Time Series Analysis of Global Price of Energy Index

Kadir Şimşek
Department of Statistics
Middle East Technical University
Ankara, Turkey
kadir.simsek@metu.edu.tr

***Abstract*—In this study, time series analyses are performed using monthly Global Energy Price Index data from 1992 to 2023. First, outlier detection and data cleaning were performed with the anomalize package, and then the data were stabilized with the Box-Cox transformation. The stationarity analysis of the series is verified by KPSS, ADF and HEGY tests, and (1,1,0)(1,0,1) is selected as the most appropriate SARIMA model. GARCH family models (sGARCH, apARCH) are tested to capture volatility and the most successful result is obtained from the sGARCH(2,2) model. In addition, ETS, TBATS, NNETAR and Prophet models are also used for forecasting purposes, both default and hyperparameterized. Different models were compared according to the error criteria (RMSE, MAPE, etc.) and it was observed that the sGARCH(2,2) model was the most successful model, especially in terms of RMSE. The analysis and forecasts indicate that the energy price index will be on an upward trend in the coming period. As a result, it is recommended to use the sGARCH(2,2) model considering both its forecasting success and its ability to capture volatility.**

Keywords—Energy, KPSS,ADF,HEGY,SARIMA,sGARCH, Hyperparameterized.

I. Introduction

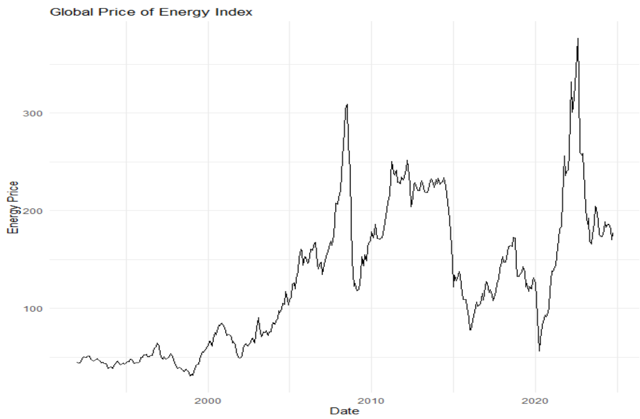
The energy market has a decisive role in global economic stability and growth. Fluctuations in energy prices due to supply-demand imbalances and geopolitical developments have increased the interest in research in this area. This study analyzes the Global Energy Price Index (PNRGINDEXM) data provided by the Federal Reserve Economic Data (FRED) to examine the monthly energy price trends between 1992 and 2024 and their likely future course.

As a first step of the analysis, outliers in the data set are identified and cleaned. Then, the statistical properties of the data are improved by applying the Box-Cox transformation and the structure and trend of the series are determined with the help of stationarity tests (KPSS, ADF, HEGY). Then, different time series forecasting methods such as SARIMA, GARCH family models (sGARCH and apARCH), ETS, TBATS, Artificial Neural Network (NNETAR) and Prophet are applied. In this way, the forecasting performance of different methodologies on energy prices is comparatively evaluated.

The results obtained are important for forecasting the possible future trends of the energy price index and to better understand the impact of price fluctuations on the economy. Thus, policymakers, investors and academics can better analyze developments in the global energy market and shape their strategic decisions accordingly.

II. Data Description and Preprocessing

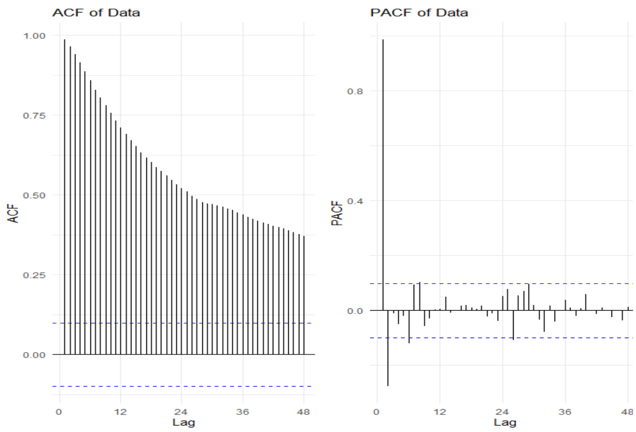
The data set was taken from https://fred.stlouisfed.org/series/PNRGINDEXM where the Federal Reserve Bank of New York shares its data sets. The dataset includes global energy price index data for 345 months from 1995 to 2023.



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

When we look at the graph, our data appears to be non-stationary. It is difficult to comment on the stability of the average. The chart shows the fluctuations of the energy price index from approximately the mid-1990s to the 2020s. Prices, which were relatively low and stable in the early period, rose rapidly from the mid-2000s and recorded their first major peak around 2008. Following the 2009 financial crisis, there was a decline, followed by a partial recovery in the 2010-2014 period. After 2014, the index started to decline again, rising significantly especially in the 2021-2022 period due to global developments (e.g., post-pandemic demand growth and geopolitical tensions), but followed a downward trend again by 2023.

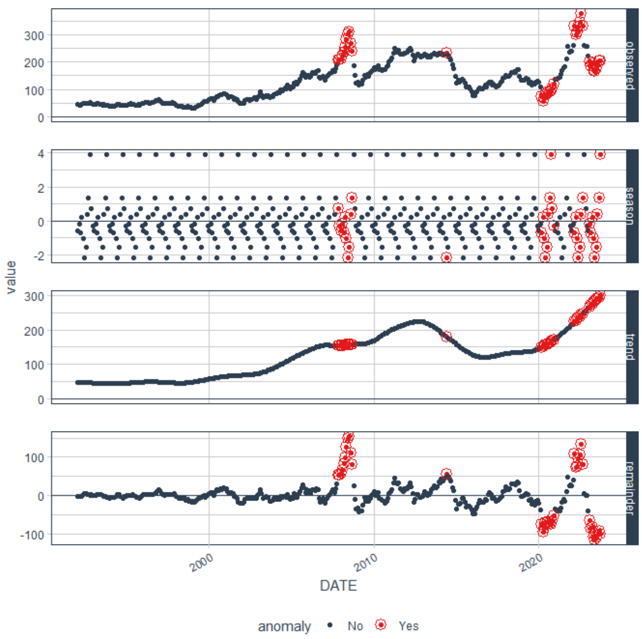
In terms of time series analysis, these trends indicate that the series has a highly volatile structure. Large price jumps and shocks emphasize that significant exogenous factors are at play in the energy market. In the literature, such fluctuations are usually handled by transformation (e.g. log or Box-Cox), differencing (to ensure stationarity) or GARCH-like volatility models. Moreover, the strong and sudden tendency of prices to fluctuate requires that seasonality, trend and structural break analysis be taken into account in model selection. In this way, future forecasts can be made more accurately and policy or strategy decisions can be better grounded



***Graph 2** : ACF and PACF Plots of Data Set*

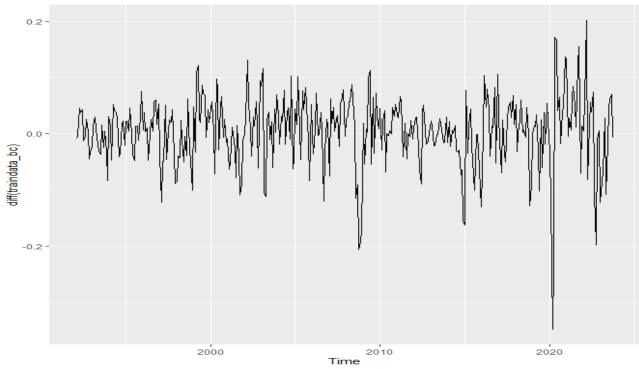
It can be said that the MA structure is not appropriate as a linear decrease is observed in the ACF graph, whereas a significant jump in the PACF at the first lag indicates a possible AR component.

At the beginning of the analysis, the dataset is divided into two subsets: training and testing. The most recent 12-month observations are used as the test set to evaluate the forecasting performance of the model. The anomalize package was used to detect and handle anomalies in the data set, ensuring that the analysis is based on a clean and reliable data set.



***Graph 3** : Anomaly Detection Plot*

Anomalies in the train data were identified and removed from the data. Then, the Box-Cox test is applied. During the test, energy price index values below zero were turned positive. Since the lambda value was far from one as a result of the test, Box-Cox transformation was performed and the lambda value approached one as a result of the test. Then, the KPSS test for “Level” was applied on the train data. Since the P-value < 0.05 as a result of the test, the H0 hypothesis was rejected and it was determined that our series was non-stationary. Subsequently, the “Trend” test revealed that the series has a stochastic trend. Subsequently, the HEGY test was applied and it was found that the series has a normal unit root. The time series data were differenced once. The subsequent KPSS test revealed that the series is stationary and has a deterministic trend. Then the “Augmented Dickey-Fuller test” was applied. After applying it to the train data, it was found that the series was not stationary, so it was applied again to the differentiated series. The test result of the differentiated data was stationary.

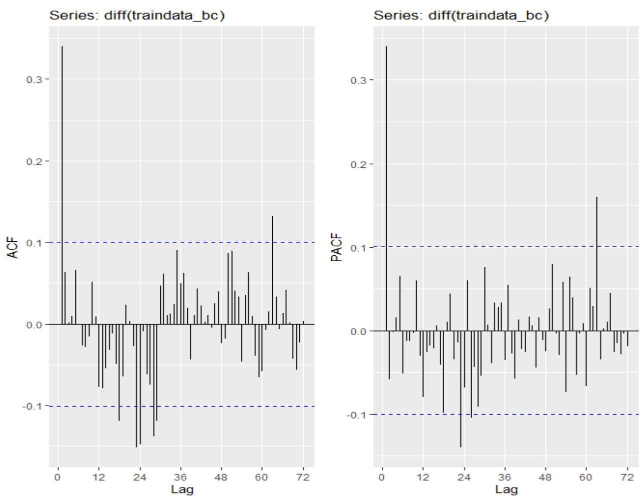


***Graph 4** : Time Series Plot of Differenced Data Set*

KPSS and Augmented Dickey-Fuller tests show that the differencing data are stationary and have a deterministic trend.

MODEL SUGGESTION

Once it is clear that the series is stationary, a model should be proposed using the ACF and PACF plots.



The ACF and PACF plots of the differenced data reveal the main characteristics of the series for model identification. In the ACF plot, significant increases in the first lags followed by a rapid decline indicate the presence of short-term dependencies. The PACF plot, on the other hand, shows a sharp break after the first lag, indicating the potential suitability of an AR(1) process.

Given this behavior, an ARIMA(1,1,1,0) model could be a strong initial candidate, as differentiation has probably reached stationarity. Despite the absence of a seasonal unit root, a SARIMA(1,1,0)(1,0,1) model is proposed given the seasonality effect in the ACF and PACF plots.

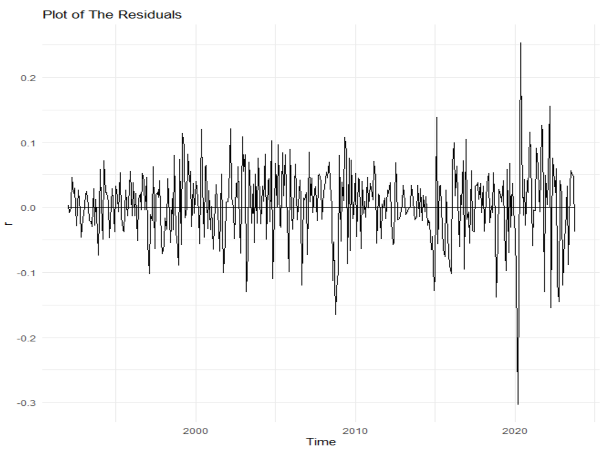
MODELLING AND DIAGNOSTIC CHECKING

After determining the models that could fit the series, it was determined which one was more suitable by looking at certain criteria. To find the most suitable model,the lowest AIC value was examined and SARIMA(1,1,0)(1,0,1)[12] was determined as the most suitable model.

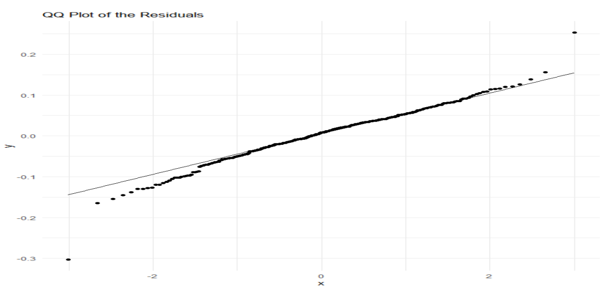
***Table 1**: Summary of Model*

ARIMA(1,1,1)(0,0,1)[12]			
Coefficients			
	AR1	MA1	SMA1
s.e.	0.0482	0.1263	0.1050
sigma^2 = 0.003369: log likelihood = 544.94 AIC=-1081.88 AICc=-1081.78 BIC=-1066.11			

After the model was proposed, we proceeded to check the diagnostic assumptions and first checked whether the residuals were normally distributed.



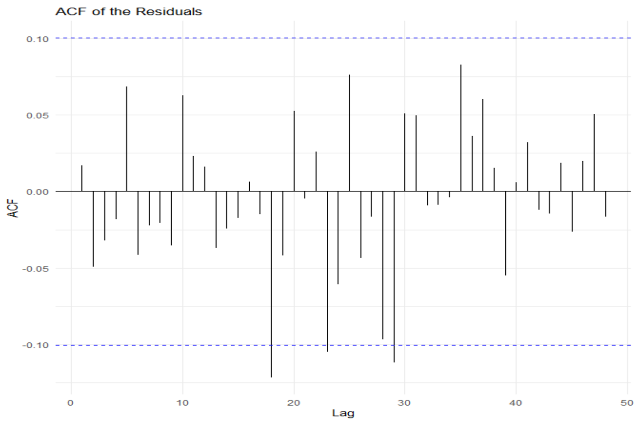
The time series plot of residuals shows a random distribution around zero, suggesting that the model can explain the data set well and that there is no significant structure or autocorrelation in the model residuals. However, volatility spikes are noticeable in some periods with extreme values, which may require additional model evaluation.



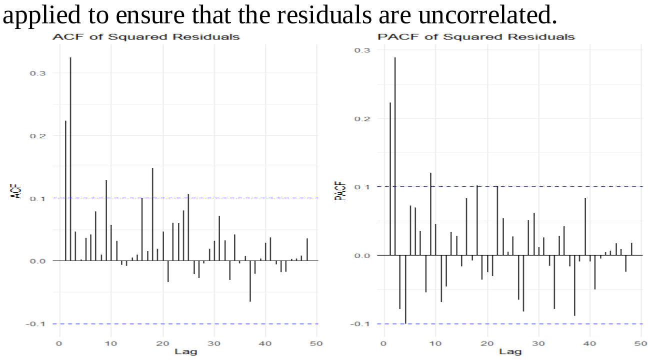
When we examine the QQ plot, it is seen that most of the residuals conform to the normal distribution and are arranged on a linear line. However, some deviations were observed at the extremes. This indicates that the model generally satisfies normality, but there are slight deviations for the extreme values.

After the normality was tested in the graph, the Jarque-Bera test and the Shapiro-Wilk test were used to test whether it was normally distributed.

After applying the Shapiro-Wilk and Jarque-Bera tests, Ho was rejected because the p-value was smaller than alpha. Therefore, it was observed that the residuals of the data set do not follow a normal distribution.



The ACF plot of the residuals shows that most of the autocorrelations are within the confidence intervals and there is no significant autocorrelation in the residuals of the model, indicating that the model adequately fits the data. Formal tests were applied to ensure that the residuals are uncorrelated. According to the results of the Breusch-Godfrey Test, since the p-value is greater than α, we are 95% confident that the residuals of the model are uncorrelated. The Ljung-Box and Box-Pierce tests are then applied to ensure that the residuals are uncorrelated.

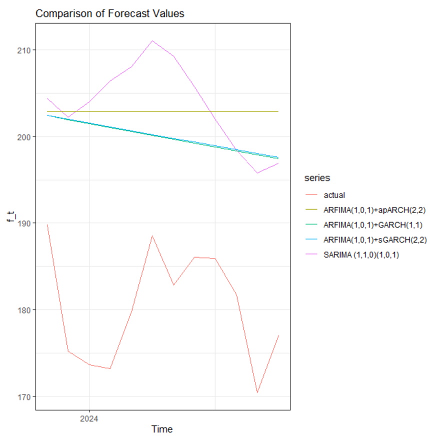


The ACF and PACF plots of the squares of the residuals show that there are significant autocorrelations especially in the first few lags. This suggests that there may be a time-varying structure (heteroskedasticity) in the residual variance of the model. Therefore, applying a GARCH or similar model can better explain the time-varying variance and improve the model. Considering the possibility of heteroskedasticity, white test and arch.test() tests were applied and heteroskedasticity was observed in the model.

```
> #For SARIMA model
> accuracy(f_t, test_data)
                 ME     RMSE      MAE       MPE     MAPE      ACF1 Theil's U
Test set -23.33745 24.03742 23.33745 -13.03193 13.03193 0.3728415  3.522421
>
> #For GARCH model
> accuracy(f_vec1, test_data)
                 ME     RMSE      MAE       MPE     MAPE      ACF1 Theil's U
Test set -19.55324 20.52358 19.55324 -10.96976 10.96976 0.3523545  3.006678
>
> #For sGARCH(2,2) Model
> accuracy(f_vec_2, test_data)
                 ME     RMSE      MAE       MPE     MAPE     ACF1 Theil's U
Test set -19.66609 20.62841 19.66609 -11.03251 11.03251 0.348581  3.022347
>
> #For apARCH(2,2) model
> accuracy(f_vec_3, test_data)
                 ME     RMSE      MAE       MPE     MAPE     ACF1 Theil's U
Test set -22.53794 23.37245 22.53794 -12.62939 12.62939 0.276022  3.432111
```

ARCH and GARCH models are fitted to resolve heteroskedasticity. According to the procedures, sGARCH(1,1) sGARCH(2,2) apARCH(2,2) models were tested. In the observations made, SARIMA model: It shows the highest RMSE (24.04%) and MAPE (13.03%), indicating a weaker forecasting performance. GARCH(1,1): It was the most successful model on the test set, providing the lowest RMSE (20.52) and MAPE (10.97%) values. sGARCH(2,2) it is performed similarly to GARCH(1,1), with RMSE = 20.63 and MAPE = 11.03%. apARCH(2,2): Despite capturing asymmetries, the RMSE (23.37) and MAPE (12.63%) values slightly underperformed the GARCH(1,1) model.

After the comparisons,one may claim that the GARCH(1,1) is the best model amoung the others, since it has lowest error rates s (RMSE and MAPE). Therefore, it is recommended to use the GARCH(1,1) model to model and forecast the volatility of energy prices. However, in the context of this analysis, the GARCH(1,1) model is the best choice in terms of accuracy and simplicity.



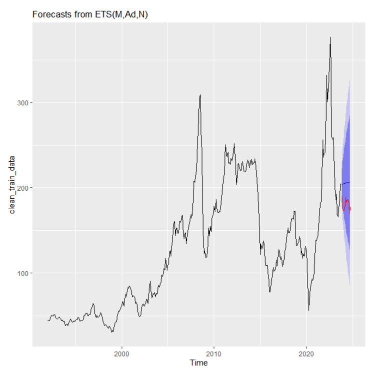
The GARCH(1,1) model outperformed the other models both in terms of overall fit and in terms of capturing fluctuations. Therefore, GARCH(1,1) may be preferred as the most appropriate model for energy price forecasts. The performance of the SARIMA model, on the other hand, is quite poor.

As a result of the model comparisons, the GARCH(1,1) model was found to perform the best; therefore, we proceeded to forecast analysis to predict the future trends of energy prices.

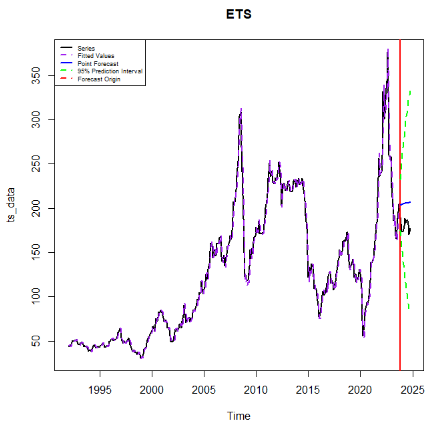
Table 2 : Summary of ETS Model

ETS(M,Ad,N)
Call: ets(y = clean_train_data, model = "ZZZ")
Smoothing parameters:
alpha = 0.9999
beta = 0.097
phi = 0.8
Initial states:
l = 44,6932
b = 0.9116
s = -0.0992 0.0372 0.0377 0.1595 0.1045 -0.0065 -0.0084 -0.0289 -0.0265 -0.0977 -0.0452 -0.0265
sigma: 0.0714
AIC AICc BIC 3814.824 3815.048 3838.496

The ETS(M,Ad,N) model is identified as an appropriate model for forecasting energy prices. The model includes Multiplicative error, Additive damped trend and non-seasonal structure components. A high α value (0.9999), indicating that the model adapts quickly to the data, β value (0.097) indicates that the trend update occurs more slowly. Damped trend showing the effect ϕ value (0.8) indicates that the trend decreases over time and tends to stabilize. The initial level of the model (l=44.6932) and trend (b=0.9116) and the standard deviation of the error components (0.0714), indicating that the model has a consistent error structure. The AIC, AICc and BIC values were calculated as 3814.824, 3815.048 and 3838.496, respectively, providing a reference for the overall suitability of the model. Looking at the error measures in the training set, the MAPE value was found to be 5.30%, indicating that the predictions are quite accurate. However, the ACF1 value (0.237) indicates a slight autocorrelation in the error components. Overall, the ETS(M,Ad,N) model is found to be an effective forecasting tool, reflecting the non-seasonal nature of energy prices and their decreasing trend over time.



Forecasts based on the ETS(M,Ad,N) model project a slight upward trend in the energy price index in the future. While the forecasts are consistent with historical data, the forecast intervals shown by the blue shading widen into the future. This indicates an increase in uncertainty in long-term forecasts. The model can be considered a reliable tool, especially for short-term forecasts.

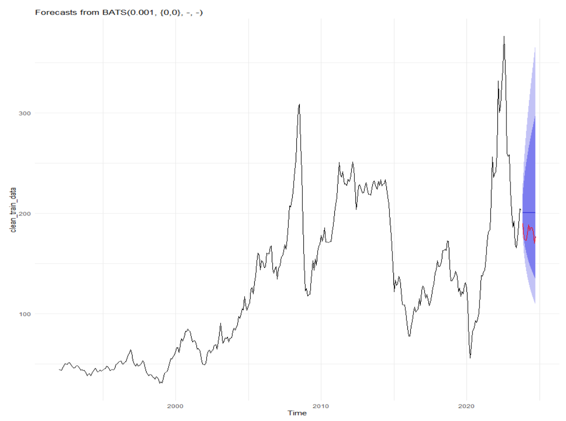


This graph shows the comparison of the predictions made using the ETS model with historical data. The black line represents the historical data, the blue line represents the values predicted by the model and the green dashed line represents the 95% prediction interval. The red line marks the point where the prediction starts, i.e. the separation of the training and test set. The forecasts are fairly close to the historical data and reflect the general trend in energy prices. However, the widening of the 95% prediction interval indicates an increase in uncertainty, especially for future dates. The model has once again demonstrated that short-term forecasts are in line with historical data.

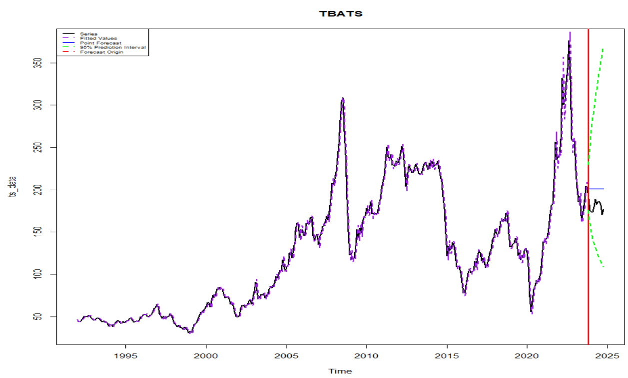
***Table** : Summary of Tbats Model*

BATS(0.001, {0,0}, -, -)
Call: tbats(y = clean_train_data)
Parameters Lambda: 0.0011 Alpha: 1.339849
Sigma: 0.06890861
AIC: 3785.99

The results of the analysis using the TBATS model provide a forecast for the future movements of the energy price index. The parameters of the model indicate that the lambda value is 0.001 and therefore the data do not need to be transformed. The low sigma value of the model (0.0689) indicates that the data fit the model properly. Moreover, the AIC value of the model (3785.99) is an important metric to evaluate its performance compared to other forecast models.



The graph shows the model's predictions and their fit with the historical data. The black line represents the historical data, the blue line represents the predicted values, and the blue shaded area represents the 95% confidence interval of the predictions. While the forecasts closely follow the data trend in the short term, there are widening forecast intervals, indicating increasing uncertainty in the future.



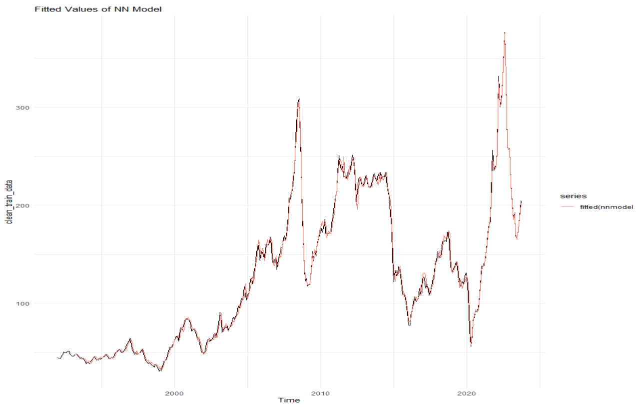
This graph shows a comparison of the forecasts made by the TBATS model with historical data and projections for the future. The black line represents the historical data, the blue line the predictions made by the model and the green line the 95% confidence interval. The general trend of the forecasts and their closeness to the historical data indicate a good fit of the model to the data.

The red line indicates the forecast origin. The forecast values following this line indicate that the model is successful in predicting the future movements of the energy price index, but uncertainties are increasing.

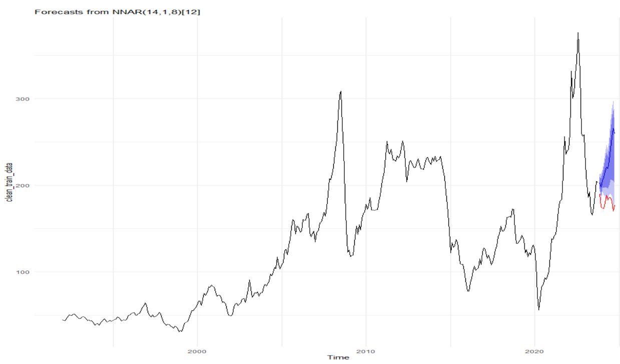
***Table 4**: Summary of NNETAR Model*

Series: clean_train_data
Model: NNAR(14,1,8)[12]
Average of 20 networks, each of which is a 14-8-1 network with 129 weights options were - linear output units
sigma^2 estimated as 26.1

The Neural Network Autoregressive (NNAR) model is applied for time series forecasting. The model is designed in the structure of NNAR(14,1,8)[12] and takes 14 past periods as input and produces a single forecast value. In the hidden layer, 8 neurons are used and a total of 129 weights are calculated. The model is built by averaging 20 different networks, each with the same structure, and trained on annual frequency (monthly) data. This model is preferred for time series with complex structure such as energy price index in order to capture nonlinear relationships and provide more accurate forecasts.



It shows how well the predictions produced by the Neural Network model fit the training data. As can be seen, the model successfully learned the past trends of the data and reproduced them with high accuracy. Especially in regions with high volatility, the model is able to produce results consistent with the data. This shows that the model is able to capture complex structural relationships and past trends. However, some high volatilities may not be fully consistent, which may point to the model's performance limitations for very long periods of complex data.

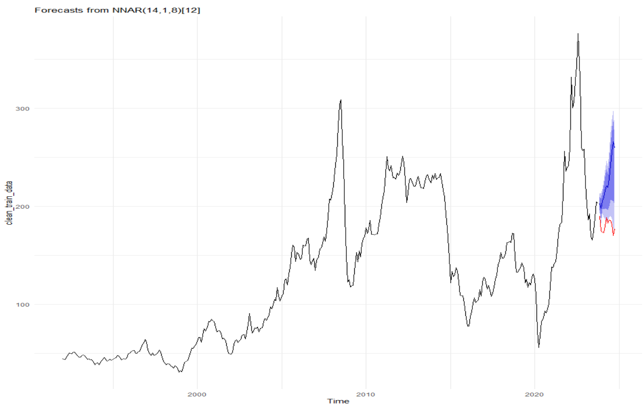


The plot presents 12-month forward forecasts generated by the NN model. Blue shaded areas represent 80% and 95% confidence intervals, indicating the level of uncertainty in the model's predictions. The predicted values are in line with the current trend; This indicates that the model effectively captures underlying patterns in the data. However, confidence intervals widen in regions where uncertainty is higher, revealing the difficulties of predicting fluctuations in the energy price index.

After NN forecast model using Prophet for the forecast.

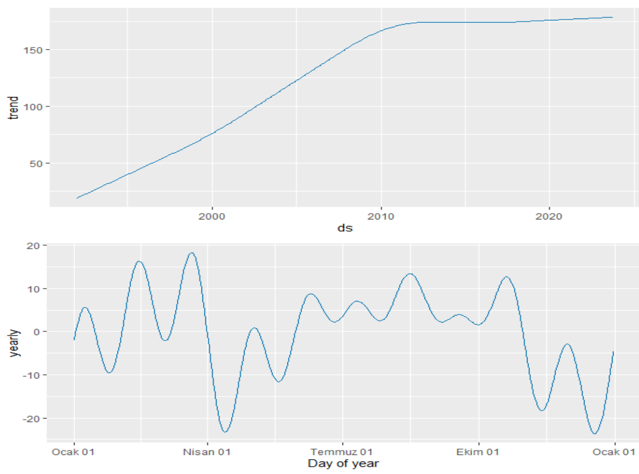
```
> accuracy(tail(forecast$yhat,12),test_data)
                ME     RMSE      MAE       MPE     MAPE      ACF1 Theil's U
Test set -2.484203 7.286095 5.928677 -1.502554 3.329729 0.3077425   1.00945
```

The performance evaluation of the Prophet model for the test set shows that the model is generally successful in forecasting. The mean error (ME) is low at -2.48, indicating a slight downward trend. The root mean squared error (RMSE) of 7.29 and the mean absolute error (MAE) of 5.93 indicate that the predictions are quite close to the actual values. Moreover, the percentage error rates (MPE: -1.52, MAPE: 3.33) are quite low and support the accuracy of the model. The first lag autocorrelation value (ACF1: 0.31) indicates that the residuals are independent and there is no autocorrelation problem in the estimation of the model. The Theil's U value is calculated as 1.00, indicating that the forecasts are generally acceptable. The Prophet model has demonstrated an effective forecasting performance with low error rates on a dynamic and uncertain data set such as the energy price index.



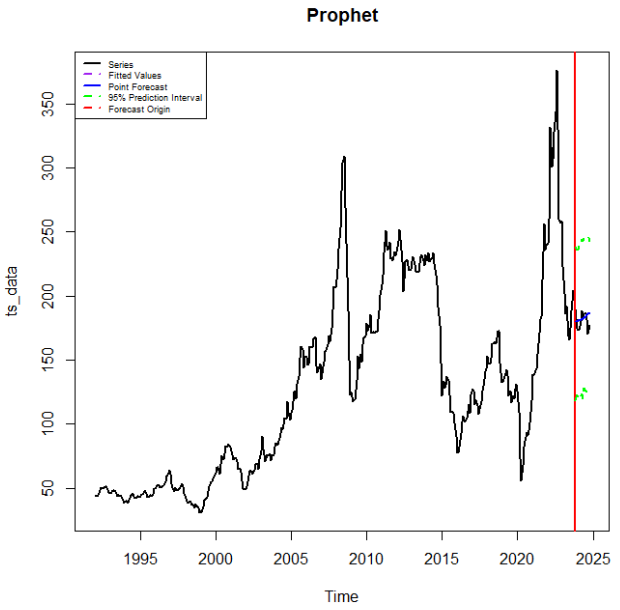
Graph 16: Forecast from NNAR

It shows the trend and forecast components obtained using the Prophet model. The top graph shows that the energy price index has increased steadily over time and stabilized after 2010. The lower graph reveals seasonal fluctuations, peaking in certain months, which can be associated with annual economic or climatic cycles. This analysis provides valuable insights for understanding recurring patterns in the dataset.

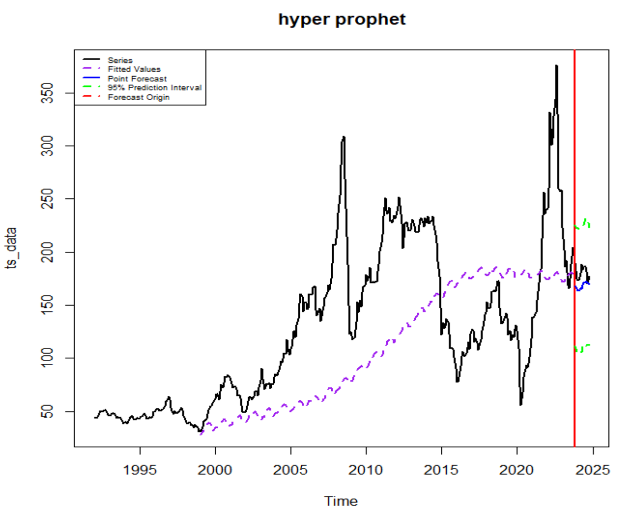


Graph 17: Prophet prediction plot.

Visualizes the forecasts produced by the Prophet model. The blue line represents the predicted values, while the blue shaded areas represent the 80% and 95% confidence intervals. The forecasts predict a gradual decline in the energy price index, following the general trend of historical data. However, the widening confidence intervals emphasize the increasing uncertainty in future forecasts and illustrate the difficulties of forecasting energy prices.



The graph compares the predicted and actual values of the energy price index.The black line shows the actual data and the forecast line shows the model's predictions.In general, the model captures the trend well, but there are slight deviations during extreme fluctuations.While this comparison demonstrates the success of the Prophet model in forecasting, it also points to areas of improvement that can be made to increase the accuracy of the model.



The “Hyper Prophet” model includes optimized hyperparameters to improve forecast accuracy.

The model's forecasts, indicated by the purple line, tend to be more conservative compared to historical data.This may be due to changes in seasonal or trend dynamics.This model clearly demonstrates the importance of hyperparameter tuning in improving forecast accuracy.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord “Format” pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.