefore, it is said that the errors are homoscedastic. However, a formal test should be applied to check this judgment. The errors are homoscedastic according to the Studentized Breusch-Pagan test (p>0.05). Thus, we have constant variance over time and don't need to use the GARCH-type model.

After the ARIMA model, we'll try to find the best exponential smoothing model. There are different forecasting methods, such as TBATS, PROPHET, etc.; however, these methods are used if the data represents multiple seasonality,..etc. Therefore the data examined in this paper don't suffer from this kind of problem, so applying these methods is meaningless; only the ETS method and NN are applied by using ets and nnetar functions under the forecast package in R. The best exponential smoothing model for the series is given below.

***Table 3:*** *Summary of the ETS Model*

|  |
| --- |
| ETS(M,N,N) |
| Smoothing parameters: |
| alpha = 0.9693 |
| Initial states: |
| l = 3.3629 |
| sigma: 0.1389 |
| AIC AICc BIC |
| 1206.071 1206.158 1216.964 |

It is seen that we have an exponential smoothing model having a multiplicative error. After fitting the model, the residuals of the ETS model are checked by the Shapiro-Wilk test and seen that they do not follow the normal distribution. (p<0.05)

As a third model in the analysis, the Neural Network model fits where past observations are considered input variables. The model details are

***Table 5:*** *Summary of the NNETAR Model*

|  |
| --- |
| model |
| Series: train |
| Model: NNAR(16,1,8)[12] |
| Call: nnetar(y = train) |
| Average of 20 networks, each of which is |
| a 16-8-1 network with 145 weights |
| options were - linear output units |
| sigma^2 estimated as 0.03249 |

The model is NNAR(16,1,8)[12]. It is a neural network with the last observation as input for forecasting output and one neuron in the hidden layer. If we are looking at NN residuals, we have non-normal residuals according to the Shapiro-Wilk test. (p<0.05)

After fitting the models, we obtain forecast values from each method using the forecast function and calculate their accuracy. The accuracy of the models is given below.

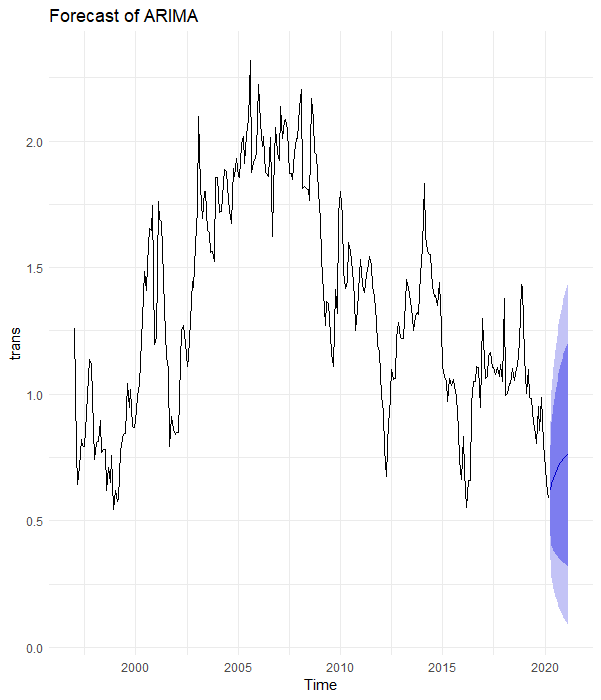
***Table 4:*** *The train accuracy of models*

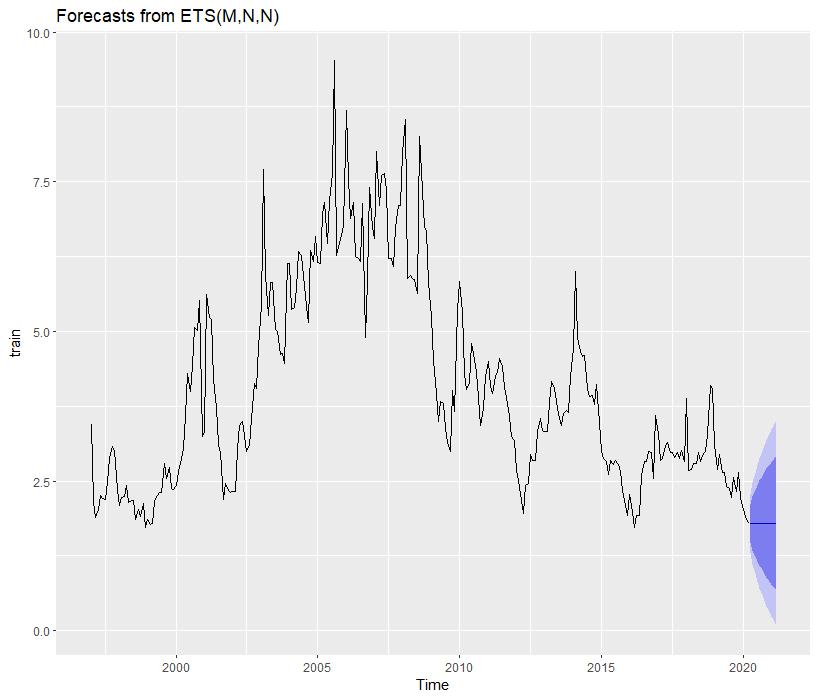
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | ME | RMSE | MAE | MAPE | ACF1 |
| ARIMA | -0.0029 | 0.1385 | 0.1014 | 8.3886 | -0.0007 |
| ETS | -0.0058 | 0.6401 | 0.408 | 9.8149 | -0.117 |
| NN | 0.0001 | 0.1812 | 0.1230 | 3.6397 | -0.0028 |

***Table 5:*** *The forecasting performance of models*

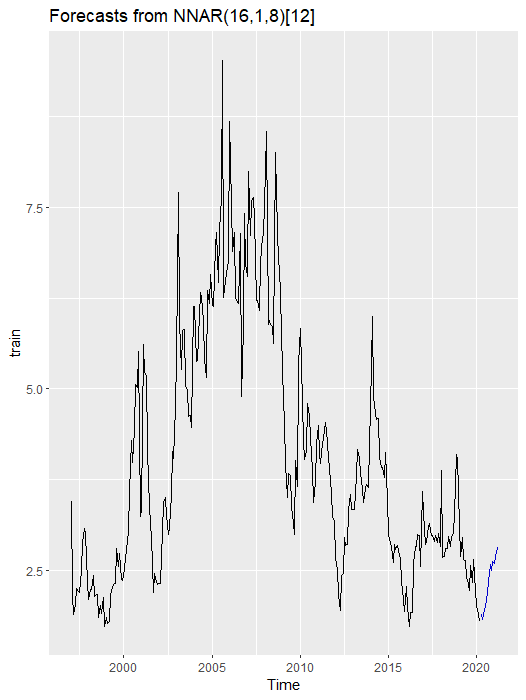
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | ME | RMSE | MAE | MAPE | ACF1 |
| ARIMA | 0.4209 | 0.9955 | 0.5699 | 18.1750 | 0.2115 |
| ETS | 0.6545 | 1.1593 | 0.7020 | 22.0286 | 0.2856 |
| NN | 0.1021 | 0.8820 | 0.4738 | 16.4665 | 0.1534 |

Both tables show that the NN model outperforms the other methods in comparison to the train and test set concerning all measures. The forecasting performance of the models can also be observed from the following plots.

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



***Graph 11****: Forecast Plot of ETS*



***Graph 12****: Forecast Plot of NN*

# DISCUSSION AND CONCLUSION

The stationary check of the series was the first step in this study after data cleaning and dividing; however, after looking at time series, ACF&PACF plots, and results of the KPSS, HEGY, and ADF tests, it is clear that this requirement was not met, because there was a stochastic trend and unit root in the series. To solve these problems, differencing methods were applied. After making the process stationary and removing unit root and trend, some tentative models were suggested using specific techniques such as ACF&PACF plots, ESACF table, and AIC table. The best model with the fewest significant parameters was then picked from among these preliminary models using AIC comparison and t statistics.

After fitting the best model, remaining diagnostic checks are performed. The problem of non-normality of errors emerged at this stage, and Boxcox transformation was used. Still, these have not given a satisfactory solution since the series has heavy-tailed residuals. Nonetheless, because the series features heavy-tailed residuals, these have not provided a viable solution. However, using visual inspection tools and formal testing, it is confirmed that the errors are uncorrelated and homoscedastic. Aside from the top ARIMA models, ETS and NN forecasting approaches were investigated, and forecasts were generated. Finally, when compared to other models, NN has the best performance in modeling series and forecasting future values. Overall, during the analytical process, we encountered some issues, some of which had solutions while others did not. Despite these issues, the best model was obtained.

# REFERENCES

Chen, J. (2022, December 13). *What is Henry Hub? definition, location, owner, and connections*. Investopedia. Retrieved January 22, 2023, from https://www.investopedia.com/terms/h/henry\_hub.asp