**Personal Saving Rate Time Series Analysis**

Boran Aslan

Statistics, Middle East Technical University  
Türkiye Ankara/Çankaya

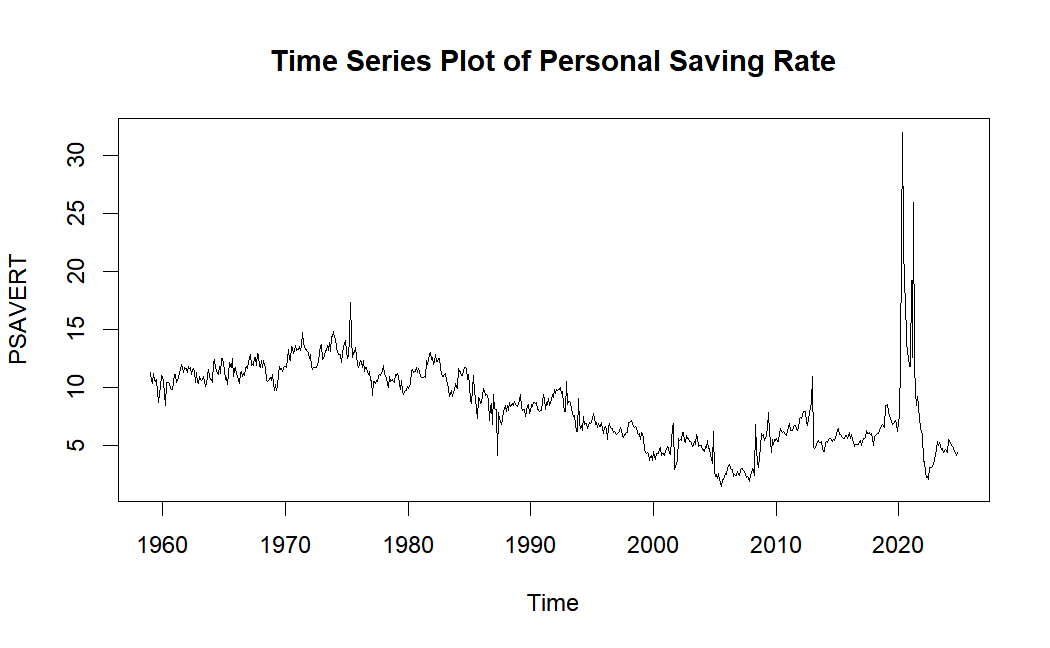
aslan.boran@metu.edu.tr

*Abstract— Using time series techniques, this study examines the US Personal Saving Rate (PSAVERT). One important economic metric is PSAVERT, which stands for the percentage of disposable personal income that households save. Significant variations in PSAVERT are seen in the data, with a notable spike occurring during the 2020–2021 pandemic. Time series analysis was carried out, encompassing diagnostic checks, model selection, and stationarity tests.  
Alternative forecasting techniques including neural networks (nnetar, TBATS, Prophet) are also examined in the study.  
The results indicate that neural network models especially nnetar models perform better in terms of predicting accuracy, even though ARIMA models offer a respectable fit.*

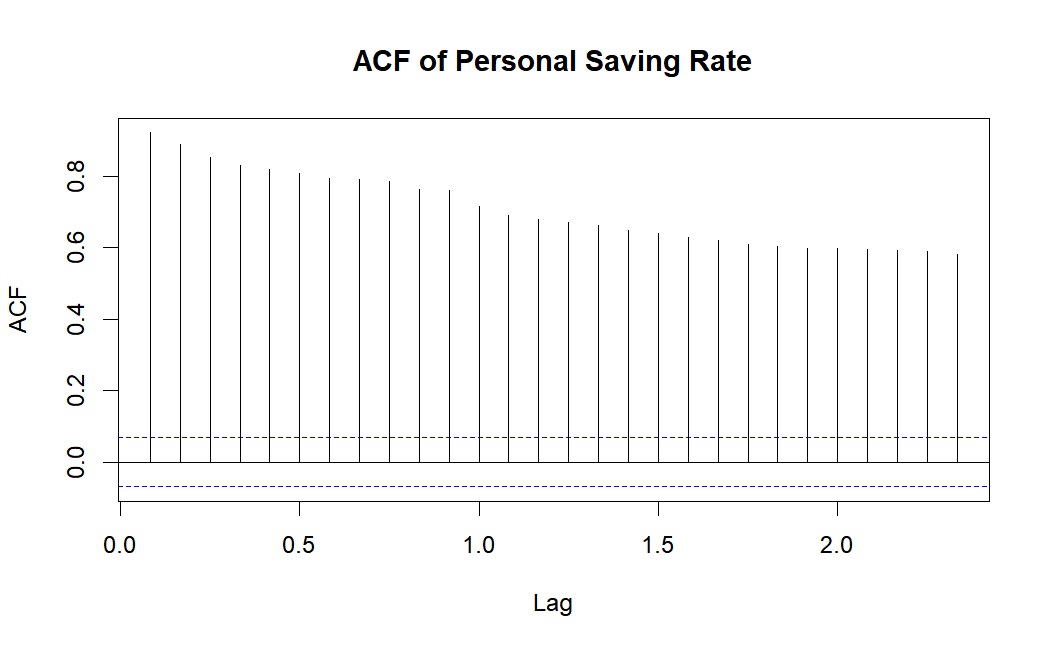
1. *introductıon*

*Our aims is to analyze the Personal Saving Rate data set, using time series. Personal Saving Rate(PSAVERT) is taken as the proportion of the personal income that households holds rather than spend. It is a very essential indicator that will help policymakers, investors, and businesses. My source is; Federal Reserve Economic Data | FRED | St. Louis Fed from here I chose Personal Saving Rate. Percentage Saving rate is between 1.4 and 32, mean is 8.43. Data starts from 1959 and ends in late 2024. There is no missing value. It is a monthly data set*

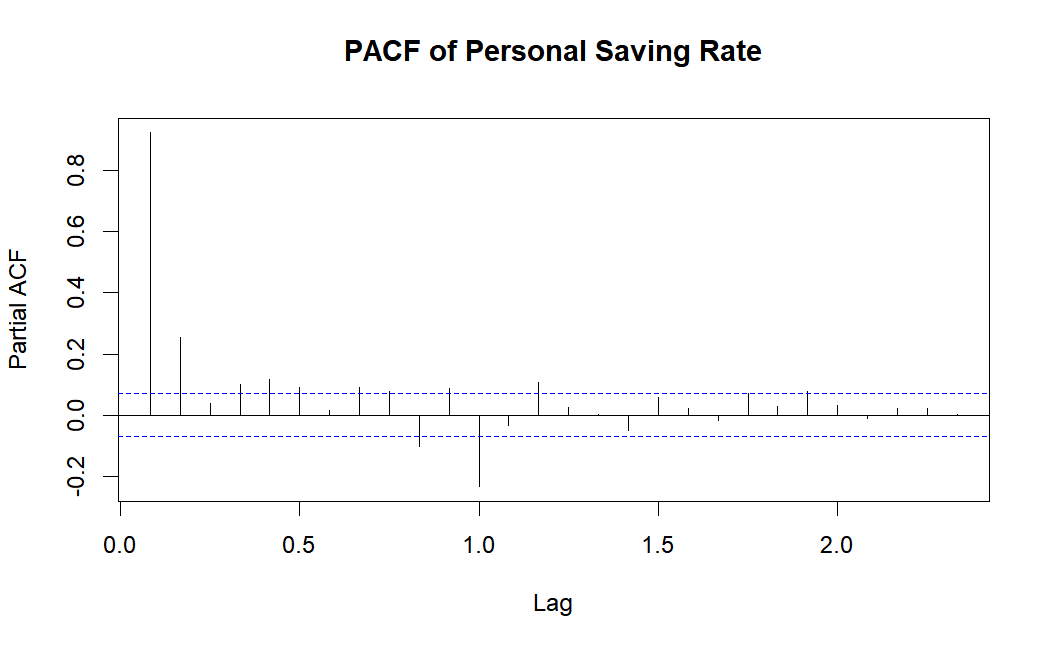
1. *Plot ınterpretatıon*

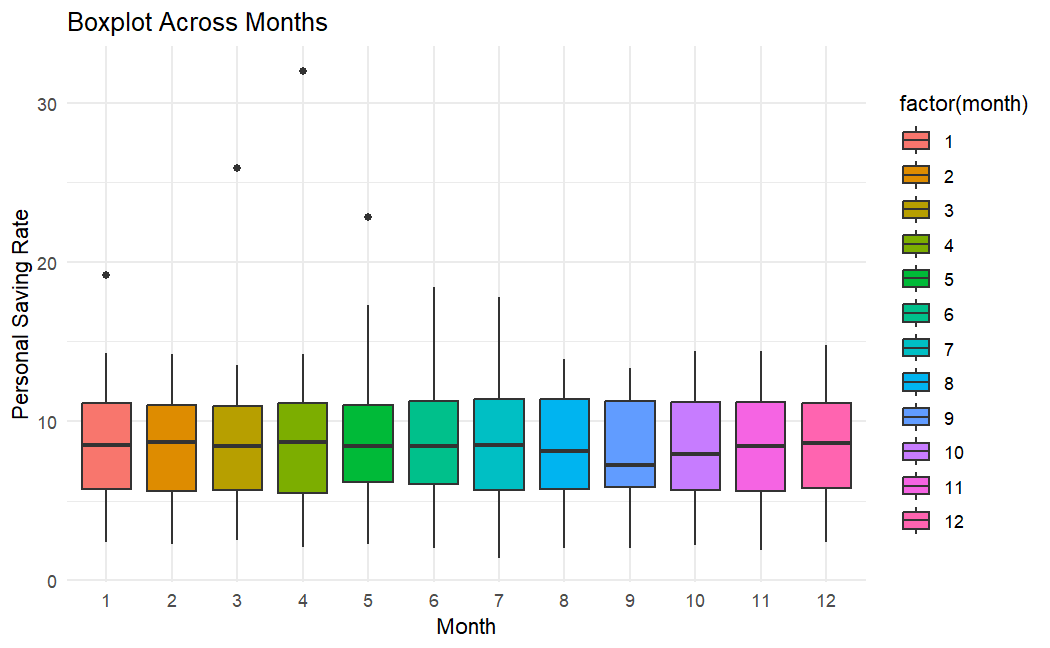
**

*The x-axis represent the time, y-axis represent the percentage of the personal saving rate. In the plot there is a huge spike in 2020, which is because of the pandemic. Data prior to 1975 shows an upward trend. Subsequently, there was a decline until approximately 2008. From 2008 to 2021, an increase was observed, after the very high spikes in 2020-2021, another decline comes.*

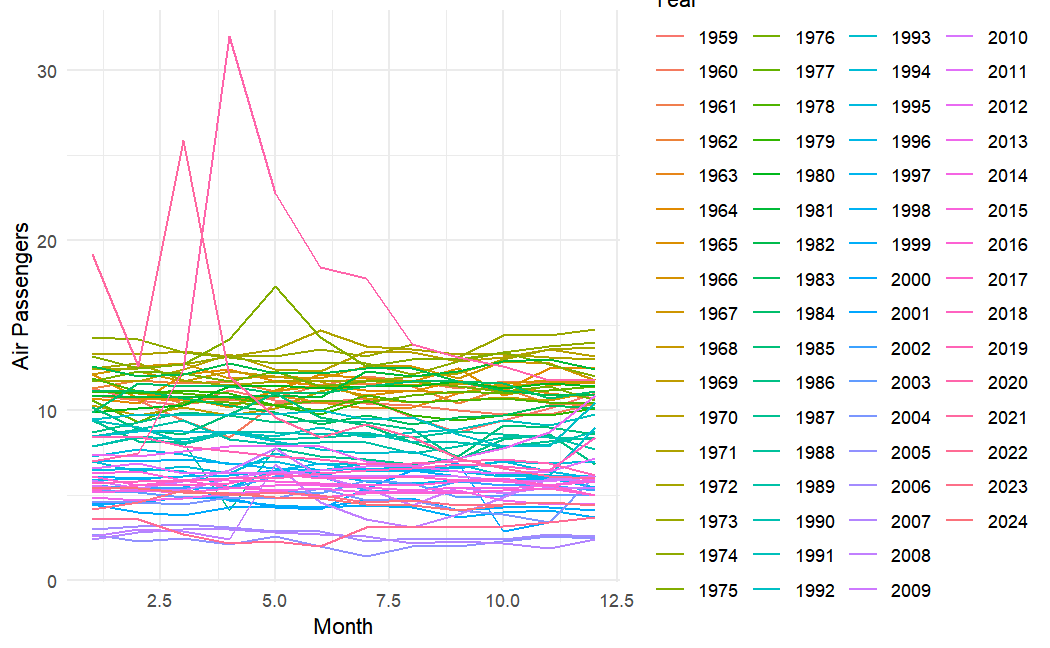
**

*There is a slow linear decay in ACF so it is not stationary but let's check with the tests as well.*

**

**

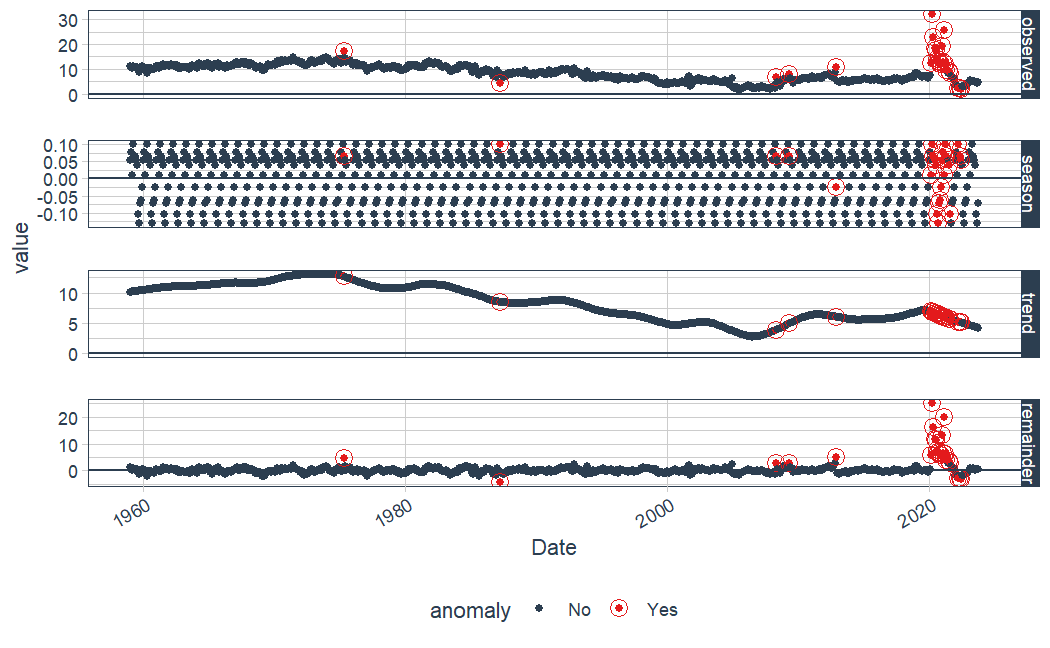
*As we can see, the means of the months are very similar to each other. Seasonality is not observed from the box-plots.*

**

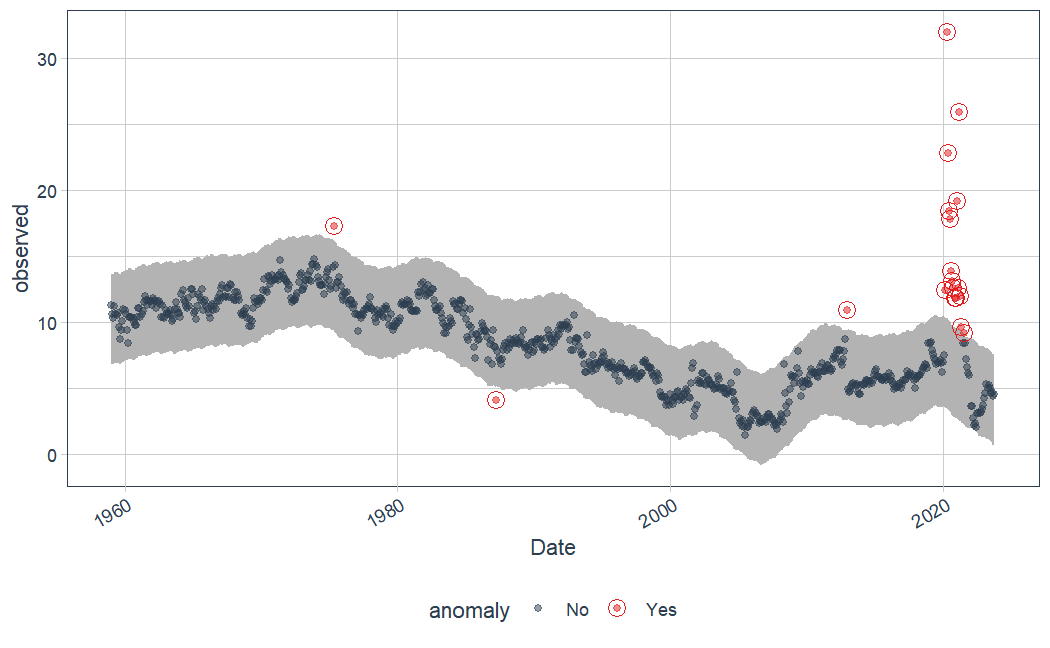
*We can see from this plots, pandemic years 2020 and 2021 differs from the rest of the years.*

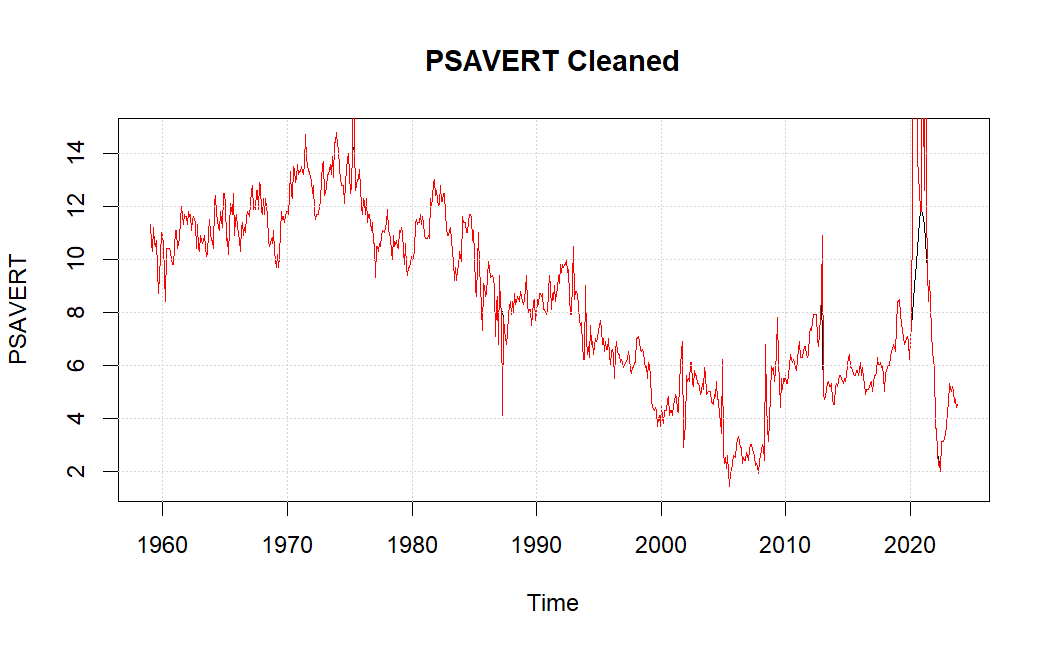
*We will split our data as train and test by taking 12 entries for the test data since it is monthly data.*

1. *anomaly detectıon*

**

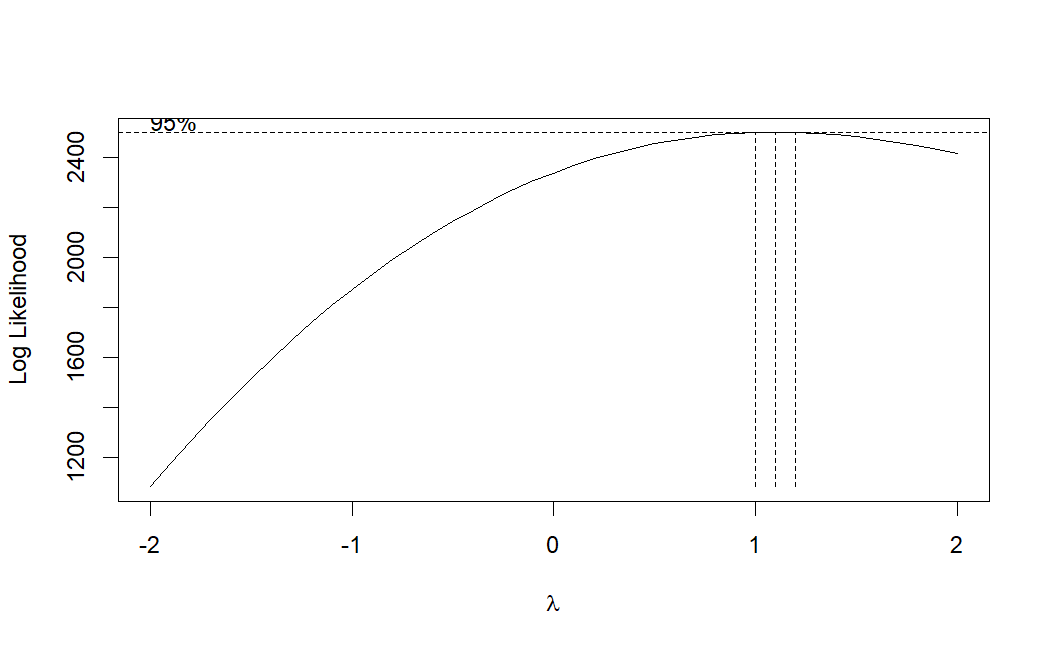
*We can see that there are lots of anomalies in the 2020-2021.*

**

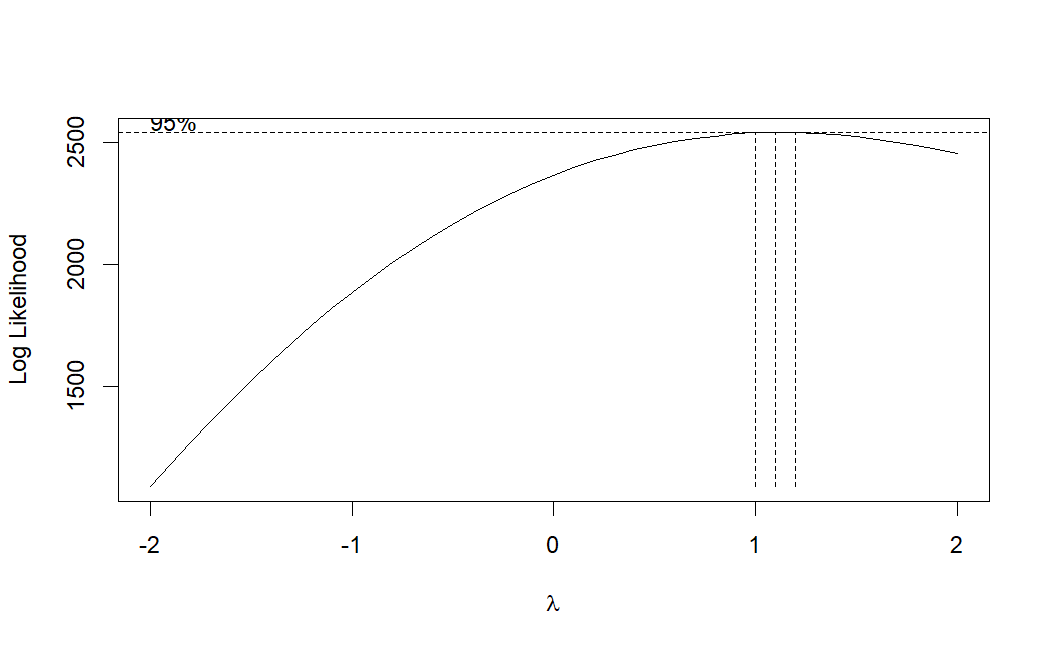
**

*We cleaned the anomalies, mostly the pandemic, we will continue with the cleaned train data.*

1. *box cox analysıs*

**

*Yule-walker*

**

*ols*

*Both of them suggests very close numbers to 1. To be exact, lambda score is 1.19. Applying box-cox does not change our strurcture significantly. Thus we will not apply box-cox. No transformation is needed in this data.*

1. *test*

*KPSS TEST;*

*KPSS Test for Level Stationary p-value:0.01*

*KPSS Test for Trend Stationary p-value:0.01*

*Both of the p-values are smaller than 0.05 so we Reject H0 for both of them. That means, series is not stationary and it has stochastic trend.*

*ADF TEST;*

*ADF Test for Stationary p-value:0.02045*

*ADF Test for Trend p-value:0.03*

*Both of the p-values are smaller than 0.05 so we Reject H0 for both of them. That means, series is not stationary and it has deterministic trend*

*PP TEST;*

*PP Test for Stationary p-value: 0.01*

*It implies the series is stationary*

*HEGY TEST;*

*HEGY Test for Regular Unit Root p-val:0.1*

*HEGY Test for Seasonal Unit Root p-val:0.01*

*HEGY test implies that;*

*series is not stationary(0.1>0.05)*

*contains regular unit root*

*series is not seasonal(0.01<0.05)*

*not contains seasonal root*

*Canova-Hansen;*

*ch Test for seasonal stability p-value: 0.02*

*P-val is smaller than 0.05, this suggests that original model has stochastic seasonality.*

*Our data seems to be not stationary so we will take a regular difference. There seems to be no seasonal root.*

*We will apply the unit root tests again with the differenced data set.*

*KPSS TEST;*

*KPSS Test for Level Stationary p-value:0.1*

*The p-value is smaller than 0.05 so we are not Rejecting H0. That means, series is stationary.*

*ADF TEST;*

*ADF Test for Stationary p-value:0.01*

*The p-value is smaller than 0.05 so we Reject H0. That means, series is stationary*

*PP TEST;*

*PP Test for Stationary p-value: 0.01*

*It implies the series is stationary*

*HEGY TEST;*

*HEGY Test for Regular Unit Root p-val:0.01*

*HEGY test implies that;*

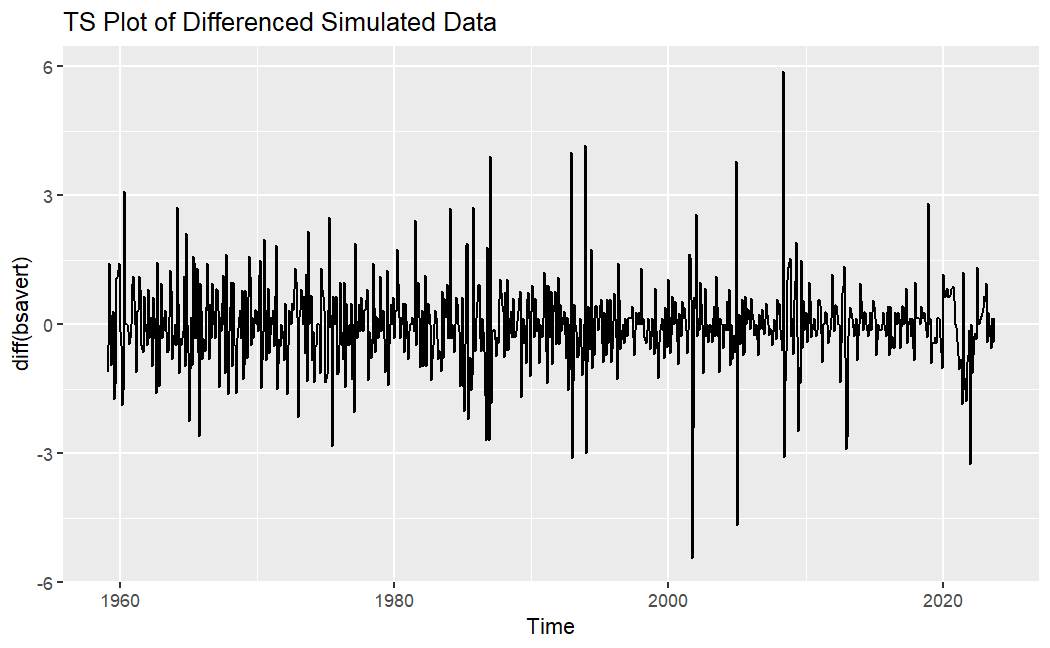
*series is not stationary(0.01<0.05)*

*Does not contain regular unit root*

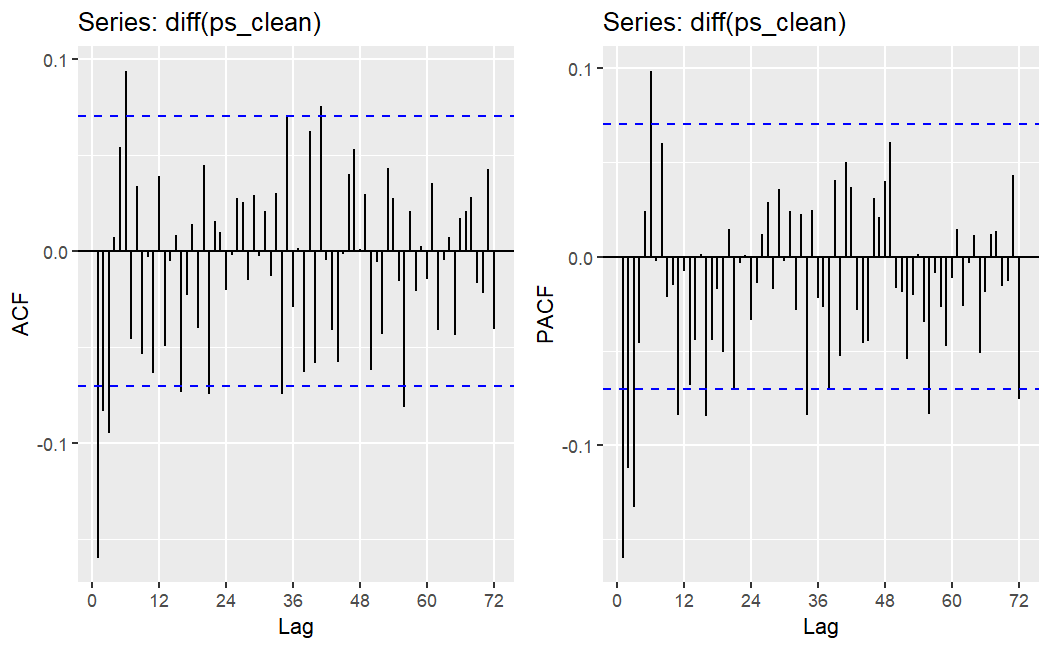
*Canova-Hansen;*

*ch Test for seasonal stability p-value: 0.43*

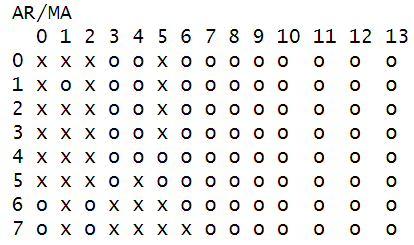
*P-val is bigger than 0.05, this suggests that the series is deterministic and stationary after regular difference.*

*The values suggests that after the regular differencing series become stationary, thus we will continue with the differenced series.*

*Mean is now stationary.*

**

*No decay is observed at acf plot.*

**

*We will suggest ARIMA(4,1,5)*

*Information table find the value 2573.461 which responds to ARIMA(2,1,5)*

1. *Suggestıng models*

*We capturated the significant models in here.*

*ARIMA(3,1,3)(2,0,0)[12]*

*(ar3) |-0.5795/0.0756| > 2 ar3 is Significant*

*(ma3) |0.5932 /0.0979| > 2 ma3 is Significant*

*(sar3) |-0.0221/0.0371| < 2 sar2 is not Significant*

*sigma^2 = 0.3736: log likelihood = -716.11 AIC=1450.22 AICc=1450.45 BIC=1492.12*

*ARIMA(3,1,3)*

*(****ar3)****∣-0.5756/0.0751∣>2 ar3 is S****ignificant***

***(ma3):****∣0.5891/0.0938∣>2 ma3 is* ***Significant***

*sigma^2 = 0.3728: log likelihood = -716.29 AIC=1446.58 AICc=1446.72 BIC=1479.17*

*ARIMA(1,1,1)*

*(****ar1)****∣0.3816/0.1001∣>2 ar1 is S****ignificant***

***(ma1):****∣−0.5870/0.0859∣>2 ma1 is* ***Significant***

*sigma^2 = 0.3798: log likelihood = -725.49 AIC=1456.97 AICc=1457 BIC=1470.94*

*ARIMA(1,1,1)(2,1,3)[12]*

*· (****ar1):****∣0.3595/0.1042∣>2 ar1 is* ***Significant***

*·* ***ma1:****∣−0.5634/0.0903∣>2 ma1 is* ***Significant***

***sar2:****∣−0.9741/0.0454∣>2 sar2 is* ***Significant***

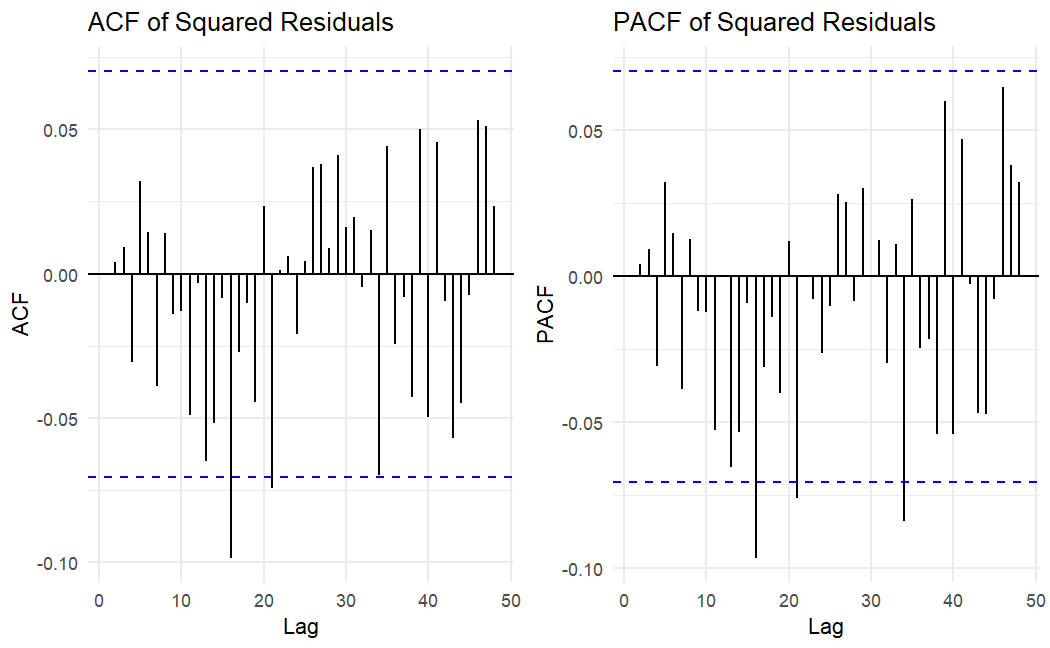
***sma3:****∣−0.9594/0.0596∣>2 sma3 is* ***Significant***

***sigma^2 = 0.3786: log likelihood = -736.76 AIC=1489.52 AICc=1489.71 BIC=1526.64***

*These are the significant models*

*Simpler models are more prefable, ARIMA(3,1,3) , suggests a better fit according to the AIC&BIC scores. Log-likelihood suggests that ARIMA(3,1,3)(2,0,0)[12] and ARIMA(3,1,3) fits better, however adding parameters can increase likelihood, which may cause over-fitting. So we will choose ARIMA(3,1,3)*

1. *dıagnostıc checkıng*

**

*In the acf-pacf plots of the residuals there are some spikes that are outside of the band.*

*Box-Ljung test for the residuals;*

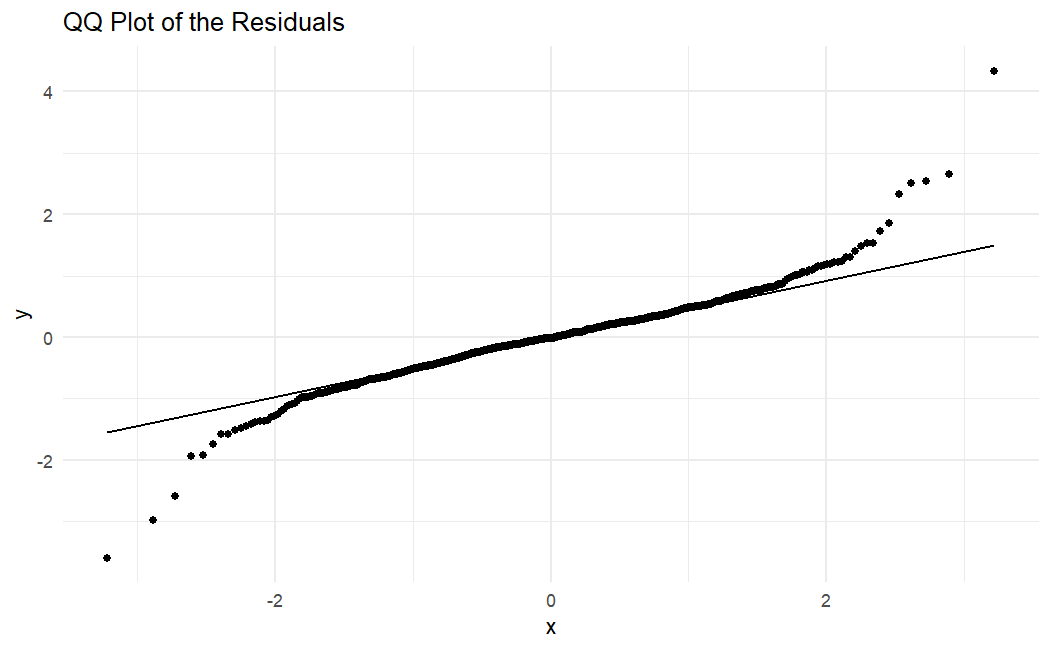
1. *value: 0.8872 We are not rejecting H0 with a 95% confidence. Residuals of the model are uncorrelated*

*Box-Pierce test;*

*P-value: 0.8898 We are not rejecting H0 with a 95% confidence. Residuals of the model are uncorrelated*

*Breusch-Godfrey test for serial correlation of order up to 7;*

1. *value:0.7657 We are not rejecting H0 with a 95% confidence. The residuals of the model are uncorrelated.*

**

*QQ Plot shows that the residuals of the model seems to have tailed distribution.(indicates S shape slightly.)*

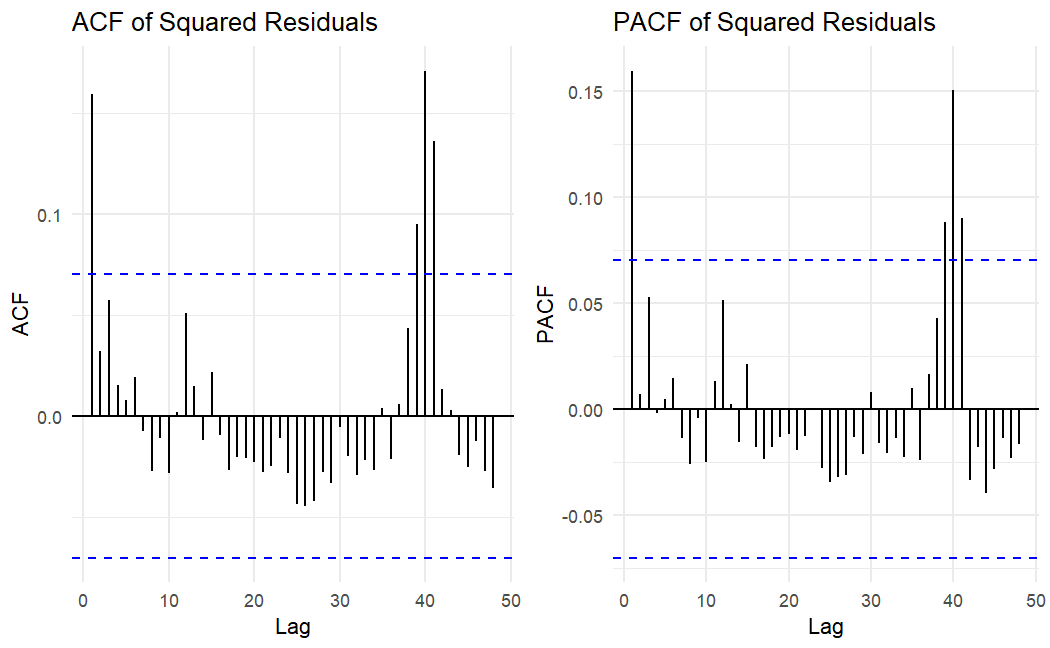
*Jarque Bera Test;*

*P-value: < 2.2e-16*

*Shapiro-Wilk normality test*

*P-value: < 2.2e-16*

*Both of the test suggests rejecting H0. Thus this implies not normal distribution.*

**

*Squared Residuals plot contains some spikes that are outside of the band in both acf and pacf plot. There might be arch/garch effect so we are going to test it.*

*studentized Breusch-Pagan test;*

1. *value:0.07463*

*White Test;*

*P-value:0.1359*

*In both of these tests H0 is Residuals are homoscedastic. (The variance is constant.)*

*Since both of the p-value’s are bigger than 0.05. We can conclude that there is no heteroscedasticity problem.*

*ARCH LM-test*

*P-value:1*

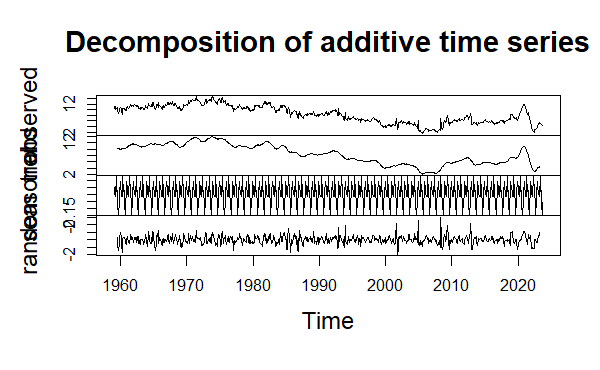
*We cannot reject H0. Therefore, we can conclude there is no ARCH effects.*

1. *forecastıng*

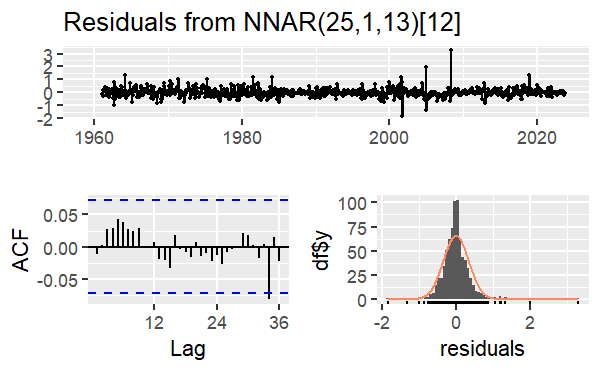
*NN MODEL*

*We are going check first the assumptions;*

*Length of the data is 778, it is more than enough to run nnmodel*

**

*There seems to be no seasonality or trend.*

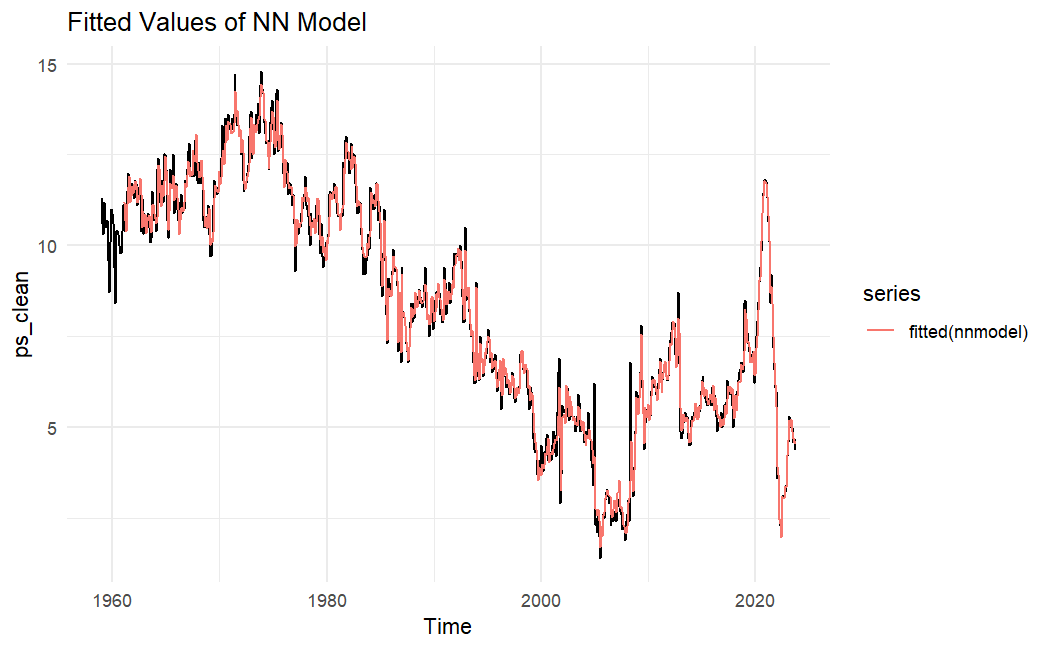
**

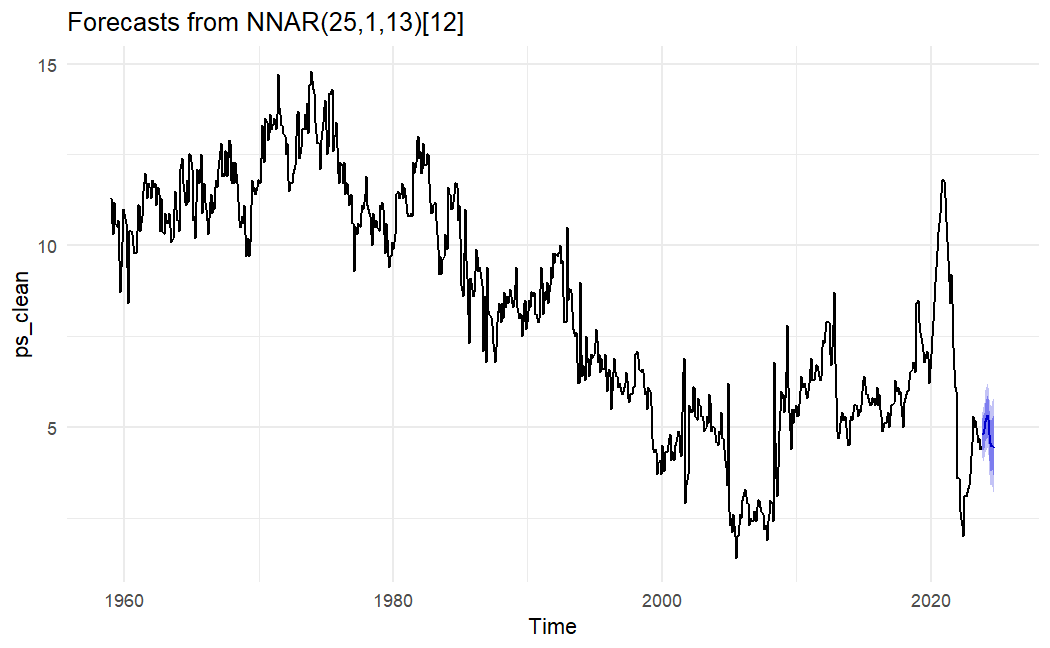
*There is one spike at near lag 36 so it is not WN.*

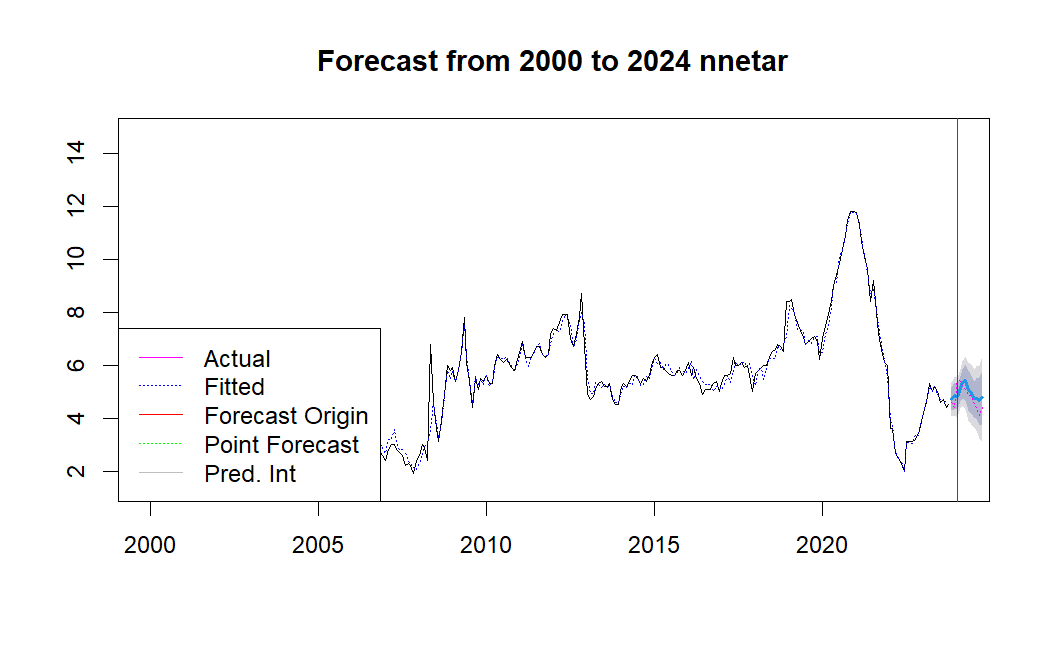
*Residuals seems normal. Errors has stable mean and variance.*

*We will apply studentized Breusch-Pagan test;*

*P-value:0.1523 Thus no residuals relations.*

**

**

**

*TBATS*

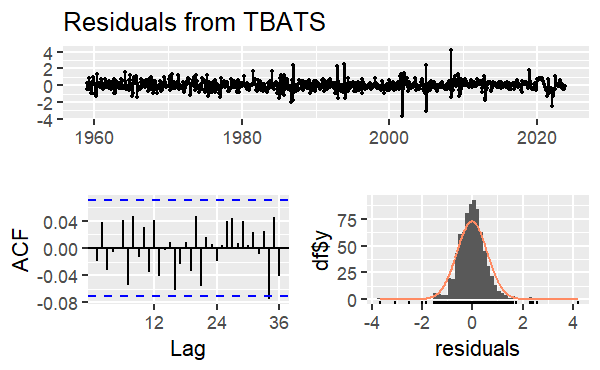
*Assumptions;*

*There is no seasonal patterns.*

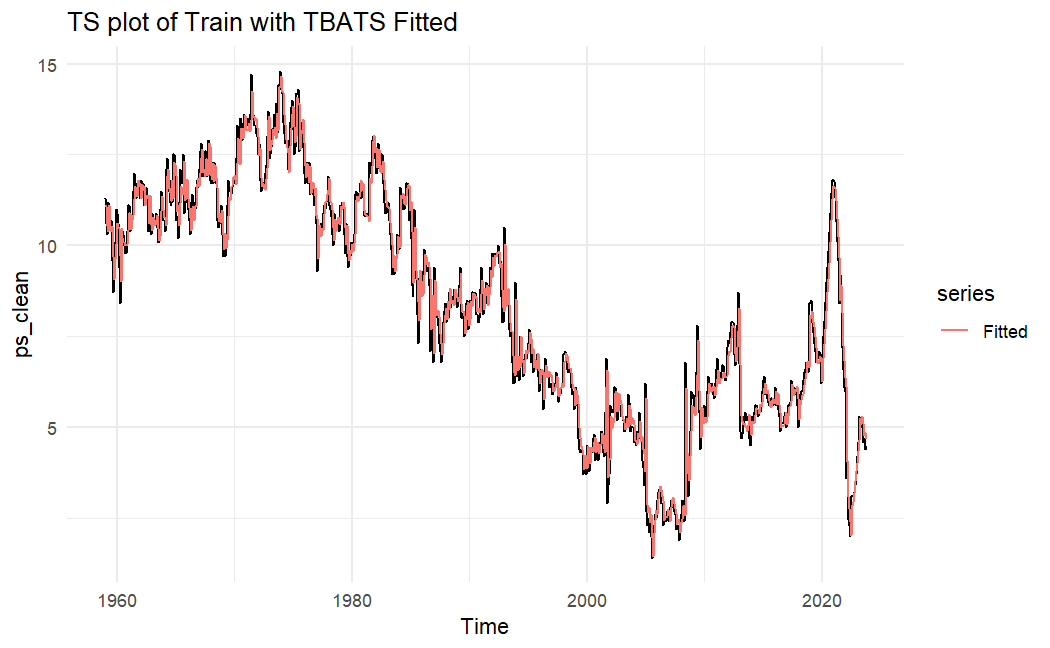
*No need to box cox transformation as well, because variance is stable enough.*

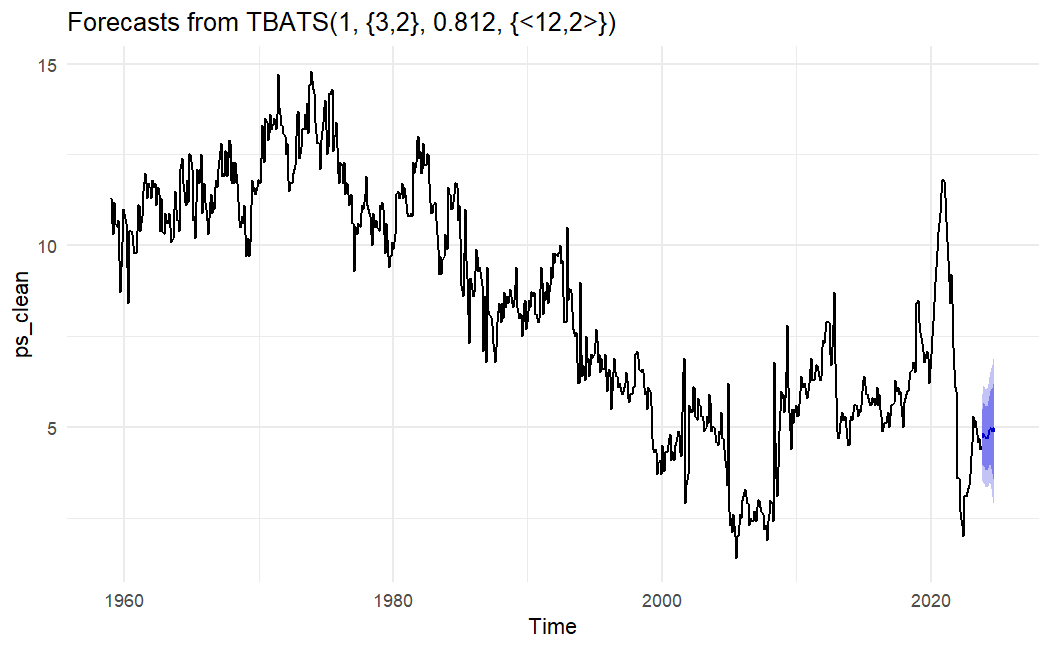
*Augmented Dickey-Fuller Test:   
p-value:0.01*

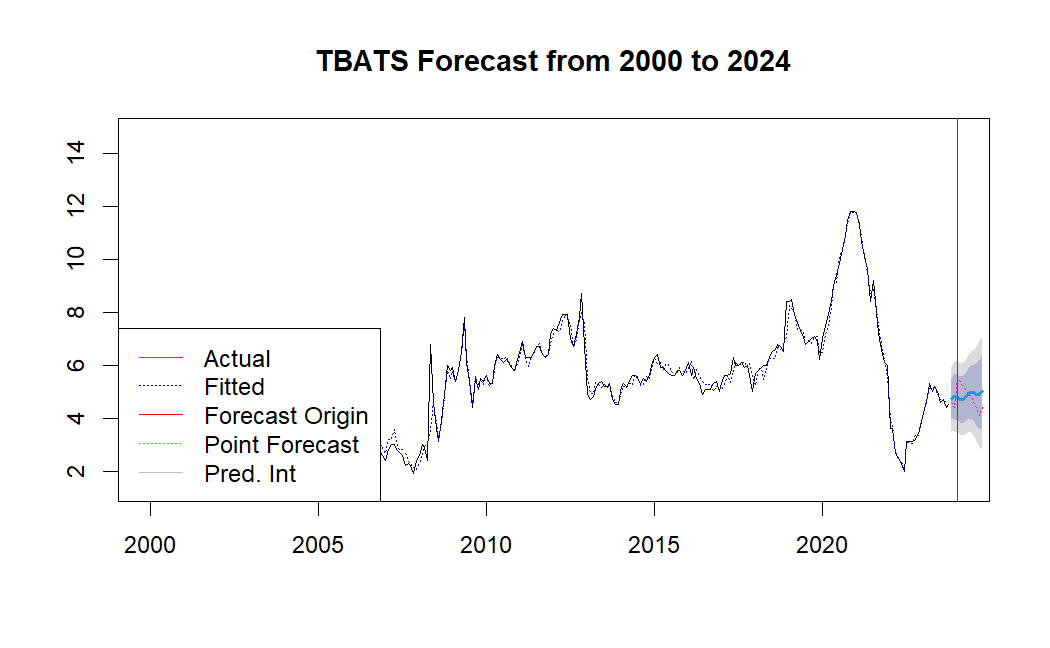
*So the residuals are stationary*

**

*This also confirms that the residuals are stationary and follows a normal distribution.*

**

**

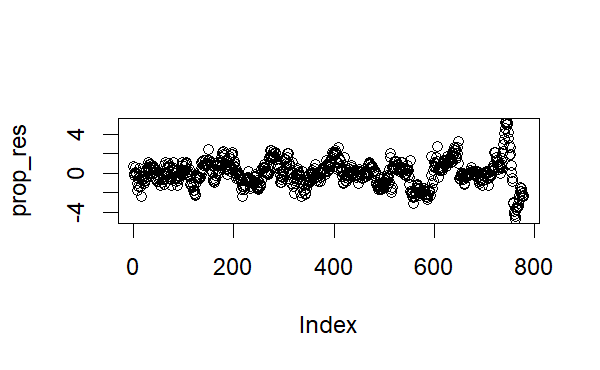
* **

*PROPHET*

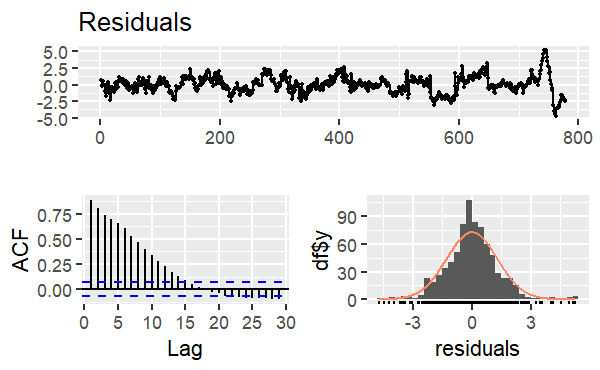
*Assumptions;*

*Data set is in regular frequency*

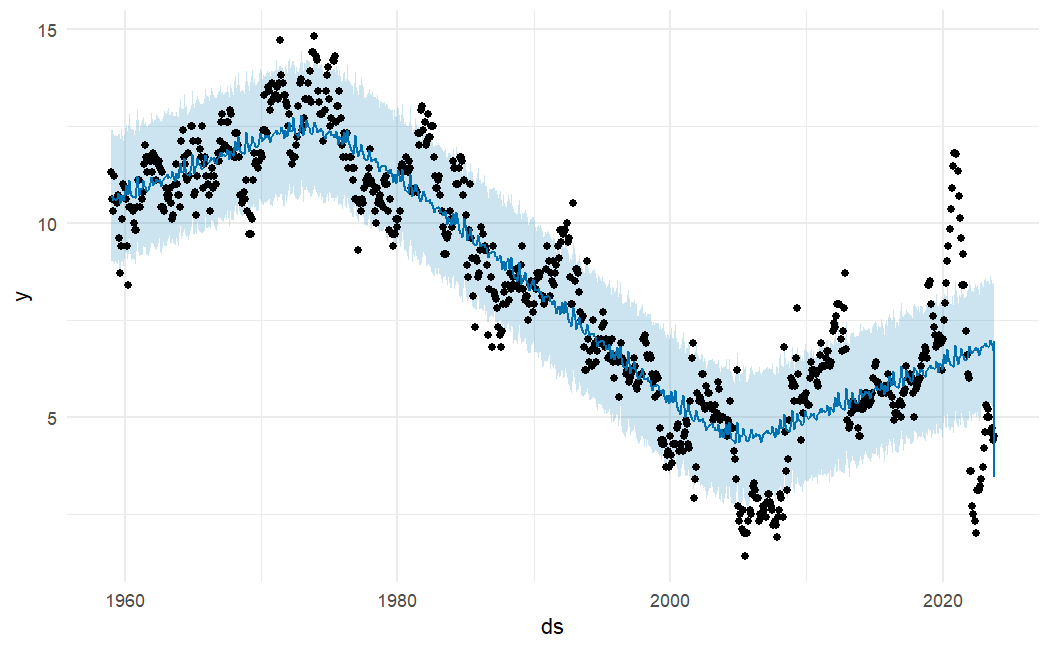
*The other seasonalities is eliminated automaticly*

**

*Residuals have different values.*

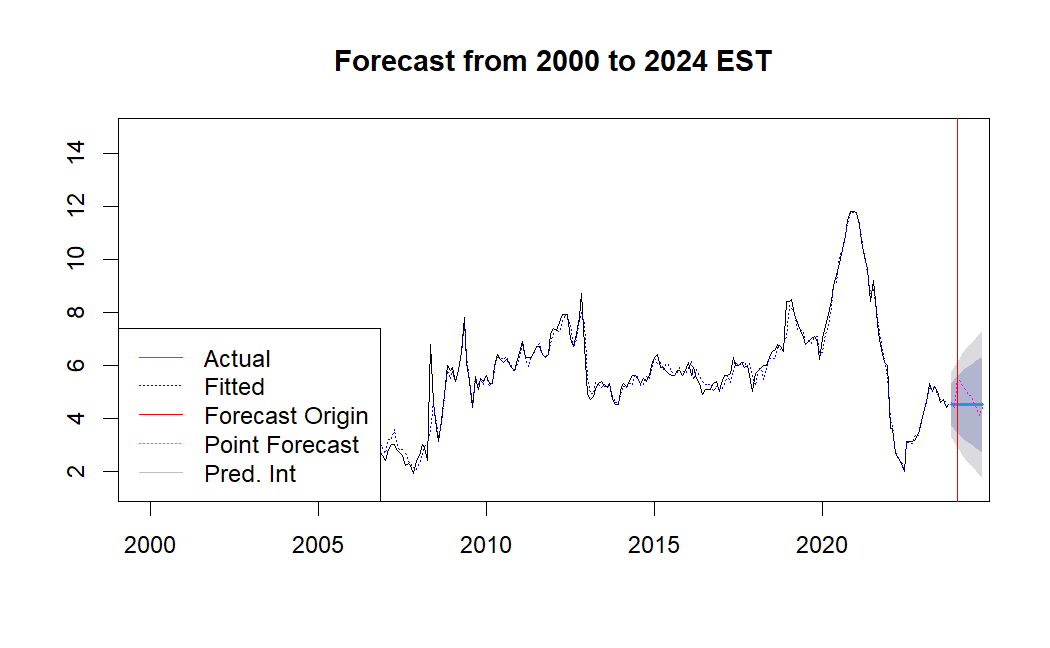
**

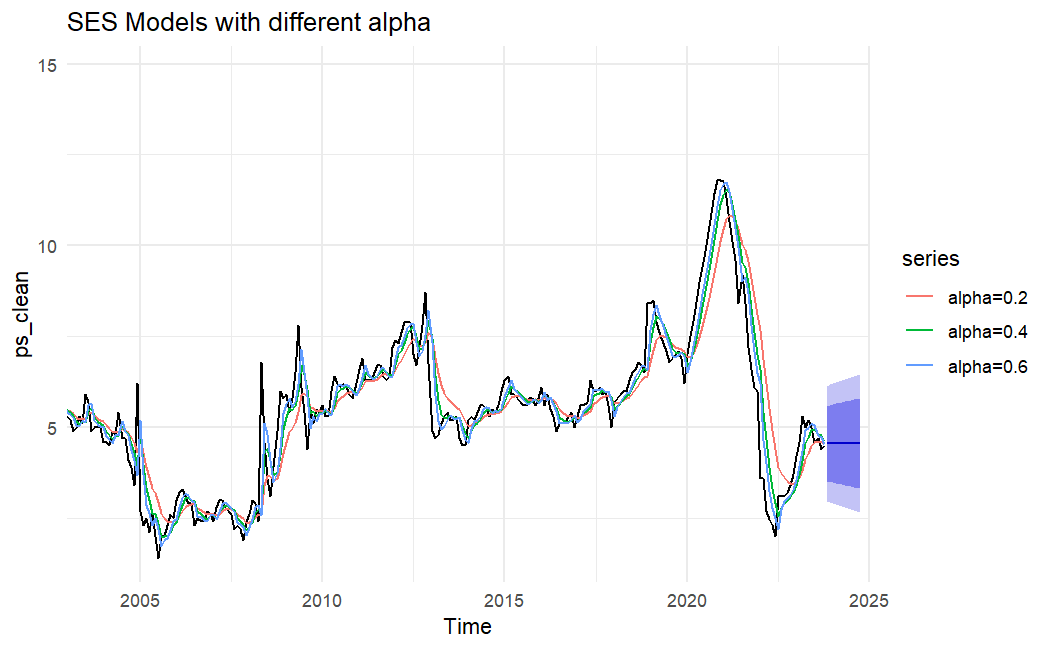
*ACF has a slow linear decay, also in residuals have some different values, so there might be a non-constant variance.*

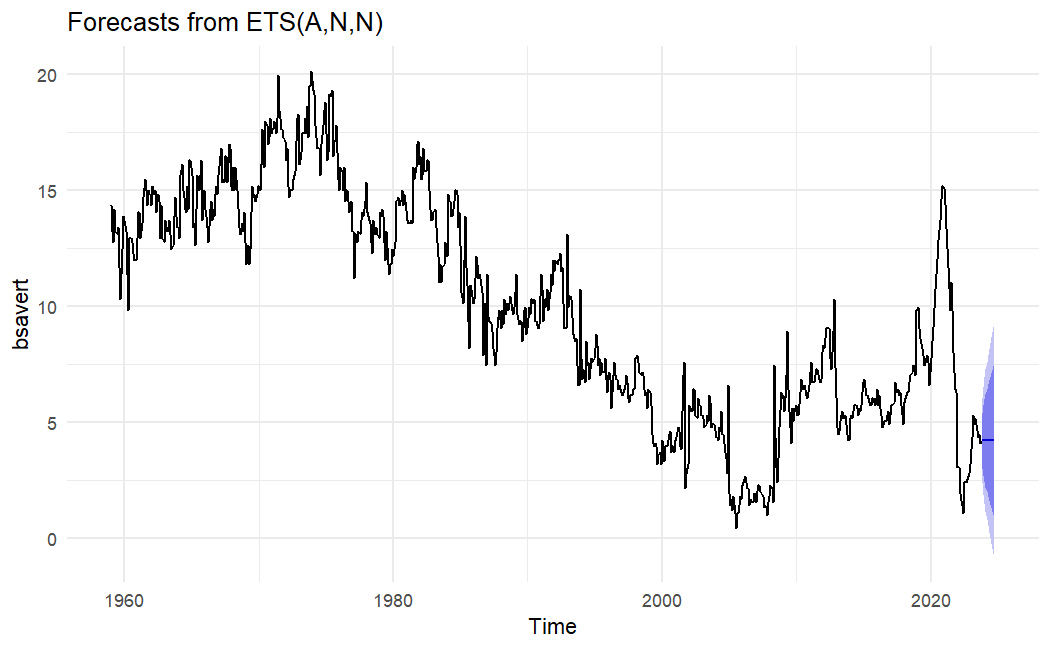
**

*Dots are the original value, blue lines are the prediction. It is not that successful.*

*EST*

**

**

**

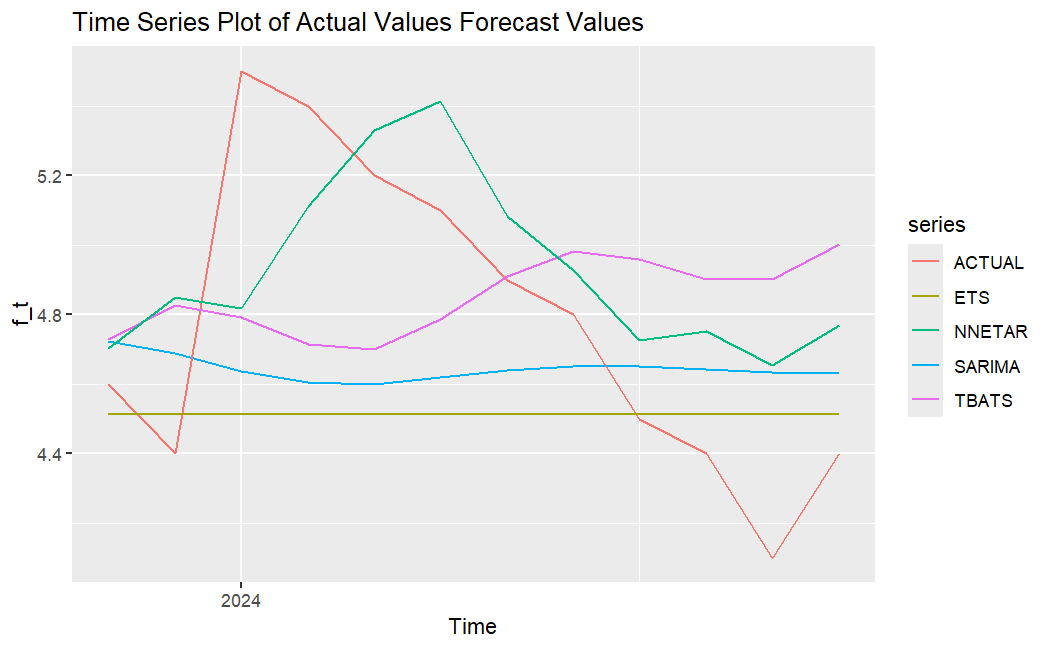
*A,N,N is suggested from the Z,Z,Z*

1. *accuracy*



*nnetar gives the best results according to the errors.*

1. *plots*

**

*We can also see from the plot that the best fit is nnetar.*

1. *conclusıon*

*This study offers important new information about the dynamics of the US personal savings rate. The research emphasizes how the COVID-19 epidemic affected saving behavior. The results imply that, in comparison to more conventional time series models like ARIMA, neural network models especially nnetar may provide more reliable and accurate forecasts.To find out how reliable these models are in various economic scenarios and to look into possible ways to increase forecasting accuracy, more study is necessary. Businesses, investors, and regulators may all benefit from this research's ability to help them understand and predict shifts in consumer saving habits.*

1. *refefences*

**[1]** Federal Reserve Economic Data (FRED), "Personal Saving Rate (PSAVERT)," Federal Reserve Bank of St. Louis. Available: [https://fred.stlouisfed.org/series/PSAVERT](https://fred.stlouisfed.org/series/PSAVERT" \t "_new).