# TIME SERIES ANALYSIS AND FORECASTING OF LOSSES OF RUSSIAN TANKS IN 2022 RUSSIAN-UKRAINE WAR

STAT 497 – APPLIED TIME SERIES ANALYSIS
DEPARTMENT OF STATISTICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

**BATUHAN SAYLAM-2429264**

January 2024

## 1.Abstract

The daily losses of Russian tanks in the Russian-Ukrainian War which is started in 2022 are predicted in this research by using a variety of forecasting models, including ARIMA, ETS, TBATS, NNETAR, and PROPHET. R-Studio is used in the study. Preprocessing is done on the data before forecasts are generated. Unit root checking, data cleansing, and outlier analysis are carried out. Following the handling of noisy data, the forecast models are fitted, and various comparison criteria are used to compare how well they perform on both train and test data.

## 2. Introduction

I worked with a dataset that details Russian prisoners of war, death toll, military wounded, and equipment losses during the 2022 conflict between Ukraine and Russia.

In order to visually identify trends and seasonality, I first began my research by interpreting time series plots related to Russian tank losses. I split the dataset into train and test. After that, detection of outliers was performed. Next, I experimented with techniques like Box-Cox transformation methods to obtain normal data. Based on this, ACF, PACF plots, KPSS and ADF or PP test results for zero mean, mean and trend situations and their interpretation. I performed unit root tests after trend was eliminated using differencing. Next, I examined the information table, ACF and PACF plots, the time series plot of a stationary series, and the ESACF. I then determined the appropriate ARMA, ARIMA, or SARIMA model. Once the possible models have been ranked, estimate the parameters using MLE, conditional or unconditional LSE. I compared information criteria of the models. Then, I conducted diagnostic checking. Lastly, I used techniques like ets, prophet, TBATS, nnetar, ARIMA/SARIMA, and others to perform forecasting. I then assessed the models' accuracy.

## 3. DATA DESCRIPTION AND PREPROCESSING

The Kaggle online data platform is where we discovered the dataset. The data will be updated every week. Each new record is the result of compiling information from previous records. This dataset details Russian military deaths, equipment losses, death tolls, and prisoners of war during the 2022 conflict between Ukraine and Russia.

Data Sources:

The General Staff of the Ukrainian Armed Forces and the Ukrainian Ministry of Defense are the primary data sources.

Information gathered from multiple fronts is combined. The high level of hostility complicates the calculation.

https://www.kaggle.com/datasets/piterfm/2022-ukraine-russian-war

Variables:

Aircraft, Helicopter, Tank, Armored Personnel Carrier (APC), Multiple Rocket Launcher (MRL), Field Artillery, Military Auto - none since 2022-05-01; combined with Fuel Tank into Vehicles and Fuel Tanks, Fuel Tank - none since 2022-05-01; combined with Military Auto into Vehicles and Fuel Tanks, Anti-aircraft warfare, Drone - UAV+RPA, Naval Ship - Warships, Boats, Anti-aircraft Warfare, Mobile SRBM System - not since 2022 05-01; combined with Cruise Missiles, Vehicles and Fuel Tanks - appear since 2022-05-01 as a sum of Fuel Tank and Military Auto, The Direction of Greatest Losses appears since 2022-04-25, while Cruise Missiles appear since 2022-05-01.

Acronyms:

MRL - Multiple Rocket Launcher,

APC - Armored Personnel Carrier,

SRBM - Short-range ballistic missile,

UAV - Unmanned Aerial Vehicle,

RPA - Remotely Piloted Vehicle.

Dataset History:

2022-03-02 - dataset is created (after 7 days of the War).

We selected Tank variable to examine.

The first six line of the data:

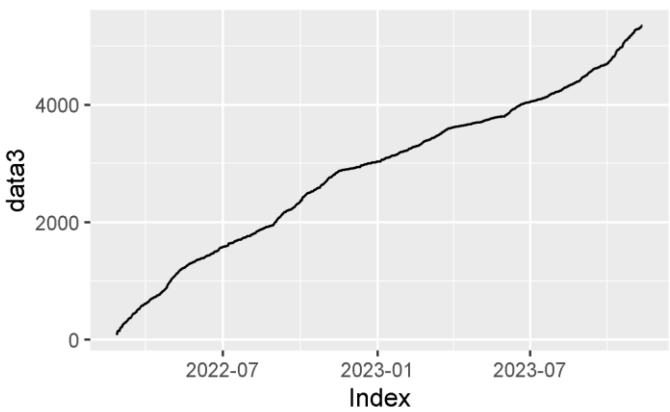
| | date <date> | tank <int> |
|---|---|---|
| 2022-02-25 | 2022-02-25 | 80 |
| 2022-02-26 | 2022-02-26 | 146 |
| 2022-02-27 | 2022-02-27 | 150 |
| 2022-02-28 | 2022-02-28 | 150 |
| 2022 03 01 | 2022 03 01 | 198 |
| 2022-03-02 | 2022-03-02 | 211 |

The structure of the data:

```
## 'data.frame':    626 obs. of  2 variables:
##  $ date: Date, format: "2022-02-25" "2022-02-26" ...
##  $ tank: int  80 146 150 150 198 211 217 251 269 285 ...
```

The summary of the data:

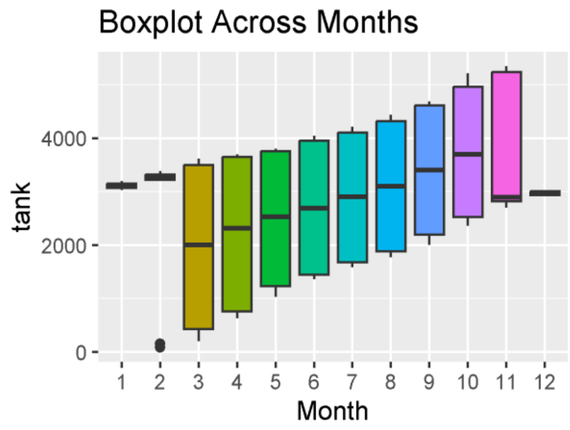
```
##       date                 tank
##  Min.   :2022-02-25   Min.   :  80
##  1st Qu.:2022-07-31   1st Qu.:1764
##  Median :2023-01-03   Median :3037
##  Mean   :2023-01-03   Mean   :2888
##  3rd Qu.:2023-06-08   3rd Qu.:3898
##  Max.   :2023-11-12   Max.   :5349
```




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

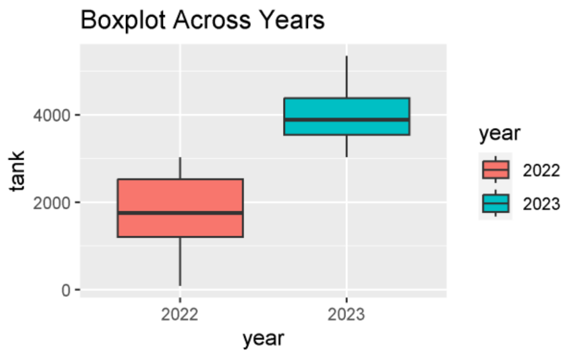
It is evident that the plot appears to be non-stationary. The mean changes. Time has an impact on it. Furthermore, the trend appears to be an increasing trend, and the linear trend shows us that the trend is the stochastic.

### 3.1 Explatory Data Analysis



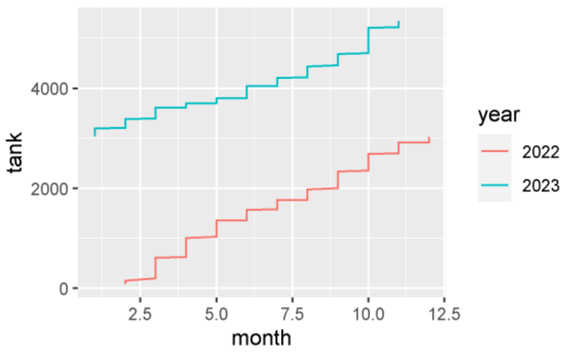

***Graph 2:*** *The boxplot across months*

We can observe that the process have no seasonal component but have an increasing trend every year because some of the median values for each month increase. We can also observe that there are some outliers.




**Graph 3:** The boxplot across years

According to the plot above, there appears to be a significant trend. throughout the series. Sales are rising annually.



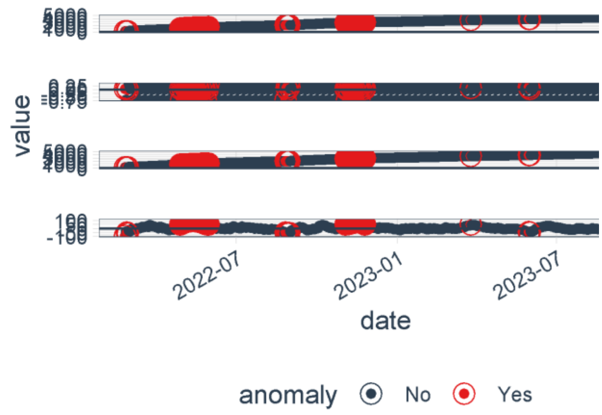

***Graph 4****: Losses of tank by year*

The time series plot for each year makes it clear that Russian tank losses have risen over time because the lines for each year do not overlap, indicating a growing tendency. Every year, the line's shape is not essentially the same. This suggests that the series does not exhibit seasonal patterns since it does not repeat itself year.

The data set is split into a test set and a train set at the start of the analysis. The last month is kept as the test set during this process.

### 3.2 Anomaly Detection

Next, we use stl decomposition to examine the data anomalies and demonstrate the anomalies in the series.

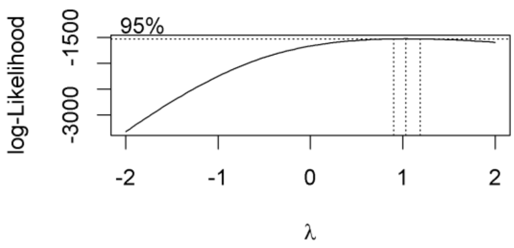



***Graph 5**: Anomaly Detection Plot*

The tsclean function eliminates anomalies in the train data and substitutes them with interpolated values.

### 3.3 Normalization

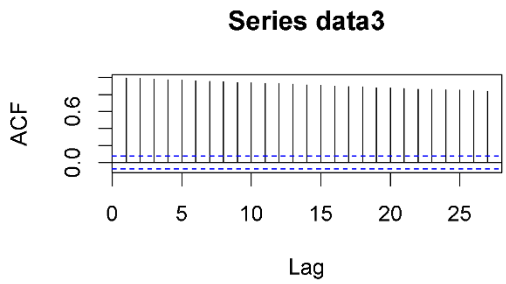
We need to transform train data if the data is not normally distributed.




***Graph 6**: Boxcox plot*

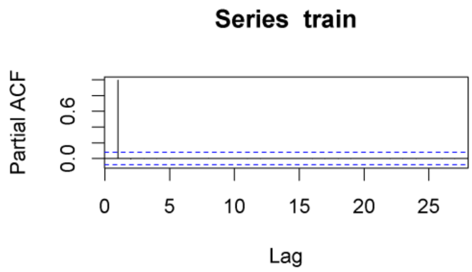
Since the vertical intervals include one, the lambda value is close to one or equals to one. Hence, we do not need transformation.

### 3.4 Model Suggestion




***Graph 7**: ACF Plot of Data Set*

The ACF plot exhibits a steady, linear decay, verifying the analysis presented. As a result, our process is said to as nonstationary.




***Graph 8**: PACF Plot of Data Set*

One can see that PACF stops working after the initial lag. However, by looking at the ACF and time series plot of the data, we already know that the process is non-stationary, hence there is no need to analyze the PACF plot of the data set.

```
        KPSS Test for Level Stationarity

data:  train
KPSS Level = 8.4175, Truncation lag parameter = 6, p-value = 0.01
```

According to the KPSS Test for Level Stationarity, the process is not stationary.

```
        KPSS Test for Trend Stationarity

data:  train
KPSS Trend = 1.8388, Truncation lag parameter = 6, p-value = 0.01
```

According to the KPSS Test for Trend Stationarity, the process has stochastic trend.

Since the mean value of the data is 2766.877, we need to apply ADF for regression model with an intercept (constant) but no time trend. Also, the test is applied with lag 1 since the PACF cuts off the first lag.

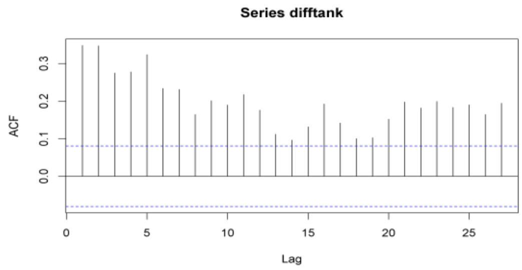
```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 1
  STATISTIC:
    Dickey-Fuller: -5.7602
  P VALUE:
    0.01
```

According to ADF test, since p-value is less than 0.05, the process has no unit root.

```
        Phillips-Perron Unit Root Test

data:  train
Dickey-Fuller Z(alpha) = -4.8213, Truncation lag parameter = 6, p-value = 0.8408
alternative hypothesis: stationary
```

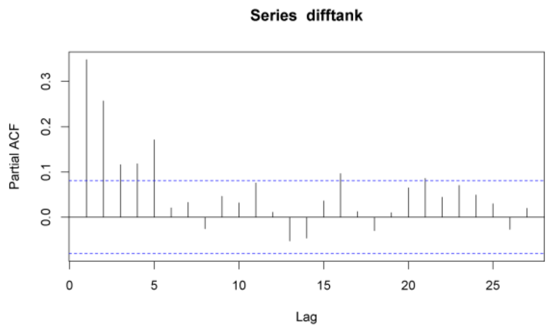
Unlike the ADF test, since p-value is greater than 0.05, the process has unit root.

There is contradictory between kpss test & pp-test and adf test so we need to check plots. Since ACF plot has linear decay, we need to differentiate the data.




***Graph 9****: ACF Plot of Differenced Data Set*

After differencing, the ACF plot exhibits that the process is still nonstationary.




***Graph 10****: PACF Plot of Differenced Data Set*

After the first five lags, PACF cuts off. But since we have already established that the process is non-stationary by looking at the ACF plot of differenced data, there is no need to interpret the PACF plot of the differenced data set.

```
        KPSS Test for Level Stationarity

data:  difftank
KPSS Level = 2.0185, Truncation lag parameter = 6, p-value = 0.01
```

According to the KPSS Test for Level Stationarity, since the p-value is less than 0.05, we reject the null hypothesis and conclude that the differenced dataset is not stationary.

Since the mean of the differenced time series (7.883838) is not zero or close to zero, we need to apply ADF test for regression model with an intercept (constant) but no time trend. Also, the test is applied with lag 5 since the PACF cuts off the first five lags.

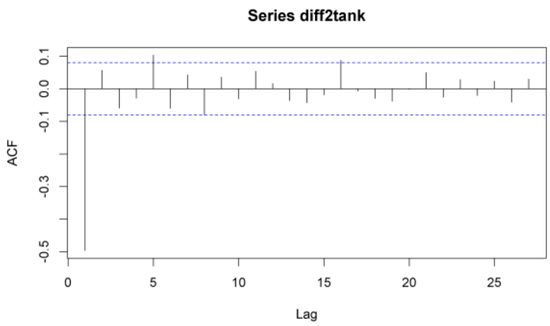
```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 5
  STATISTIC:
    Dickey-Fuller: -5.5029
  P VALUE:
    0.01
```

According to ADF test, since p-value is smaller than 0.05, the differenced process has no unit root.

```
	Phillips-Perron Unit Root Test

data:  difftank
Dickey-Fuller Z(alpha) = -554.04, Truncation lag parameter = 6, p-value = 0.01
alternative hypothesis: stationary
```

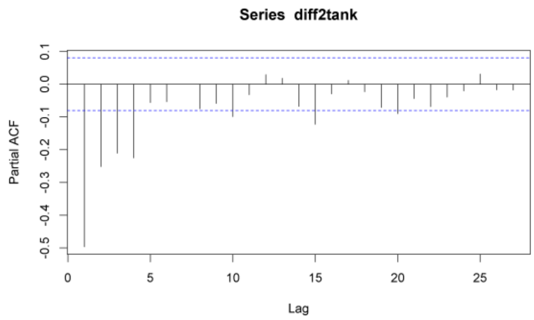
In addition to ADF test, the pp test suggests us since p-value is smaller than 0.05, the differenced process has no unit root.

There is contradictory between kpss test and pp-test & adf test so we need to check plots. Since ACF plot shows us the process is not stationary, we need to differentiate the data again.




***Graph 11****: ACF Plot of Two Differenced Data Set*

After two differencing, the ACF plot cuts off first lag or fourth lag.




***Graph 12****: PACF Plot of Two Differenced Data Set*

PACF cuts off after the first four lags. According to both plots, the process is stationary after two differencing.

```
	KPSS Test for Level Stationarity

data:  diff2tank
KPSS Level = 0.020304, Truncation lag parameter = 6, p-value = 0.1
```

The null hypothesis cannot be rejected because the p-value is greater than 0.05, leading us to the conclusion that the two differenced datasets are stationary.

Since the mean of the time series is zero, we need to apply ADF test for regression model with no intercept (constant) nor time trend.

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 4
  STATISTIC:
    Dickey-Fuller: -16.7239
  P VALUE:
    0.01
```

Since p-value is less than 0.05, the two differenced process has no unit root.

```
        Phillips-Perron Unit Root Test

data:  diff2tank
Dickey-Fuller Z(alpha) = -738.73, Truncation lag parameter = 6, p-value = 0.01
alternative hypothesis: stationary
```

In additon to ADF test, according to PP test, since p-value is less than 0.05, the two differenced process has no unit root. Hence, there is enough evidence to decide that the two differenced process is stationary.

EACF of the data:

```
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x o o o x o o o o o o  o  o  o
1 x x o o x o o o o o o  o  o  o
2 x x x o x o o o o o o  o  o  o
3 x x x x o o o o o o o  o  o  o
4 x x o x x o o o o o o  o  o  o
5 x x o x o o o o o o o  o  o  o
6 x x x x o o o o o o o  o  o  o
7 x x o x o o o x o o o  o  o  o
```

Also, the EACF, ACF and PACF plots shows us that the possible models of the data is ARIMA(4,2,1) and ARIMA(4,2,4).

```
Series: train
ARIMA(2,2,2)

Coefficients:
         ar1     ar2      ma1     ma2
      0.7691  0.0390  -1.5921  0.6035
s.e.  0.1523  0.0584   0.1478  0.1388

sigma^2 = 26.97:  log likelihood = -1817.17
AIC=3644.35   AICc=3644.45   BIC=3666.28
```

Also, auto arima suggests us that the model must be ARIMA(2,2,2).

We need to check overdifferencing by setting MA part as 1.

```
Series: train
ARIMA(4,2,1)

Coefficients:
         ar1     ar2     ar3     ar4      ma1
      0.1364  0.1347  0.0266  0.0509  -0.9516
s.e.  0.0467  0.0457  0.0450  0.0451   0.0208

sigma^2 = 27.09:  log likelihood = -1817.94
AIC=3647.88   AICc=3648.02   BIC=3674.19
```

Since the estimation of ma1 is not close to 1 or equals to 1, the process has no overdifferencing. Also, the AR4 of the model is not significant. Also, the model is one of the our possible models so let us decrease the order of AR.

```
Series: train
ARIMA(3,2,1)

Coefficients:
         ar1     ar2     ar3      ma1
      0.1276  0.1319  0.0243  -0.9419
s.e.  0.0470  0.0463  0.0455   0.0217

sigma^2 = 27.1:  log likelihood = -1818.58
AIC=3647.15   AICc=3647.26   BIC=3669.08
```

The AR3 of the model is not significant.

```
Series: train
ARIMA(2,2,1)

Coefficients:
         ar1     ar2      ma1
      0.1252  0.1297  -0.9370
s.e.  0.0474  0.0466   0.0212

sigma^2 = 27.07:  log likelihood = -1818.72
AIC=3645.44   AICc=3645.51   BIC=3662.98
```

The both AR2 and MA1 of the model is significant. Hence, we can check the model.

Now, we check the significance of the second possible model which is ARIMA(4,2,4).

```
Series: train
ARIMA(4,2,4)

Coefficients:
          ar1      ar2     ar3     ar4      ma1     ma2      ma3     ma4
      -0.0955  -0.2698  0.7856  0.0431  -0.7213  0.2082  -1.0641  0.6089
s.e.   0.1525   0.0989  0.1168  0.0582   0.1482  0.0158   0.0292  0.1395

sigma^2 = 26.78:  log likelihood = -1814.63
AIC=3647.26   AICc=3647.56   BIC=3686.72
```

The AR4 of the model is not significant. Let us decrease the order of MA, then decrease the order of AR respectively.

```
Series: train
ARIMA(4,2,3)

Coefficients:
         ar1      ar2     ar3     ar4     ma1     ma2      ma3
      0.1574  -0.8454  0.1250  0.1581  -0.972  1.0286  -0.9402
s.e.  0.0478   0.0461  0.0476  0.0463   0.023  0.0174   0.0222

sigma^2 = 26.49:  log likelihood = -1812.4
AIC=3640.8   AICc=3641.05   BIC=3675.88
```

The both AR4 and MA3 of the model is significant. Hence, we can check the model.

Let us decrease the order of AR, then decrease the order of MA respectively.

```
Series: train
ARIMA(3,2,4)

Coefficients:
         ar1     ar2     ar3      ma1     ma2      ma3     ma4
      0.2321  0.0154  0.4459  -1.0472  0.2096  -0.5496  0.4042
s.e.  0.6129  0.4715  0.2347   0.5938  0.8854   0.5089  0.1803

sigma^2 = 26.96:  log likelihood = -1815.59
AIC=3647.18   AICc=3647.43   BIC=3682.26
```

AR3 of the model is not significant.

```
Series: train
ARIMA(3,2,3)

Coefficients:
         ar1     ar2      ar3      ma1     ma2     ma3
      0.5009  0.3118  -0.0289  -1.3234  0.1270  0.2091
s.e.  0.4497  0.3236   0.0635   0.4479  0.6814  0.2606

sigma^2 = 27.03:  log likelihood = -1816.86
AIC=3647.72   AICc=3647.91   BIC=3678.42
```

Both AR3 and MA3 are not significant.

```
Series: train
ARIMA(2,2,3)

Coefficients:
Warning: NaNs üretimi         ar1      ar2       ma1       ma2      ma3
      0.0777  0.6394  -0.8882  -0.5889  0.4933
s.e.     NaN     NaN      NaN      NaN     NaN

sigma^2 = 27.03:  log likelihood = -1817.37
AIC=3646.74   AICc=3646.88   BIC=3673.05
```

Since the model produced NA values, we pass the model.

```
Series: train
ARIMA(2,2,2)

Coefficients:
         ar1     ar2      ma1     ma2
      0.7691  0.0390  -1.5921  0.6035
s.e.  0.1523  0.0584   0.1478  0.1388

sigma^2 = 26.97:  log likelihood = -1817.17
AIC=3644.35   AICc=3644.45   BIC=3666.28
```

The model is also suggested by auto arima but the AR2 of the model is not significant.


```
Series: train
ARIMA(2,2,1)

Coefficients:
         ar1     ar2      ma1
      0.1252  0.1297  -0.9370
s.e.  0.0474  0.0466   0.0212

sigma^2 = 27.07:  log likelihood = -1818.72
AIC=3645.44   AICc=3645.51   BIC=3662.98
```

The both AR2 and MA1 of the model is significant. Hence, we can check the model.

Also, we need to check ARIMA(4,1,5) and ARIMA(1,1,5).

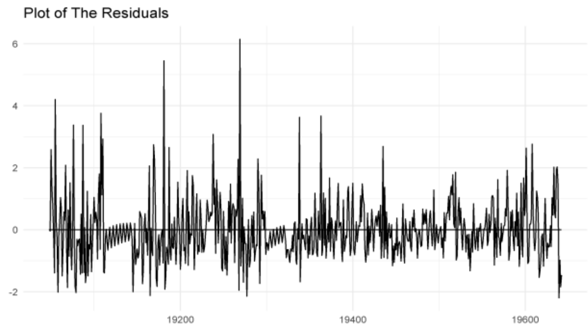
```
Series: train
ARIMA(4,1,5)

Coefficients:
         ar1      ar2     ar3      ar4      ma1     ma2      ma3     ma4      ma5
      1.3234  -0.9494  1.3522  -0.7272  -1.1291  0.9044  -1.3126  0.6042  -0.0206
s.e.  0.1336   0.0812  0.0775   0.1268   0.1400  0.0767   0.0837  0.1243   0.0458

sigma^2 = 26.82:  log likelihood = -1817.26
AIC=3654.52   AICc=3654.9   BIC=3698.39
```

```
Series: train
ARIMA(1,1,5)

Coefficients:
         ar1      ma1     ma2      ma3     ma4      ma5
      0.9989  -0.8178  0.0278  -0.1118  0.0234  -0.0292
s.e.  0.0015   0.0431  0.0552   0.0598  0.0588   0.0449

sigma^2 = 27.23:  log likelihood = -1822.66
AIC=3659.32   AICc=3659.51   BIC=3690.02
```

The MA5 values of the both model are not significant. As a Result, we check ARIMA(4,2,3) and ARIMA(2,2,1).

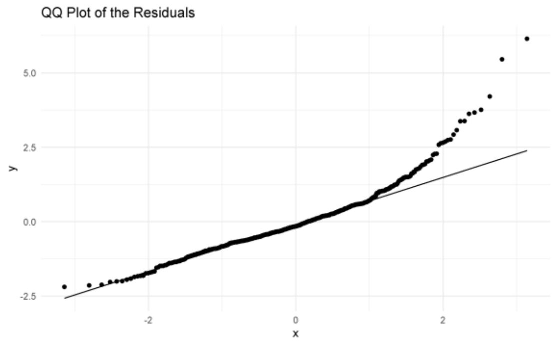
### 3.5 Diagnostic Checking

Since the ARIMA(4,2,3) model has lowest AIC, we check the model firstly.



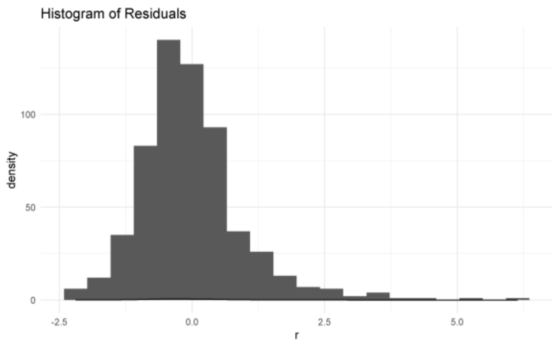

***Graph 13****: Plot of the residuals*

Since residuals are dispersed around zero, zero mean can be inferred.




***Graph 14****: QQ-Plot of the residuals*

The majority of the model's residuals are not located on a straight line on the QQ Plot, indicating that the residuals are right skewed and not normally distributed.




***Graph 15****: Histogram of the residuals*

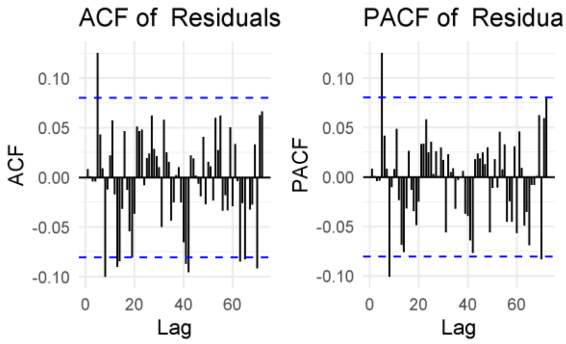
In addition to QQ-plot, the histogram of the residuals shows that the distribution is right skewed.,

```
        Shapiro-Wilk normality test

data:  r
W = 0.91643, p-value < 2.2e-16
```

Also, Shapiro-Wilk test verifies that the residuals do not have normal distribution so there exist enough evidence to decide that the normality assumption is not satisfied.

```
        Ljung-Box test

data:  Residuals from ARIMA(4,2,3)
Q* = 17.197, df = 3, p-value = 0.0006437

Model df: 7.   Total lags used: 10
```

Also, according to Ljung- Box test, the p-value is lower than 0.05 so we need to reject null hypothesis which is that the model does not exhibit lack of fit. As a result, the model exhibits lack of fit.

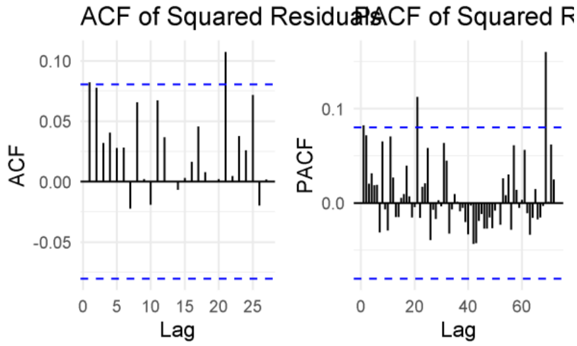



***Graph 16***: *ACF and PACF Plots of Residuals*

The residuals show no correlation since almost every spike in the ACF is located in the WN band. Let's apply a formal test to be sure.

```
        Breusch-Godfrey test for serial
        correlation of order up to 10

data:  m
LM test = 17.222, df = 10, p-value =
0.06959
```

The Breusch-Godfrey Test results show that we have a 95% confidence level that the model's residuals are uncorrelated since the p value is greater than alpha.



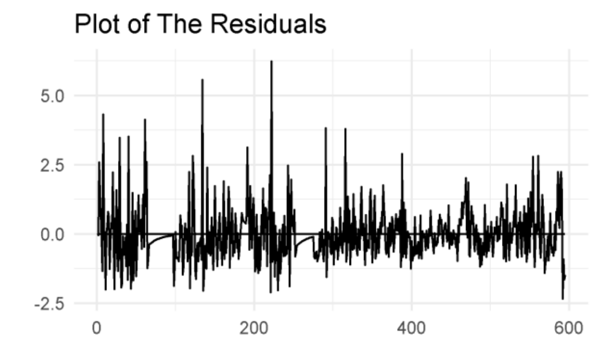

***Graph 17***: *ACF and PACF Plots of Squared Residuals*

Also, the almost every spike in both plot are between WN band so the heteroscedasticity of the residuals assumption is satisfied. However, to be sure, we need to conduct formal test.

```
        ARCH LM-test; Null hypothesis: no ARCH
        effects

data:  r
Chi-squared = 17.147, df = 12, p-value =
0.1441
```

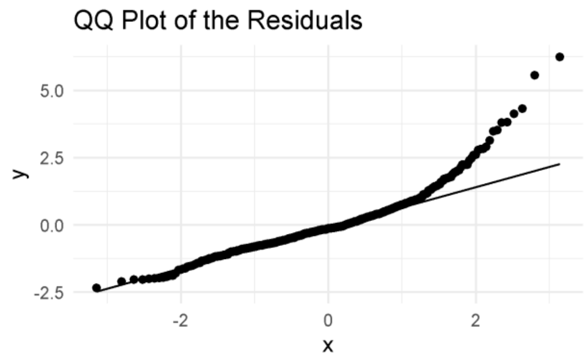
We are unable to reject H0 since p values are bigger than α. Consequently, we may say that ARCH effects are not present.

Hence, the model satisfied the assumptions except the normality assumption and lack of fit so we continue with checking the another possible model which is the ARIMA(2,2,1).



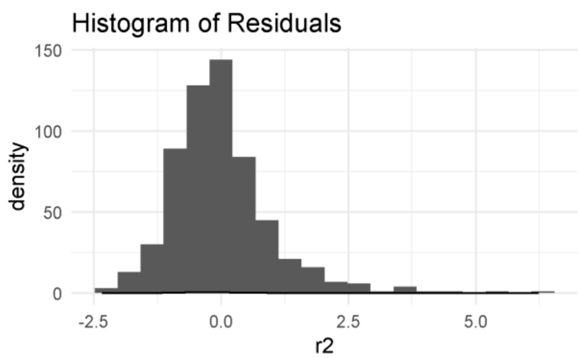

***Graph 18**: Plot of the residuals*

Since residuals are dispersed around zero, zero mean can be inferred.




***Graph 19**: QQ-Plot of the residuals*

The majority of the model's residuals are not located on a straight line on the QQ Plot, indicating that the residuals are right skewed and not normally distributed.




***Graph 20**: Histogram of the residuals*

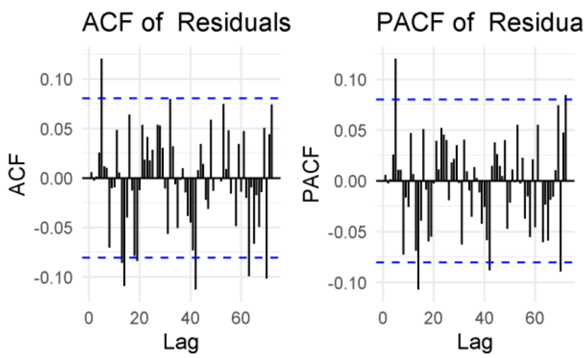
In addition to QQ-plot, the histogram of the residuals shows that the distribution is right skewed.,

```
	Shapiro-Wilk normality test

data:  r2
W = 0.90164, p-value < 2.2e-16
```

Also, Shapiro-Wilk test verifies that the residuals do not have normal distribution so there exist enough evidence to decide that the normality assumption is not satisfied.

```
	Ljung-Box test

data:  Residuals from ARIMA(2,2,1)
Q* = 12.457, df = 7, p-value = 0.08649

Model df: 3.   Total lags used: 10
```

Additionally, the model does not show a lack of fit, according to the Ljung-Box test, where the p-value is greater than 0.05, preventing us from rejecting the null hypothesis. Consequently, there is no evidence of lack of fit in the model.

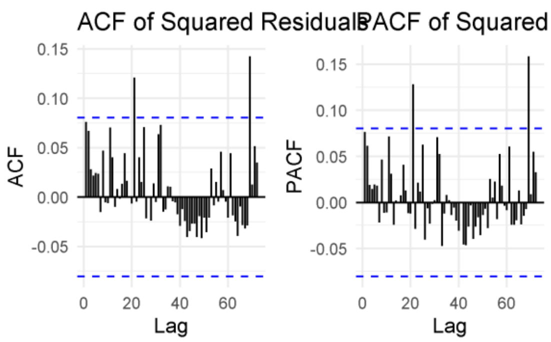



***Graph 21**: ACF and PACF Plots of Residuals*

The residuals show no correlation since almost every spike in the ACF is located in the WN band. Let's apply a formal test to be sure.

```
        Breusch-Godfrey test for serial
        correlation of order up to 10

data:  m2
LM test = 13.5, df = 10, p-value = 0.1971
```

The Breusch-Godfrey Test results show that we have a 95% confidence level that the model's residuals are uncorrelated since the p value is greater than alpha.




***Graph 22**: ACF and PACF Plots of Squared Residuals*

Also, the almost every spike in both plot are between WN band so the heteroscedasticity of the residuals assumption is satisfied. However, to be sure, we need to conduct formal test.

```
        ARCH LM-test; Null hypothesis: no ARCH
        effects

data:  r2
Chi-squared = 13.382, df = 12, p-value =
0.3419
```

We are unable to reject H0 since p values are bigger than α. Consequently, we may say that ARCH effects are not present.

Therefore, the model satisfied the assumptions except the normality assumption. Hence, the second model is better than the first model since it satisfied more assumption than the first one. As result, we are going to use the second model which is ARIMA(2,2,1).

```
Forecast method: ARIMA(2,2,1)

Model Information:
Series: train
ARIMA(2,2,1)

Coefficients:
         ar1     ar2      ma1
      0.1252  0.1297  -0.9370
s.e.  0.0474  0.0466   0.0212

sigma^2 = 27.07:  log likelihood = -1818.72
AIC=3645.44   AICc=3645.51   BIC=3662.98
```

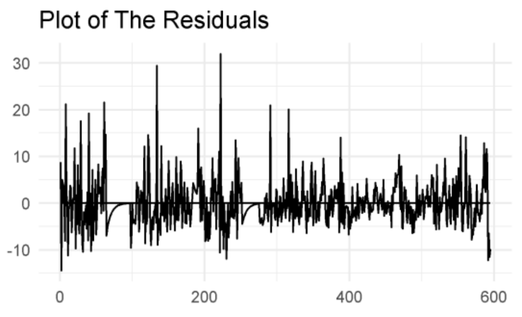
### 3.6 Forecasting

In addition ro ARIMA model, we are going to use ETS, TBATS, NNETAR, and PROPHET models. Also, we will check the assumptions of them and conduct parameter tuning for PROPHET.

We are going to begin with ETS model.

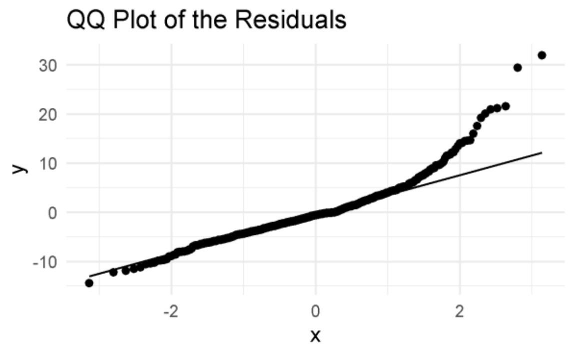
```
ETS(A,A,N)

Call:
 ets(y = train, model = "ZZZ")

  Smoothing parameters:
    alpha = 0.9999
    beta  = 0.1456

  Initial states:
    l = 124.1091
    b = 13.1785

  sigma:  5.2568

     AIC     AICc      BIC
5782.002 5782.104 5803.945
```

According to summary, the model has additive error, no trend and no seasonality. So the model is Holt's linear method with additive errors.



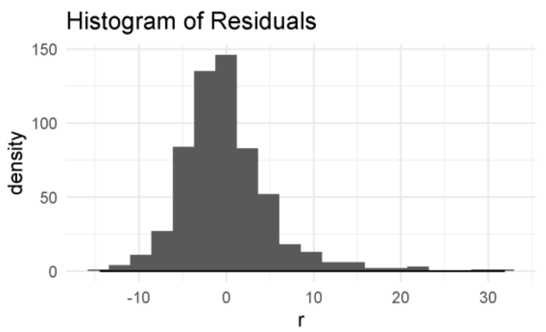

***Graph 23**: Plot of the residuals*

Since residuals are dispersed around zero, zero mean can be inferred.




***Graph 24**: QQ-Plot of the residuals*

The majority of the model's residuals are not located on a straight line on the QQ Plot, indicating that the residuals are right skewed and not normally distributed.




***Graph 25**: Histogram of the residuals*

In addition to QQ-plot, the histogram of the residuals shows that the distribution is right skewed.,

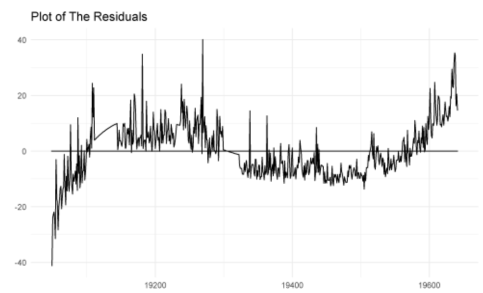
```
        Shapiro-Wilk normality test

data:  r
W = 0.91062, p-value < 2.2e-16
```

Also, Shapiro-Wilk test shows that the residuals do not have normal distribution so there exist enough evidence to decide that the normality assumption is not satisfied.

Therefore, the model did not satisfy the normality assumption. Now, we will continue with NNETAR.

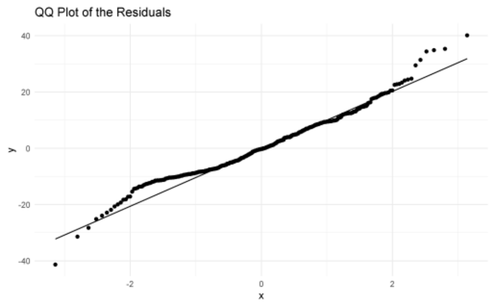
```
"NNAR(1,1)"

Average of 20 networks, each of which is
a 1-1-1 network with 4 weights
options were - linear output units
```

The model is NNAR(1,1). A single neuron in the hidden layer and the final observation, Y_(t-1) as input are used in this neural network to predict the output, Y_(t).



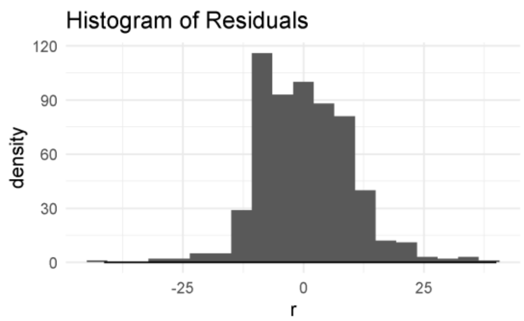

***Graph 26**: Plot of the residuals*

Since residuals are not dispersed around zero, zero mean cannot be inferred.




***Graph 27***: *QQ-Plot of the residuals*

The majority of the model's residuals are not located on a straight line on the QQ Plot, indicating that the residuals are not normally distributed.




***Graph 28***: *Histogram of the residuals*

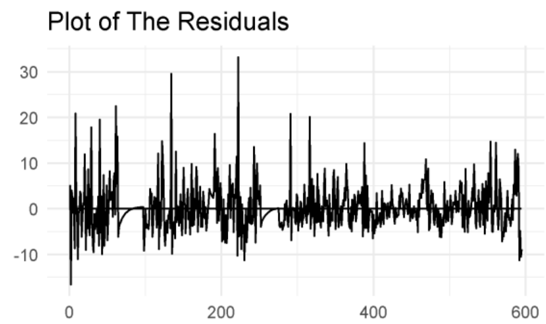
In addition to QQ-plot, the histogram of the residuals shows that the distribution is not normal.,

```
        Shapiro-Wilk normality test

data:  r
W = 0.97479, p-value = 1.391e-08
```

Also, Shapiro-Wilk test verifies that the residuals do not have normal distribution so there exist enough evidence to decide that the normality assumption is not satisfied.

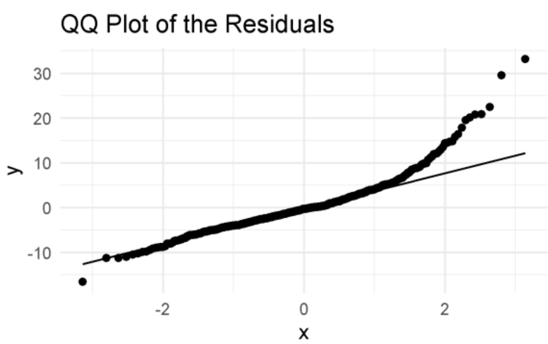
Therefore, the model did not satisfy the normality assumption. Now, we continue with TBATS model.

```
BATS(1, {0,0}, 0.995, -)

Call: tbats(y = ts(train))

Parameters
  Alpha: 1.051277
  Beta: 0.122806
  Damping Parameter: 0.994958

Seed States:
          [,1]
[1,] 125.04875
[2,]  15.85002

Sigma: 5.21842
AIC: 5777.306
```



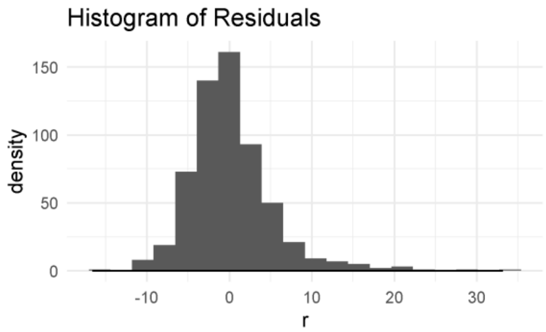

***Graph 29***: *Plot of the residuals*

Since residuals are dispersed around zero, zero mean can be inferred.



*__Graph 30__: QQ-Plot of the residuals*

The majority of the model's residuals are not located on a straight line on the QQ Plot, indicating that the residuals are right skewed and not normally distributed.




*__Graph 31__: Histogram of the residuals*

In addition to QQ-plot, the histogram of the residuals shows that the distribution is right skewed.,

```
        Shapiro-Wilk normality test

data:  r
W = 0.90448, p-value < 2.2e-16
```

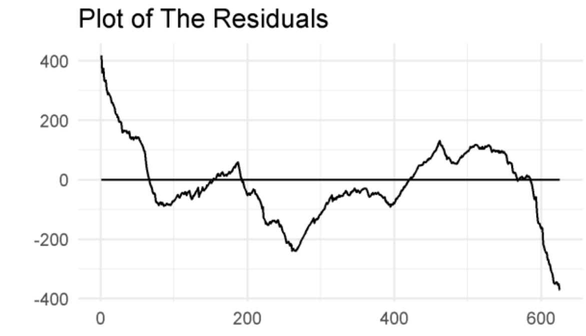
Also, Shapiro-Wilk test shows that the residuals do not have normal distribution so there exist enough evidence to decide that the normality assumption is not satisfied.

Therefore, the model did not satisfy the normality assumption. We continue with PROPHET.

Firstly, we conduct parameter tuning. For the tuning, changepoint prior scale parameter is 0.001,0.01,0.1 and 0.5, and changepoint range is 0.8,0.85,0.90 and 0.95.

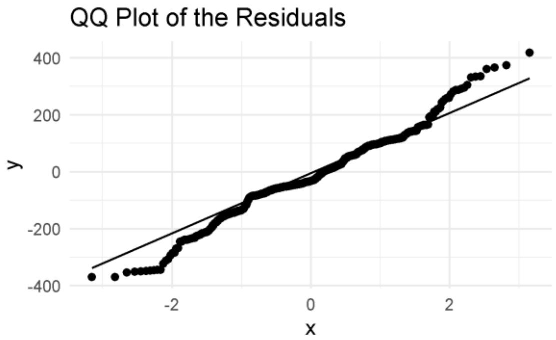
| | ME | RMSE | MAE | MPE | MAPE |
|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| 0.001-0.8 | 4534.172 | 4534.569 | 4534.172 | 88.00939 | 88.00939 |
| 0.001-0.85 | 4557.406 | 4557.785 | 4557.406 | 88.46109 | 88.46109 |
| 0.001-0.9 | 4531.847 | 4532.249 | 4531.847 | 87.96403 | 87.96403 |
| 0.001-0.95 | 4595.92 | 4596.266 | 4595.92 | 89.21016 | 89.21016 |
| 0.01-0.8 | 4794.007 | 4794.066 | 4794.007 | 93.07424 | 93.07424 |
| 0.01-0.85 | 4794.718 | 4794.776 | 4794.718 | 93.08809 | 93.08809 |
| 0.01-0.9 | 4795.329 | 4795.371 | 4795.329 | 93.10193 | 93.10193 |
| 0.01-0.95 | 4794.353 | 4794.397 | 4794.353 | 93.08295 | 93.08295 |
| 0.1-0.8 | 4792.21 | 4792.223 | 4792.21 | 93.04694 | 93.04694 |
| 0.1-0.85 | 4792.386 | 4792.397 | 4792.386 | 93.05119 | 93.05119 |
| 0.1-0.9 | 4791.264 | 4791.274 | 4791.264 | 93.02981 | 93.02981 |
| 0.1-0.95 | 4790.621 | 4790.633 | 4790.621 | 93.01801 | 93.01801 |
| 0.5-0.8 | 4792.126 | 4792.139 | 4792.126 | 93.04606 | 93.04606 |
| 0.5-0.85 | 4791.652 | 4791.663 | 4791.652 | 93.03733 | 93.03733 |
| 0.5-0.9 | 4790.899 | 4790.908 | 4790.899 | 93.02324 | 93.02324 |
| 0.5-0.95 | 4789.96 | 4789.972 | 4789.96 | 93.00599 | 93.00599 |

As a result of the tuning, the changepoint prior scale parameter is 0.001 and changepoint range is 0.9. Now, we check the assumptions.



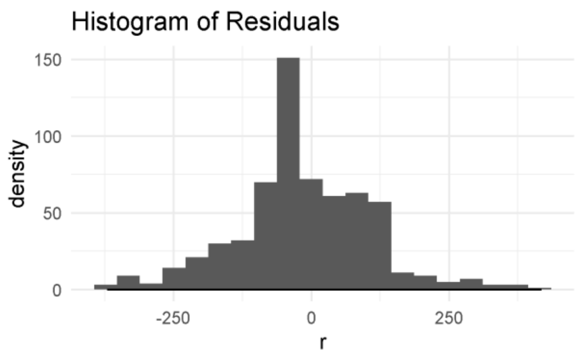

***Graph 32****: Plot of the residuals*

Since residuals are not dispersed around zero, the mean is not equal to zero.




***Graph 33****: QQ-Plot of the residuals*

The majority of the model's residuals are not located on a straight line on the QQ Plot, indicating that the residuals are not normally distributed.




***Graph 34****: Histogram of the residuals*

In addition to QQ-plot, the histogram of the residuals shows that the distribution is not normally distributed.

```
        Shapiro-Wilk normality test

data:  r
W = 0.98078, p-value = 2.523e-07
```

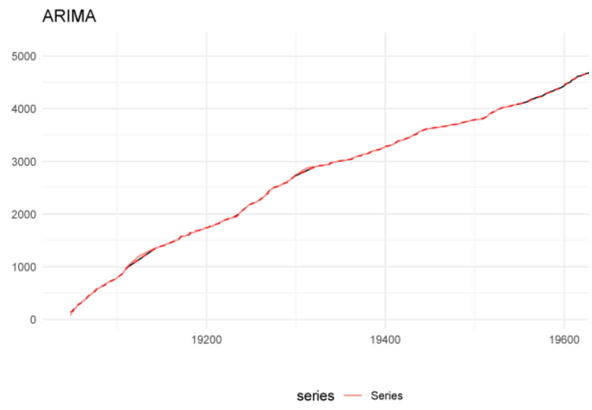
Also, Shapiro-Wilk test verifies that the residuals do not have normal distribution so there exist enough evidence to decide that the normality assumption is not satisfied.

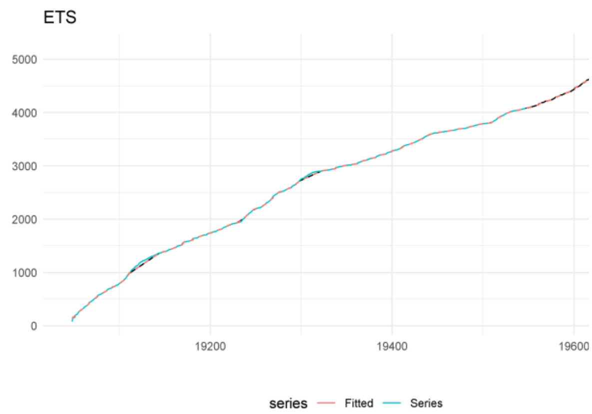
As a result, any model can not satisfied the normality assumption.

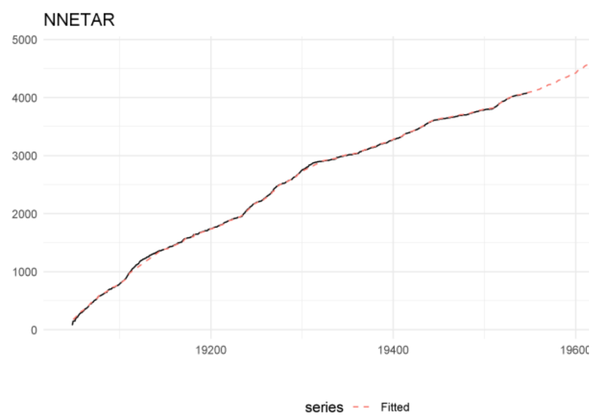
Following model fitting, we compute the accuracy of each method's forecast values by utilizing the forecast function. The models' accuracy is listed below.

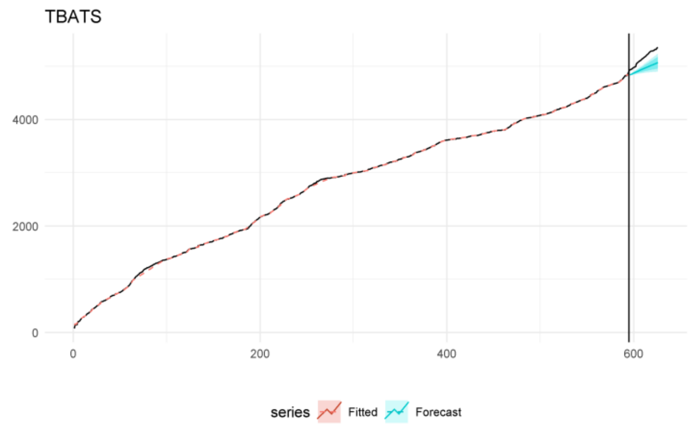
```
[1] "ARIMA(2,2,1)"
                      ME       RMSE        MAE       MPE      MAPE      MASE        ACF1
Training set  -0.05560763   5.180923   3.655655 0.0164543 0.2503019  0.463294 0.005892095
Test set     190.34392895 198.410056 190.343929 3.6681379 3.6681379 24.122956          NA
[1] "ETS"
                     ME       RMSE        MAE          MPE      MAPE       MASE       ACF1
Training set  -0.0549773   5.239054   3.688071 -0.001526009 0.2651306  0.4674022 0.03969414
Test set     189.9388230 198.572654 189.938823  3.659281250 3.6592812 24.0716153         NA
[1] "NNETAR"
                     ME       RMSE        MAE        MPE     MAPE      MASE      ACF1
Training set   0.2493625   9.524709   7.468734 -0.2165987 0.585039  0.946539 0.7388282
Test set     480.2073608 520.639393 480.207361  9.2234421 9.223442 60.858368        NA
[1] "TBATS"
                    ME      RMSE        MAE         MPE      MAPE       MASE        ACF1
Training set   0.213985   5.21842   3.609074 0.002092469 0.2553715  0.4573906 0.003401719
Test set     198.887362 208.98148 198.887362 3.829788413 3.8297884 25.2056952          NA
[1] "Prophet"
               ME     RMSE      MAE      MPE     MAPE
Test set 4531.847 4532.249 4531.847 87.96403 87.96403
```

The results demonstrate that the ARIMA model surpasses the other methods according to the comparison of the train set with respect to metrics like ME, RMSE, and MAPE, while the ETS model surpasses the other methods in the test set with respect to all measures except RMSE. Furthermore, compared to other models for the test set, ARIMA has the second-best forecasting performance. As can be seen from the preceding tables, the prophet model has the lowest forecasting accuracy. The following graphs provide more insight into the models' forecasting performance.

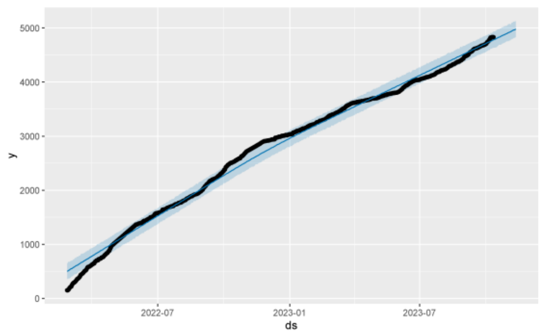



And the plot of Prophet forecasting:




## 4. DISCUSSION AND CONCLUSION

After the data was cleaned and divided, the first step in this study was to check if the series was stationary. However, it was found that this criteria was not achieved because the series exhibited a stochastic trend after examining the time series, ACF&PACF plots, and the outcomes of the KPSS, ADF, and PP tests. Twice, differencing was applied to solve this problem. Following the stationary of the process, Two possible models were offered utilizing particular techniques including ACF&PACF displays and the ESACF table. Then, using t statistics and an AIC comparison, the model that performed best with the lowest and most significant parameters was selected from these tentative models.

Diagnostic checking are applied to residuals after the optimal model's fitting. The non-normality of errors is appeared. Nonetheless, formal testing and visual inspection methods show that errors are uncorrelated and homoscedastic. Four other forecasting techniques were applied in addition to best ARIMA models , and forecasts were generated. In summary, ARIMA outperforms other models according to ME, RMSE, and MAPE metrics in terms of modeling series, and ETS outperforms other models in terms of future value prediction.

## 5. REFERENCES


*ODTUCLASS 2023-2024 FALL: Applied Time Series Analysis (n.d.).* https://odtuclass2023f.metu.edu.tr/course/view.php?id=841