Technical report

Part 1--Exponential Smoothing

1. Data Preparation

1.1 Global Mean Surface Temperature Anomaly (MSTA) in ℃

First of all, it is better to truncate the dataset to more recent conditions. The time span of MSTA is too long, dating back to 1850 at the earliest. The early data not only does not have sufficient reference but also leads to the final time plots being too dense, which has an impact on the observation of other trends. In order to be consistent with ET12, the adjusted data should start in 1995 (Figure 1.1.1.1). What's more, because this data is for calendar months, then the account might have to be taken of the length of a month, so we did calendar adjustments (Figure 1.1.1.2).

1.2 Global Monthly Atmospheric Carbon Dioxide Levels (CH4)

As above, for clearer analysis and consistency, the data were truncated from 1995. Moreover, the analysis of indicators in recent years is also more representative (Figure 1.1.2.1). This data is also for calendar months. In theory, it should be calendar adjustments. But after the adjustments, the image was found to be denser and more difficult to observe. It is decided that CH4 will not be calendar adjusted (Figure 1.1.2.2).

1.3 International Passenger Survey, UK visits abroad (GMAF)

The same thing as before, the data is also for long datasets and calendar months, so we truncated the data to 1995 (Figure 1.1.3.1) and did calendar adjustments. Although there's not much difference graphically from not calendar adjusting (Figure 1.1.3.2).

1.4 UK inland monthly energy consumption (ET12)

In this data, we also need to do calendar adjustments cause of its calendar months. After doing the adjustments, the time plot looks more smoothy than before (Figure 1.1.4.1 is before adjustment, and Figure 1.1.4.2 is after adjustment).

2. Preliminary Analysis

2.1 Global Mean Surface Temperature Anomaly (MSTA) in ℃

As shown in Figure 1.1.1.1, the MSTA in ℃ from January 1995 to December 2021 was used as the sample analysis data. The MSTA in January 1995 was 0.41360325℃, and the MSTA in December 2021 was 0.751281℃. The CHANGES of MSTA from January 1995 to December 2021 are shown in Figure 1.1.1.1. As can be seen from Figure 1.1.1.1, the overall trend of MSTA increased from 1995 to 2021. The data on MSTA contains occasional large fluctuations which are difficult to explain. These data provide preliminary evidence that the earth's temperature is gradually warming. Based on studies of factors associated with global warming, air temperatures on Earth have been rising since the Industrial Revolution. While natural variability plays some part, the preponderance of evidence indicates that human activities—particularly emissions of heat-trapping greenhouse gases—are mostly responsible for making our planet warmer.

2.2 Global Monthly Atmospheric Carbon Dioxide Levels (CH4)

See Figure 1.1.2.1, the CH4 time plot displays a clear trend and seasonality. Note that it shows a trend and a slight seasonality. The annual rate of increase in atmospheric carbon dioxide over the recent years is much faster than previous natural increases. And I think that this result is directly related to human activity. Human activities have increased the concentration of carbon dioxide in our atmosphere, amplifying Earth's natural greenhouse effect.

2.3 International Passenger Survey, UK visits abroad (GMAF)

As shown in Figure 1.1.3.1, the GMAF plot also displays a clear trend and seasonality. The magnitude of the change increases with time. In recent years, more and more people go abroad and has a very obvious seasonal. It can be seen from the seasonal plot that August is the peak time to go abroad, probably because of the summer vacation and the relatively warm weather, the temperature is very suitable for traveling. The lowest point is reached in the cold December, when the cold weather makes it less suitable to travel and most people stay at home, or because of Christmas, at home with family.

2.4 UK inland monthly energy consumption (ET12)

As shown in Figure 1.1.4.1, UK inland monthly energy consumption was relatively stable until 2008. It began to decline after 2008 and reached its lowest point in 2020, with a slight upward trend in 2021. It can also be seen from the figure that the seasonality of ET12 is obvious. It can be seen from the seasonal plot that ET12 reaches its lowest point in August and the overall consumption is low in summer. This is probably because natural gas consumption is lower in the summer, A decrease in coal consumption largely offset an increase in providing space heating and hot water petroleum consumption. as a bigger proportion of household consumption is for space and water heating. It is also possible that global warming has led to a gradual decrease in energy consumption in recent years.

3. Model Selection and Model Effect

3.1 Global Mean Surface Temperature Anomaly (MSTA) in ℃

The seasonal image of MSTA is shown in Figure 1.3.1.1. From 1980 to 1999, there were some seasonal trends, but this trend rarely slowed down in the past 20 years. Combined with its overall image (Figure 1.1.1.1), It can be considered that MSTA has additive trend effect and additive seasonal effect, so we should choose Holt-Winter’s method with additive seasonality model.

As can be seen from Figure 1.3.1.4 and Figure 1.3.1.5, the mean value of the residual tends to be 0, and there is no significant correlation in the residual sequence, so the residual variance can be regarded as a constant. From Figure 1.3.1.6 and Figure 1.3.1.7, the optimized model fits better, and the sample autocorrelation coefficient soon attenuates to a small random fluctuation around 0. Therefore, it can be judged that the series is a stationary time series.

3.2 Global Monthly Atmospheric Carbon Dioxide Levels (CH4)

As can be seen from Figure 1.3.2.1, CH4 also has some seasonality, but it is not obvious. Combined with the overall trend (Figure 1.1.2.1), we believe that CH4 has an additive trend effect and no seasonal effect, so we should choose Holt’s linear exponential smoothing model.

As can be seen from Figure 1.3.2.4 and Figure 1.3.2.5, the mean value of the residual tends to be 0, and there is no significant correlation in the residual sequence, so the residual variance can be regarded as a constant. It can be seen from Figure 1.3.2.6 and Figure 1.3.2.7 that the graphs of ACF show obvious sinusoidal fluctuation rules, which are typical characteristics of time series with periodicity. Therefore, it can be judged that the time series has periodicity and tendency, which is consistent with our previous judgment, so it can be judged that the time series is non-stationary.

3.3 International Passenger Survey, UK visits abroad (GMAF)

As can be seen from Figure 1.3.3.1, no matter which period, the seasonal trend is very obvious. Combined with Figure 1.1.3.1, it can be seen that the fluctuation range of GMAF increases with the passage of time, so we can think that GMAF has additive trend effect and multiplicative seasonal effect, and we should choose Holt-Winter's method with multiplicative seasonality model.

As can be seen from Figure 1.3.3.3, the mean value of the residual tends to be 0, and there is no significant correlation in the residual sequence except for one outlier, so the residual variance can be regarded as a constant. From Figure 1.3.3.4, the sample autocorrelation coefficient soon attenuates to a small random fluctuation around 0. Therefore, it can be judged that the series is a stationary time series.

3.4 UK inland monthly energy consumption (ET12)

As can be seen from Figure 1.3.4.1, the seasonal trend of ET12 is particularly obvious, and as can be seen from Figure 1.1.4.1, the change range of ET12 has not changed over time. It follows that ET12 has additive trend effect and additive seasonal effect, and we should choose Holt-Winter's method with additive seasonality model.

As can be seen from Figure 1.3.4.4 and Figure 1.3.4.5, the mean value of the residual tends to be 0, and there is no significant correlation in the residual sequence, so the residual variance can be regarded as a constant. From Figure 1.3.4.6 and Figure 1.3.4.7, the optimized model fits better, and the sample autocorrelation coefficient soon attenuates to a small random fluctuation around 0. Therefore, it can be judged that the series is a stationary time series.

Part 2—ARIMA

1. data preparation

-- As our data is the global mean surface temperature anomaly, the values are positive and negative. So, we cannot use the square root and log transforms. And the positive and negative signs of temperature have different meanings, so we cannot square them.

--Then we found that the time have several formats. In order to read the data easily, we change all the time into one format, such as 2000-01. And put the new data into MSTAdata_33276048_33347999_33270635.xlsx and the sheet name is MSTA.

2.preliminary analysis

2.1 With code: ARIMA_TimePlot_33276048_33347999_33270635.py, we can get a time plot of global mean surface temperature anomaly (MSTA) (Figure 2.1). As we can see the time series with upward trend is non-stationary.

2.2 From code: ARIMA_AcfPacf_33276048_33347999_33270635.py, we can be sure that the MSTA time serious is non-stationary because the value of ACF descends slowly and the second value of PACF plot is a large spike. (Figure 2.2)

2.3 Using code: ARIMA_1stDifferencing_33276048_33347999_33270635.py, we apply first order differencing. In figure 2.3, time, ACF and PACF plot has changed a lot. The obvious trend in time plot has gone, but the seasonality is still present. In the ACF and PACF plot, the large spikes accrue every 12th lag.

3. select the ARIMA model

-- We did first order and seasonal differencing with code: ARIMA_1stSeasonalDiff_33276048_33347999_33270635.py. Figure 2.4 is the plots after applying seasonal differencing, and figure 2.5 is the plots after both first order and seasonal differencing. Then we can find that there are no obvious trend and cycle in the time plot of figure 2.5. So, the time series is stationary after applying first order and seasonal differencing. According to the ACF and PACF plot in figure 2.5, many the values are outside the $\pm {1.96}/{\sqrt{n}}$ range, the new serious is not a white noise serious.

-- Besides, both autocorrelation and partial autocorrelation coefficients are trailing, so we assume $p = 1$ and $q = 1$. As we did the first order differencing, $d = 1$. It is repeated approximately after every 12th lag in the ACF and PACF, so we have $s = 12$. In this way, it is concluded that the preliminary best model is ARIMA (1,1,1) (1,1,1)12.

-- Then, we test ARIMA (1,1,1) (1,1,1)12 with code: ARIMA_forecast_33276048_33347999_33270635.py. it shows that the P value of the Q statistics is 0.20>0.05, which means the model pass the Ljung-Box tests. And the AIC is -2962.483. The MSE of this model is 0.01, close to 0, which means the model fits well.
In figure 2.6, top right figure shows the KDE is close to N (0,1), and the QQ-plot shows that the points are close to the red line. These mean residuals approximate the standard normal distribution. Moreover, top left figure shows there are no obvious trend and cycle in the time plot of residuals. And in the bottom right figure, the values are inside the $\pm {1.96}/{\sqrt{n}}$ range, so the residuals serious is a white noise serious. The model is valid.

-- In figure 2.7, we show the one-step ahead forecasts together with the original data set of models ARIMA (1,1,1) (1,1,1)12. The results seem reasonable.

In figure 2.8, we forecast 12 steps ahead in future of model ARIMA (1, 1, 1) (1, 1, 1, 12).

-- In order to find that if there is a better model, we use code: ARIMA_BestModel_33276048_33347999_33270635.py. This code runs all the possible combinations of different values of p, d, q, P, D, Q and compares their AIC. Finally, it shows that the models with the smallest AIC.ARIMA(1, 1, 1) (1, 0, 1, 12) with AIC: -3011.428083646573

-- Again, we use code: ARIMA_TestNew_33276048_33347999_33270635.py to test this model. about ARIMA (1, 1, 1) (1, 0, 1, 12), P value of the Q statistics is 0.08>0.05, which means the model pass the Ljung-Box tests. The MSE of this model is 0.1, closed to 0, which means the model fits well.

-- In figure 2.9, top right figure shows the KDE is close to N (0,1), and the QQ-plot shows that the points are close to the red line. These mean residuals approximate the standard normal distribution. Moreover, top left figure shows there are no obvious trend and cycle in the time plot of residuals. And in the bottom right figure, the values are inside the $\pm {}^{1.96}/_{\sqrt{n}}$ range, so the residuals serious is a white noise serious. The model is valid.

-- In figure 2.10, we show the one-step ahead forecasts together with the original data set of models ARIMA (1,1,1) (1,0,1)12. The results seem reasonable.
In figure 2.11, we forecast 12 steps ahead in future of model ARIMA (1, 1, 1) (1, 0, 1, 12).
Finally, we decide ARIMA (1, 1, 1) (1, 0, 1, 12) is the best model for forecasting the global mean surface temperature Anomaly (MSTA)

2.4 compare ARIMA and exponential smoothing forecasting

-- The steps of ARIMA and exponential smoothing forecasting are the same. First, we need to make the time plot. And according to the plot, analyze the trend and the seasonality of the time serious. Then select the suitable model. Finally, forecast the future values.

-- There are several exponential smoothing models with some ARIMA models which are like the exponential smoothing models.
When trend and seasonality are not present, the single exponential smoothing can be used. And it is similar to ARIMA (0,1,1).

-- When a possible linear trend is present and the seasonality is not, we can choose the Holt's linear exponential smoothing method, which is like ARIMA (0,2,2).

-- When the seasonality and the linear trend are both present, there are two versions, additive and multiplicative. Holt-winter's method, multiplicative seasonality should be used when variability is amplified with trend, whereas the additive approach is more suitable in the case of constant variability. Besides, Holt-winter's method, additive seasonality is similar to SARIMA (0,1,0)(0,1,1)S

--In this instance, we choose a taxonomy of exponential smoothing method with additive trend effect and additive seasonal effect, which is the additive Holt-winters' method. And we can find the trend and seasonality through the ARIMA forecasting.

<u>Part 3—Regression Model</u>

1.Data preparation

1.1 MSTA: data range from 1850-01 to 2021-12, CH4: data range from 1983-07 to 2021-12

GMAF: data range from 1980-01 to 2020-03, ET12:data range from 1995-01 to 2021-12.

1.2 -Use 'MSTAdata_33276048_33347999_33270635.xlsx' in regression model.
-The sheet: 'Initial data': In order to apply the regression model, the length of all dependent variable and independent variable should keep the same. Time range from 1995-01 to 2020-03 were chosen for all above four data. There are 303 rows in total.
-The sheet: 'data1' displays the cleaned data and a matrix with entries of 1s, where the columns are from D1 to D11, used for time plot and generating model summaries.
- The sheet: 'data2' displays the four columns of data, used for analyzing correlation.
- The sheet:'data3' is a single column of all the data in MSTA, use to forecast.

2.preliminary analysis

- By code: Regression_TimePlot_33276048_33347999_33270635.py, we reveal the relationship of the 4 variables against time, and it use the sheet “data2”
- In Figure 3.2.1, the MSTA shows a increasing trend, which corroborate with the fact that the global warming is getting server during these time period. But it also contains occasional large fluctuations which are difficult to explain.
- In figure 3.2.2, the time plot of CH4 indicates an obvious increasing trend and seasonality. The rate of Carbon Dioxide Levels increase is much faster after around 2010 (about halfway through). This may be the result of rapid growth on global population and human activity.
- In figure 3.2.3, the time plot of GMAF displays a clear trend and seasonality. The number of International Passenger increases more rapid in recent years. May due to economic growth and population growth. It can also be deduced that the travel peak time for is in the summer and there are much fewer people traveling in the end of the year.
- In figure 3.2.4, it is the time plot of ET 12. There is a decrease trend and seasonality. The data drops in an obvious pattern after approximately in the year of 2009. As opposite to the GMAF, the energy consumption has the lowest point in the summer but highest in the end of the year.
- By code: Regression_Correlation_33276048_33347999_33270635.py, it displays the correlation between variables. From figure 3.2.5, we can see from the scatter plot that the correlation between MSTA and CH4 is 0.770224 and it is quite close to 1. GMAF and ET12 has higher correlation of -0.803479, which indicate a negative linear relationship close to 1. No clear linear patterns can be deduced for other variable combinations.

3.Model selection and analysis

3.1. ANOVA table

-- By the code: Regression_ANOVA_33276048_33347999_33270635.py, from figure 3.3.1, the AVOVA table indicates the R-square is 0.594, which is greater than 0.5 but not really close to 1. The model also has p-value 3.44e-58, which indicate that this regression model (MSTA ~ CH4 + GMAF + ET12) is significant. From the individual p-values, the independent variable GMAF and ET12 are not significant as their p-values are 0.681, 0.986, which are much larger than 0.05. So, there is possibilities to improve the model.

3.2 Add month indicator to the model

-- By the code: Regression_MonthIndicator_33276048_33347999_33270635.py, it introduces 11 monthly indicator variables and fits a new model using the sheet “data1”. From Figure 3.3.2, the R-square has increased to 0.619, and the model p-value also getting smaller to 3.50e-52. This implies that the adding of monthly indicator improved the model performance. As for the individual variable, the significance of GMAF and ET12 have also been improved because now they have p-values of 0.126 and 0.104. Although they are still slight bigger than 0.05, they do decrease in a great extent compared to the previous model.

3.3 Add monthly indicator +Time to the model

-- By the code: Regression_MonthIndicator+Time_33276048_33347999_33270635.py It generates a new model with formular. In Figure 3.3.3, we can see that model R-square increases to 0.624. Its P-value decreases to 3.69e-52, both information shows that the model has been improved compared to 3.2. However, when comparing the individual variable, we notice that the p-value for all three independent variable increases, especially for GMAF, rise from 0.126 to 0.492.

3.4 Model comparison

-- By the code: Regression_ModelCompare_33276048_33347999_33270635.py, as the three models result above, adding time and monthly indicator fits model better. The R-square of model 3.1 to 3.3 is 0.594 ~0.624, which is getting closer to 1.P-value are both significant. But need investigate more statistics rather than decide depends only one. Individual The p-value also become smaller compare the model 3.1 to 3.3, but still GMAF and ET12 are that significant in the overall model.One reason may be that these two variables can be strongly related, so that they don't really provide linearly independent information.

4.Forecasting

-- By the code: Regression_Forecasting_33276048_33347999_33270635.py, using the sheet "data3", we use the existing 303 rows of MSTA data to predict the future trend from end of the data in 2020.Mar up to 2022 Dec, where there are 33 months to be predicted.

-- From Figure 3.4.1 to 4.3.2, we can see that the predicted data fit the original data well and shows great seasonality, especially for GMAF. Both CH4 and GMAF forecasted a reasonable trend for the upcoming months. However, in Figure 4.3.3, the forecasting for ET12 indicates that there is going to have a increasing trend in energy consumption, which does not line up with the original trend that it was decreasing before 2020 Mar. Finally in Figure 3.4.4, it displays the predicted trend for MSTA. Generally, it shows a proper increase trend for the future that Global Mean Surface Temperature Anomaly is going to increase gradually every month. Overall, since the variable shows seasonality, holt linear method is a proper way for forecast, and the linear assumption is suitable.

Appendix A: Code descriptions

Exponential smmothing

1. ExpSmooth_MSTA_33276048_33347999_33270635.py: The adjusted MSTA (after truncating and calendar adjustments) time plot and seasonal plot were obtained. Build model 1 with additive trend effect and additive seasonal effect and optimized model 1. Get residual and ACF plot to judge the efficiency of the model.
2. ExpSmooth_CH4_33276048_33347999_33270635.py: The adjusted CH4 (after truncating and calendar adjustments) time plot and seasonal plot were obtained. Build model 1 with additive trend effect and no seasonal effect and optimized model 1. Get residual and ACF plot to judge the efficiency of the model.

3. ExpSmooth_GMAF_33276048_33347999_33270635.py: The adjusted GMAF (after truncating and calendar adjustments) time plot and seasonal plot were obtained. Build optimized model 1 with additive trend effect and multiplicative seasonal effect. Get residual and ACF plot to judge the efficiency of the model.
4. ExpSmooth_ET12_33276048_33347999_33270635.py: The adjusted ET12 (after calendar adjustments) time plot and seasonal plot were obtained. Build model 1 with additive trend effect and additive seasonal effect and optimized model 1. Get residual and ACF plot to judge the efficiency of the model.

ARIMA

1.ARIMA_TimePlot_33276048_33347999_33270635.py: get a time plot of global mean surface temperature anomaly (MSTA)
2.ARIMA_AcfPacf_33276048_33347999_33270635.py: get the ACF and PACF plot.
3.ARIMA_1stDifferencing_33276048_33347999_33270635.py: apply first order differencing of MSTA
4. ARIMA_1stSeasonalDiff_33276048_33347999_33270635.py: did first order and seasonal differencing
5. ARIMA_forecast_33276048_33347999_33270635.py: test if the model we chose (ARIMA(1,1,1)(1,1,1)12) fits well.
6. ARIMA_BestModel_33276048_33347999_33270635.py: find if there is a better model.
7. ARIMA_TestNew_33276048_33347999_33270635.py: test if the new model (ARIMA(1, 1, 1)x(1, 0, 1, 12)) fits well.

Regression

1. Regression_TimePlot_33276048_33347999_33270635.py: get the time plot of all four variables for preliminary information.
2. Regression_Correlation_33276048_33347999_33270635.py: find correlation scatter plot among the four variables.
3. Regression_ANOVA_33276048_33347999_33270635.py： generate ANOVA table for the 1st model with dependent variable MSTA ~ independent variable CH4, GMAF, ET12.
4. Regression_MonthIndicator_33276048_33347999_33270635.py: generate summary for the 2nd model with 11 added monthly indicators.
5. Regression_MonthIndicator+Time_33276048_33347999_33270635.py: generate summary for the 3rd model with 11 monthly indicator and time.
6. Regression_ModelCompare_33276048_33347999_33270635.py: compare the three models with their statistics.
7. Regression_Forecasting_33276048_33347999_33270635: find the forecast data till 2020 Dec.

Appendix B: Analysis and forecast graphs

Part-1 Exponential Smoothing

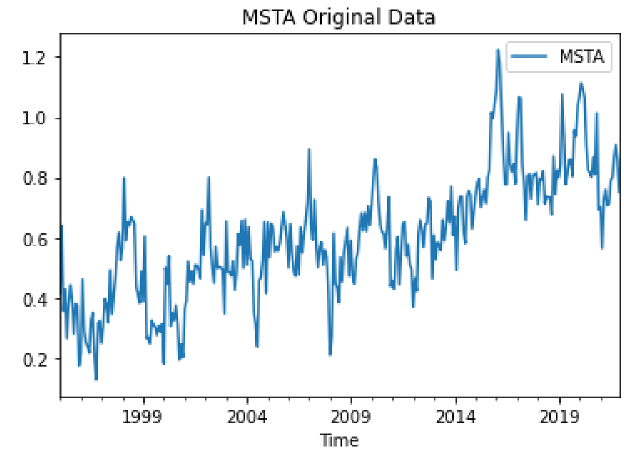


Figure 1.1.1.1

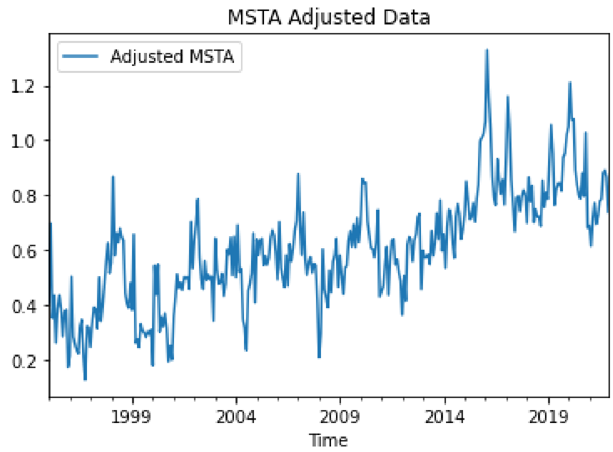


Figure 1.1.1.2

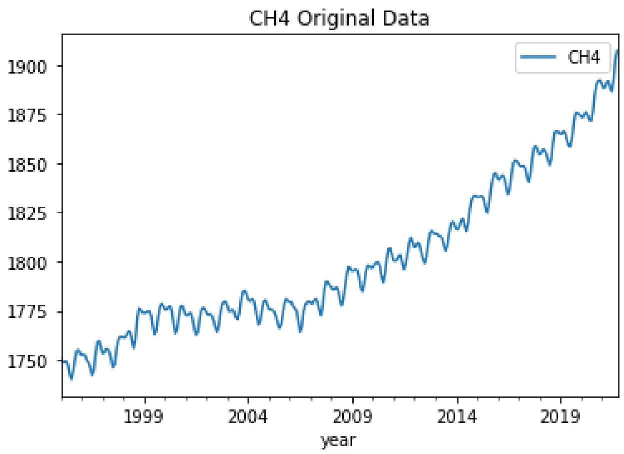


Figure1.1.2.1

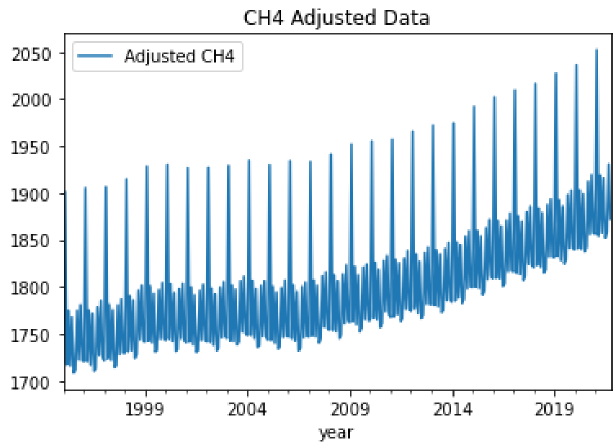


Figure1.1.2.2

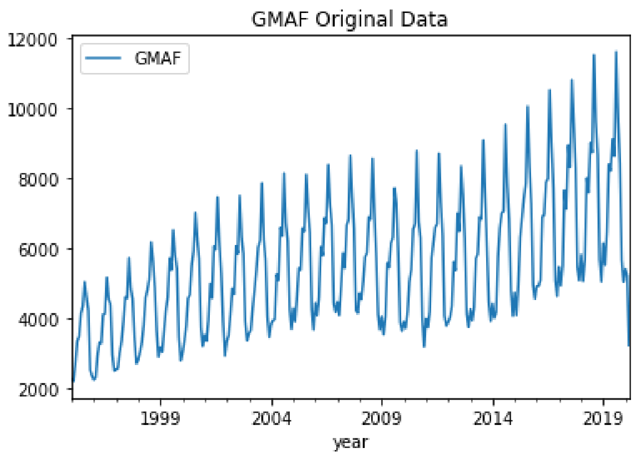


Figure1.1.3.1

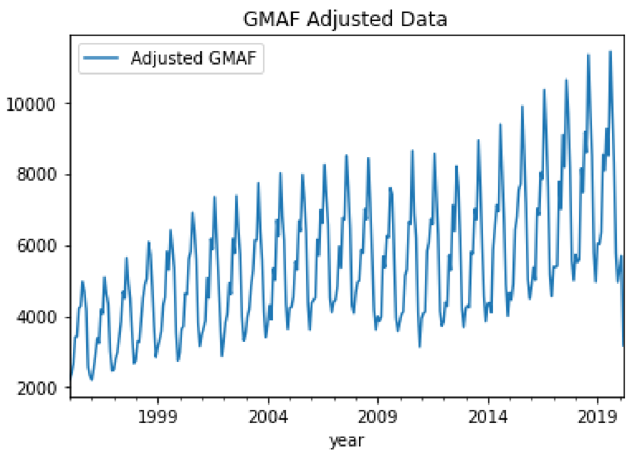


Figure1.1.3.2

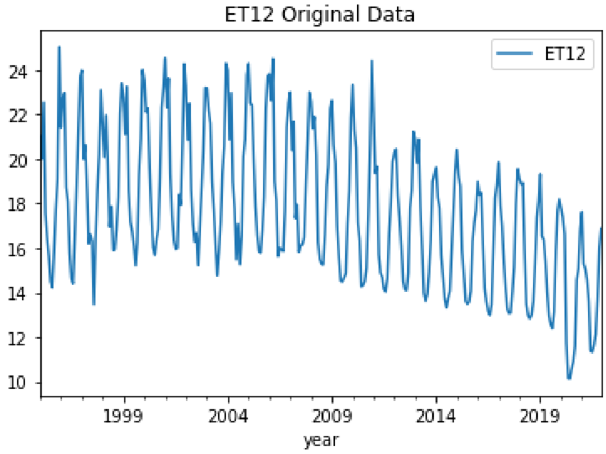


Figure 1.1.4.1

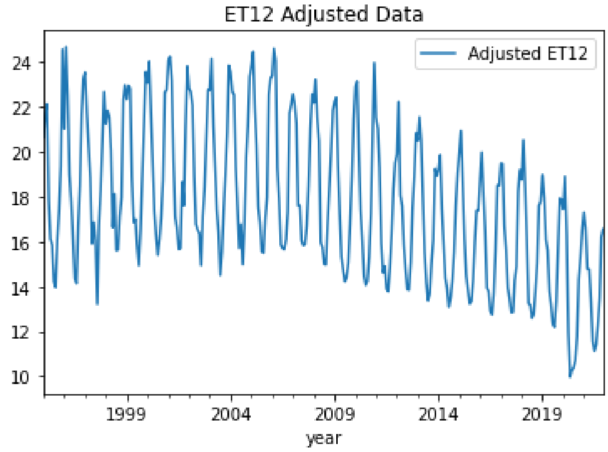


Figure 1.1.4.4

Figure 1.1.1.3

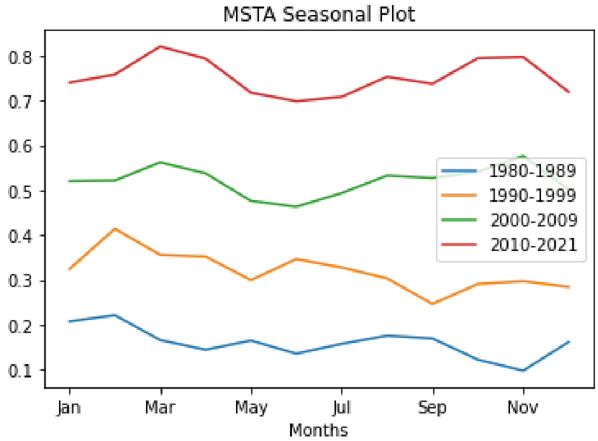


Figure 1.3.1.3

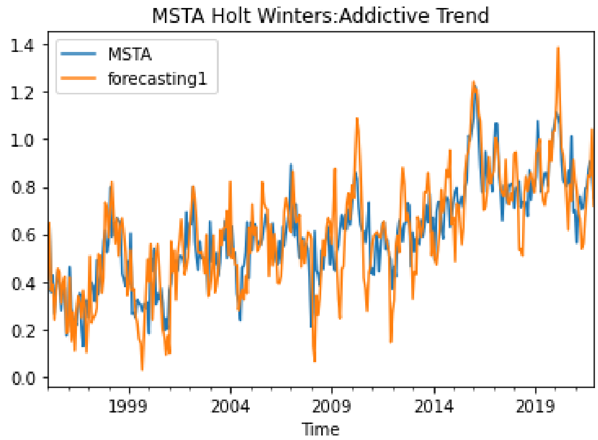


Figure 1.3.1.4

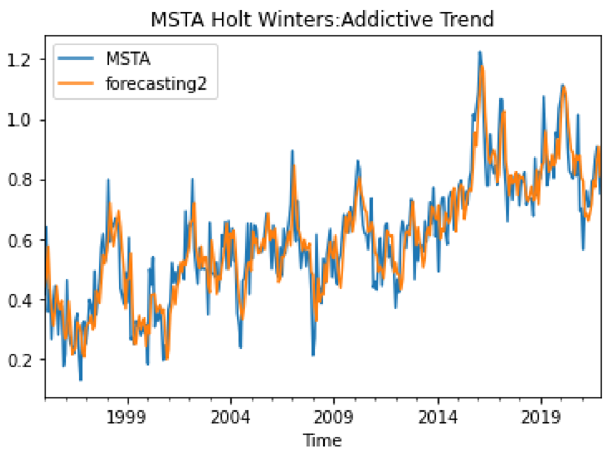


Figure 1.3.1.3

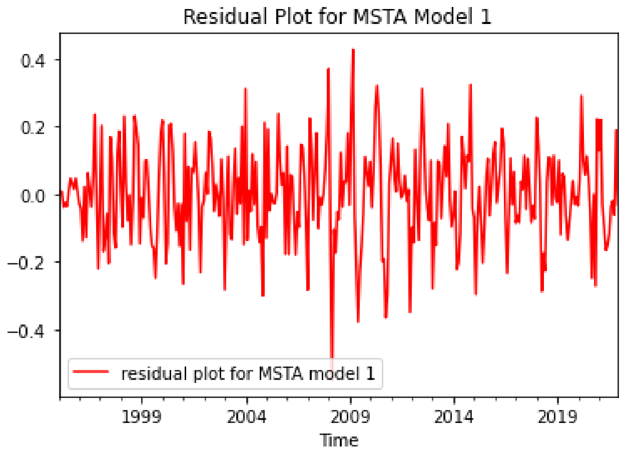


Figure 1.3.1.4

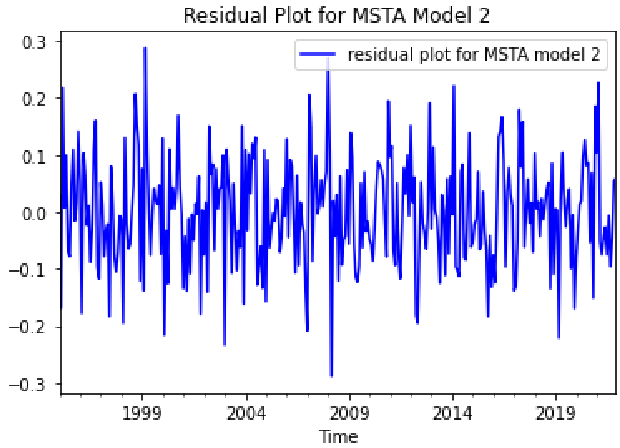


Figure 1.3.1.5

Figure 1.1.1.5

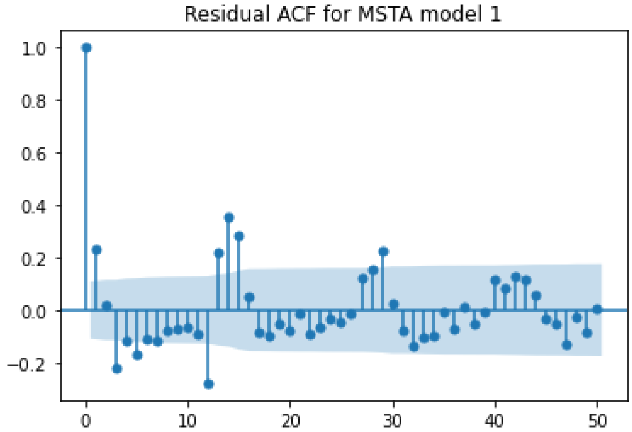


Figure 1.3.1.6

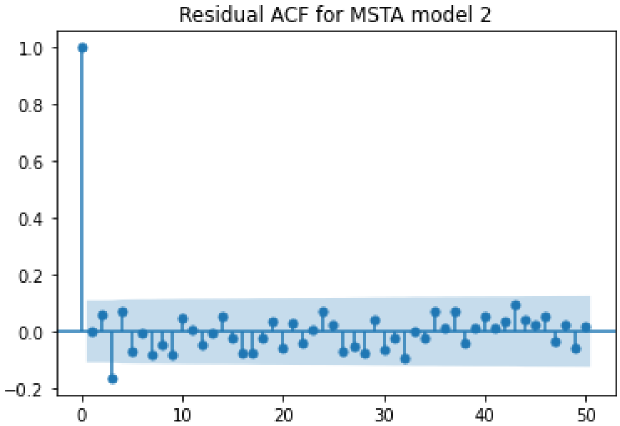


Figure1.3.1.7

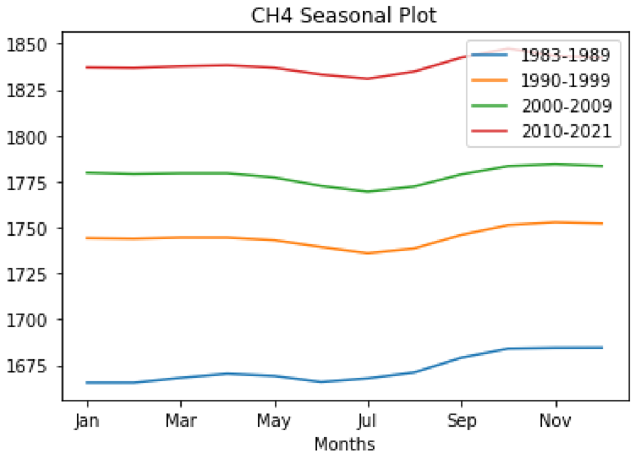


Figure1.3.2.1

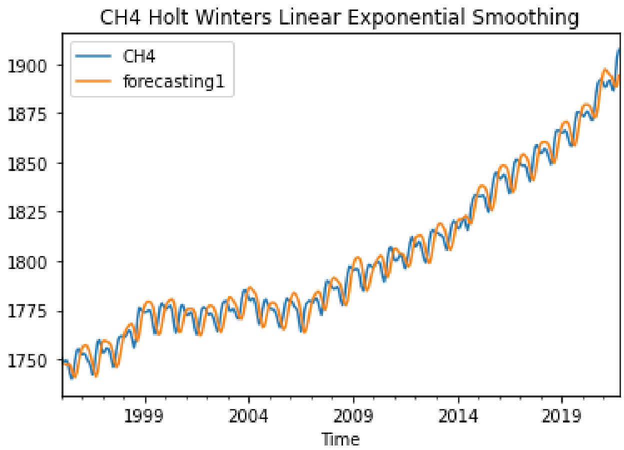


Figure1.3.2.2

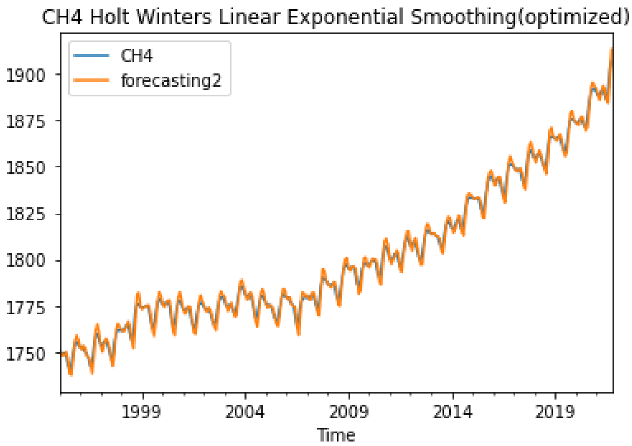


Figure1.3.2.3

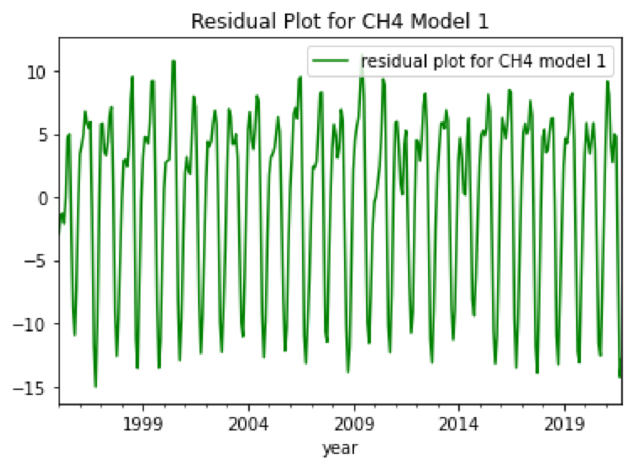


Figure 1.3.2.4

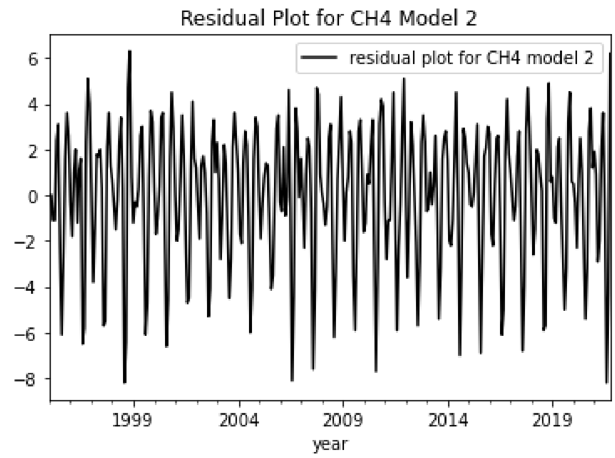


Figure 1.3.2.5

Figure 1.1.1.6

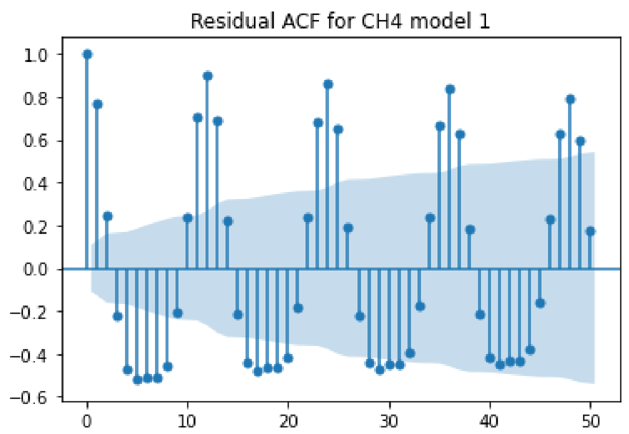


Figure1.3.2.6

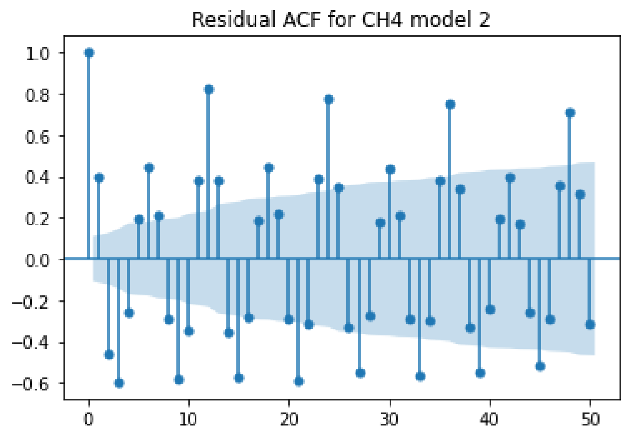


Figure1.3.2.7

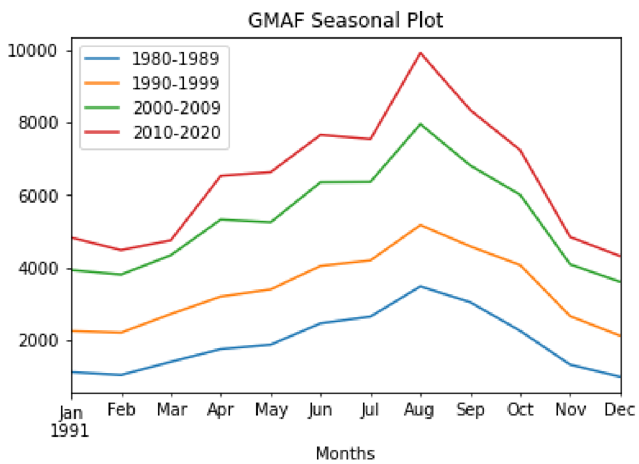


Figure1.3.3.1

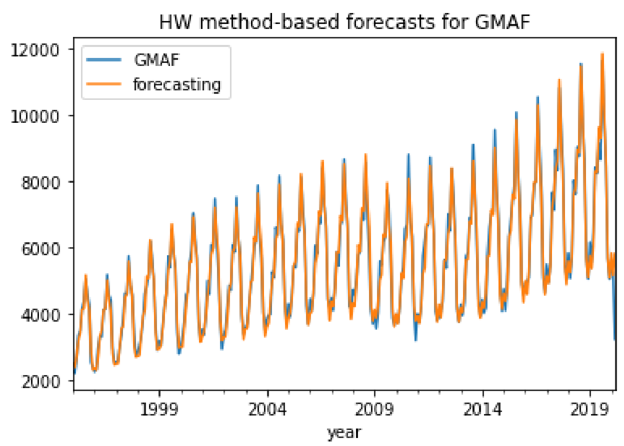
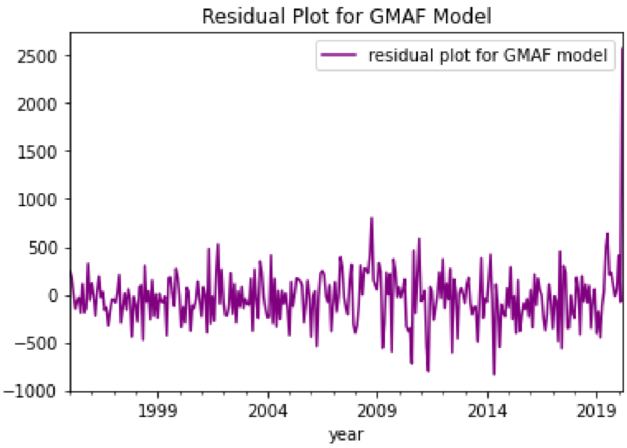


Figure1.3.3.2



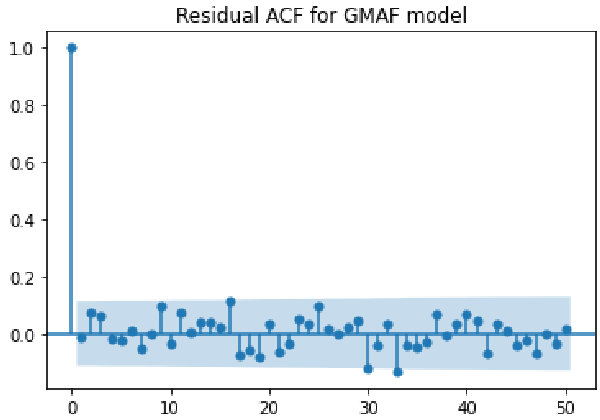
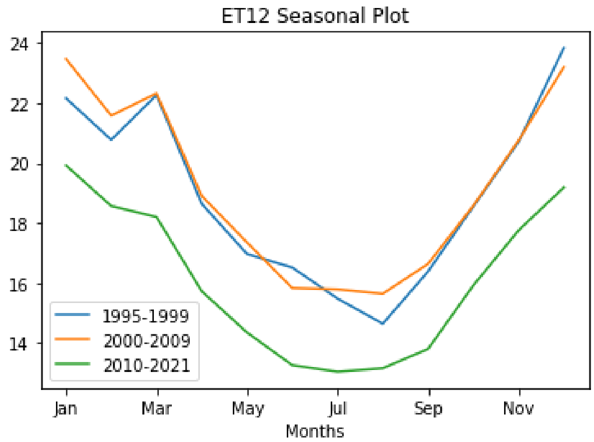


Figure 1.3.3.3

Figure 1.3.3.4

Figure 1.1.1.7



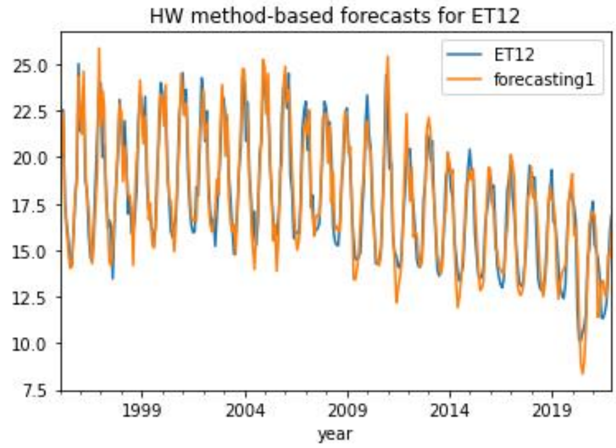
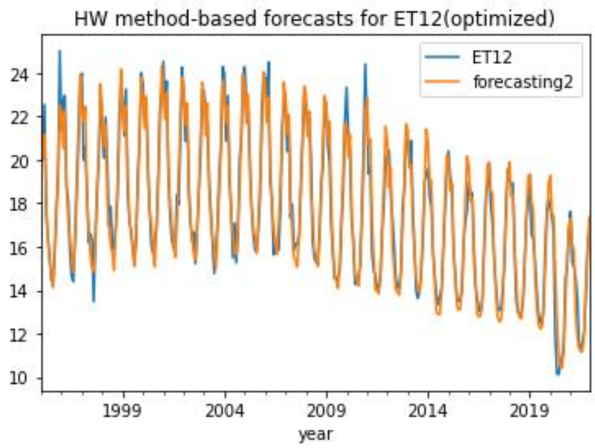


Figure1.3.4.1

Figure1.3.4.2



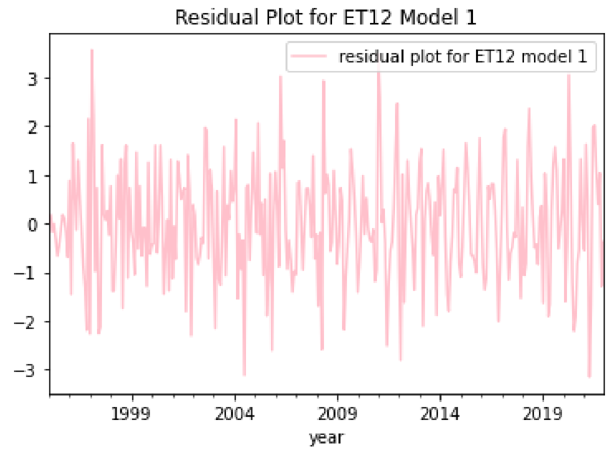


Figure1.3.4.3

Figure1.3.4.4

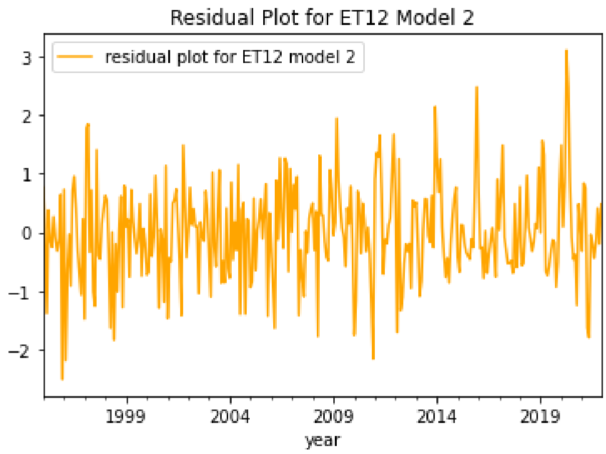


Figure 1.3.4.5

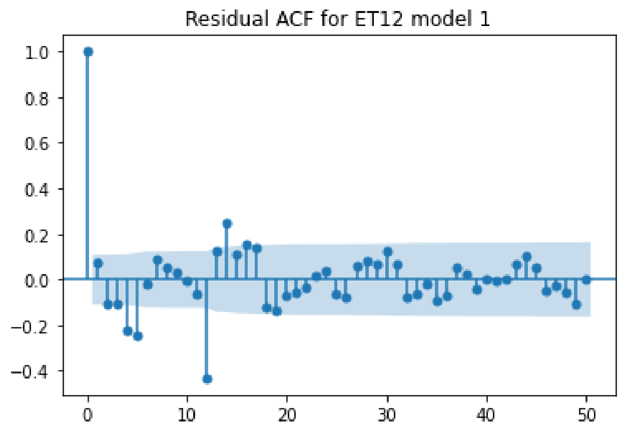


Figure 1.3.4.6

Figure 1.1.1.8

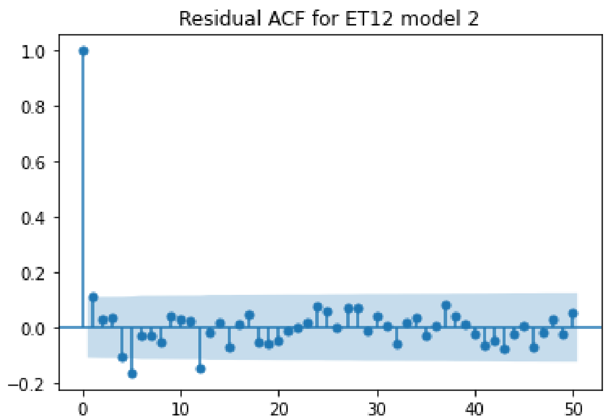


Figure1.3.4.7

Part -2 ARIMA

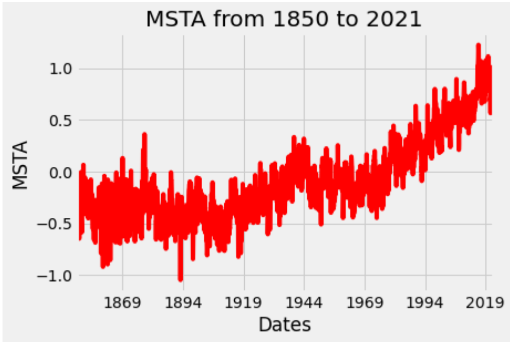


Figure 2.1 timeplot of the MSTA

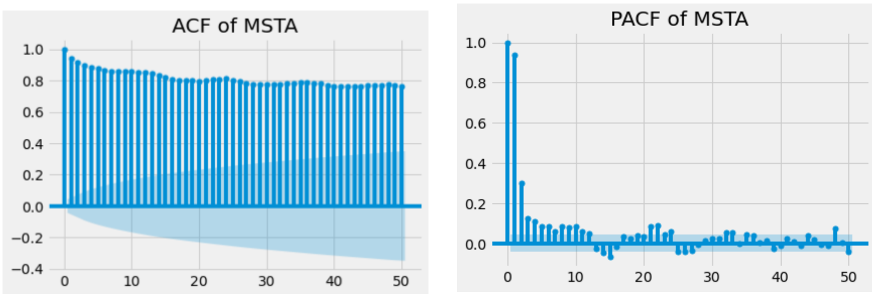


Figure 2.2 ACF plot and PACF plot

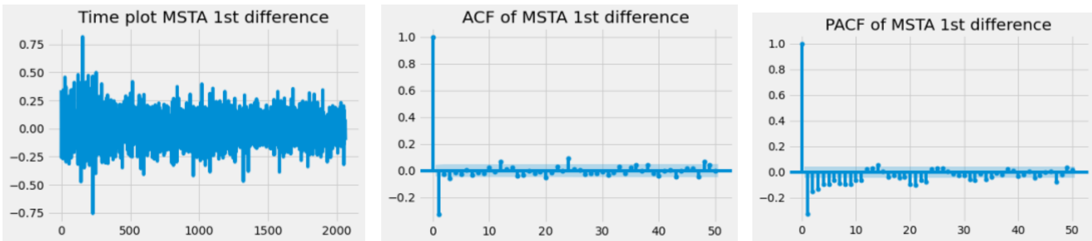


Figure 2.3 time, ACF and PACF plot for 1st order differencing

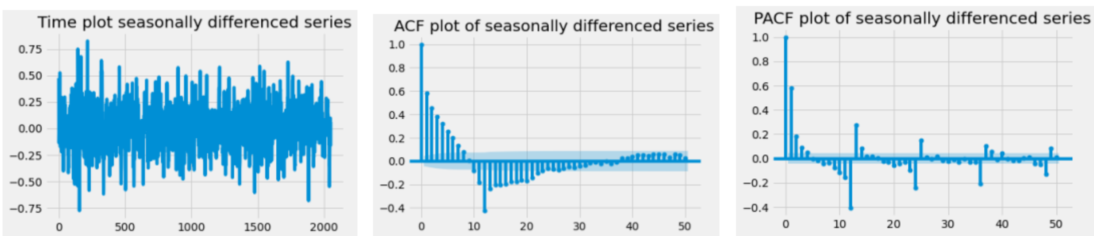


Figure 2.4 Time, ACF, and PACF plots for the seasonally differenced series

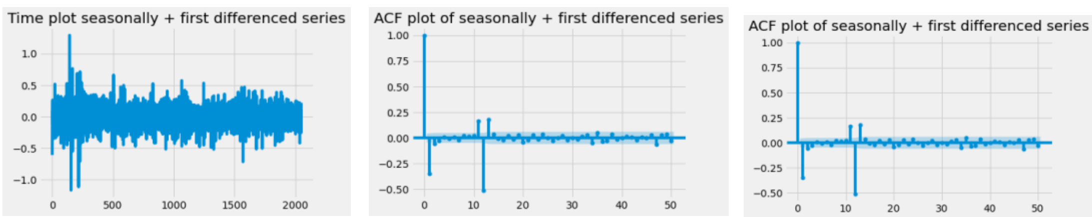


Figure 2.5 Time, ACF, and PACF plots for the Seasonal + First difference

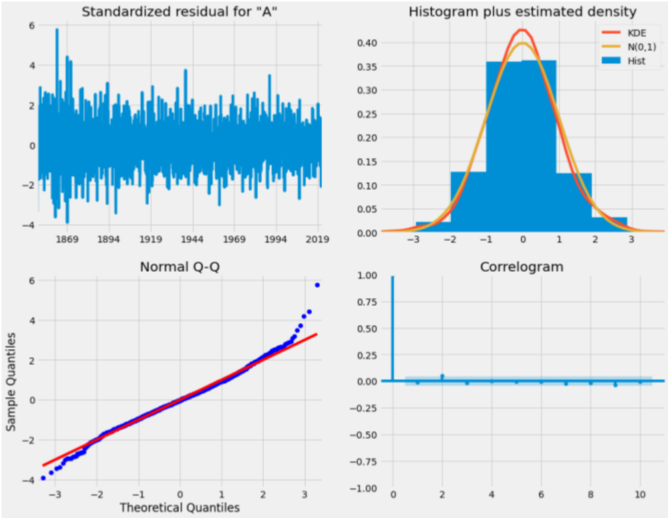


Figure 2.6 the graphical statistics of model ARIMA(1, 1, 1)x(1, 1, 1, 12)

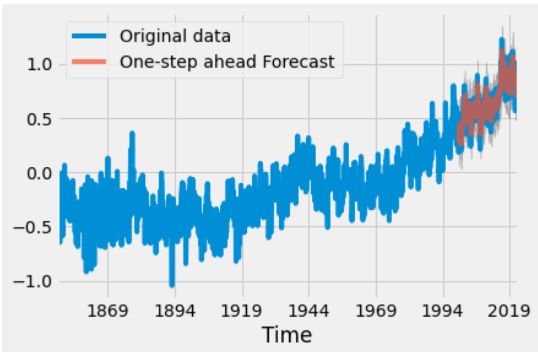


Figure 2.7one-step ahead forecasts together with the original data set of model ARIMA(1, 1, 1)x(1, 1, 1, 12)

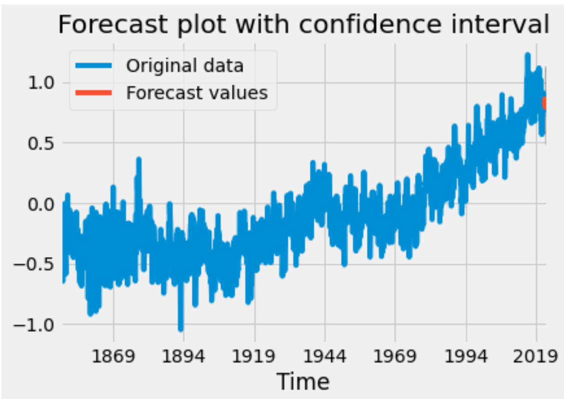


Figure2.8 forecast 12 steps ahead in future of model ARIMA(1, 1, 1)x(1, 1, 1, 12)

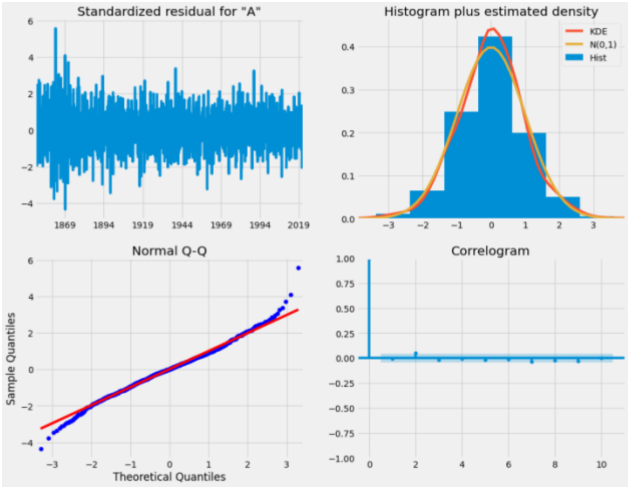


Figure 2.9 the graphical statistics of model ARIMA(1, 1, 1)x(1, 0, 1, 12)

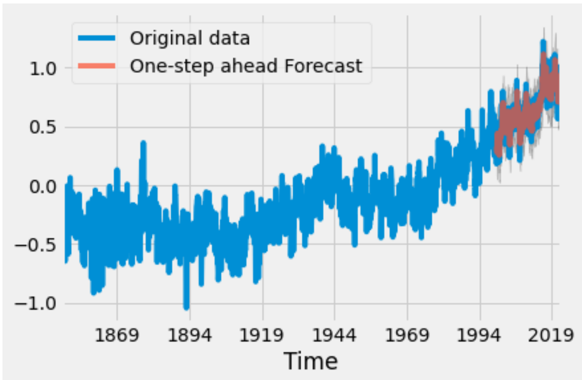


Figure 2.10 one-step ahead forecasts together with the original data set of model ARIMA(1, 1, 1)x(1, 0, 1, 12)

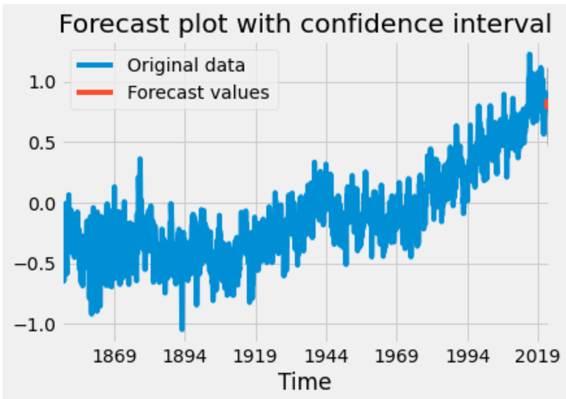


Figure 2.11 forecast 12 steps ahead in future of model ARIMA(1, 1, 1)x(1, 0, 1, 12)

Part-3 Regression

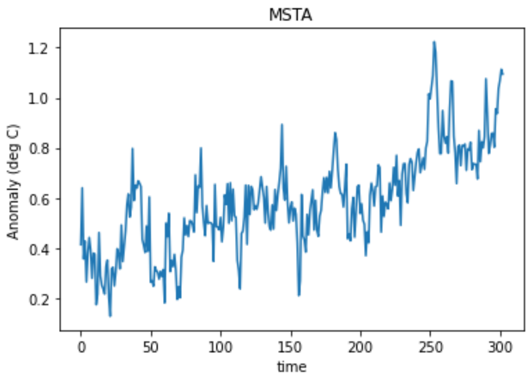


Figure 3.2.1

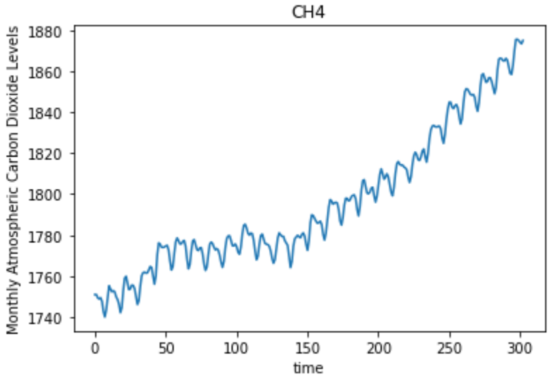


Figure 3.2.2

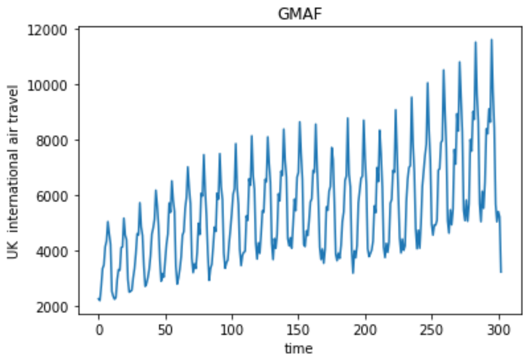


Figure 3.2.3

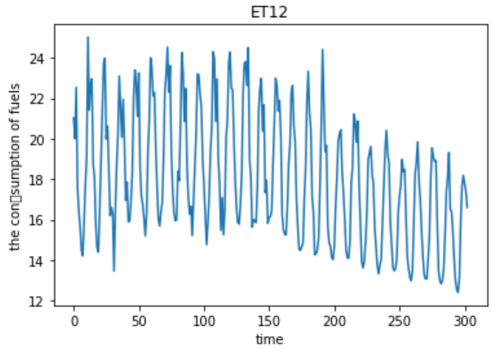
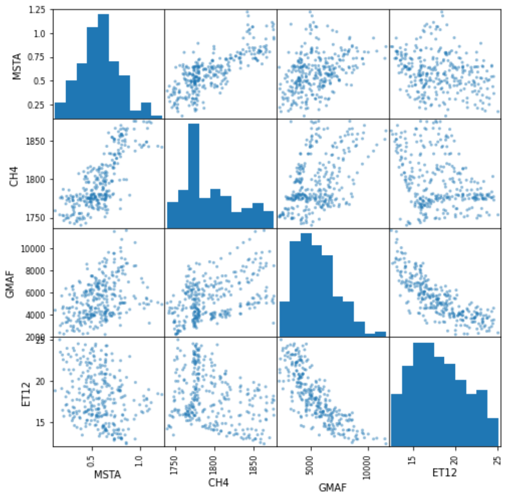


Figure 3.2.4



	MSTA	CH4	GMAF	ET12
MSTA	1.000000	0.770224	0.384289	-0.269358
CH4	0.770224	1.000000	0.472221	-0.327933
GMAF	0.384289	0.472221	1.000000	-0.803479
ET12	-0.269358	-0.327933	-0.803479	1.000000

Figure 3.2.5

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   MSTA   R-squared:                       0.594
Model:                            OLS   Adj. R-squared:                  0.590
Method:                 Least Squares   F-statistic:                     145.7
Date:                Sun, 20 Mar 2022   Prob (F-statistic):           3.44e-58
Time:                        21:35:40   Log-Likelihood:                 188.65
No. Observations:                 303   AIC:                            -369.3
Df Residuals:                     299   BIC:                            -354.5
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -7.5031      0.438    -17.114      0.000      -8.366      -6.640
CH4            0.0045      0.000     18.032      0.000       0.004       0.005
GMAF        3.046e-06   7.41e-06      0.411      0.681   -1.15e-05    1.76e-05
ET12        7.402e-05      0.004      0.018      0.986      -0.008       0.008
==============================================================================
Omnibus:                        4.841   Durbin-Watson:                   0.712
Prob(Omnibus):                  0.089   Jarque-Bera (JB):                4.611
Skew:                           0.249   Prob(JB):                       0.0997
Kurtosis:                       3.343   Cond. No.                     3.51e+05
==============================================================================
```

Figure 3.3.1

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   MSTA   R-squared:                       0.619
Model:                            OLS   Adj. R-squared:                  0.601
Method:                 Least Squares   F-statistic:                     33.44
Date:                Sun, 20 Mar 2022   Prob (F-statistic):           3.50e-52
Time:                        21:49:13   Log-Likelihood:                 198.42
No. Observations:                 303   AIC:                            -366.8
Df Residuals:                     288   BIC:                            -311.1
Df Model:                          14
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -5.2955      1.354     -3.912      0.000      -7.960      -2.631
CH4            0.0034      0.001      4.894      0.000       0.002       0.005
GMAF        2.201e-05   1.43e-05      1.535      0.126   -6.22e-06    5.02e-05
ET12          -0.0142      0.009     -1.629      0.104      -0.031       0.003
D1             0.0261      0.037      0.707      0.480      -0.047       0.099
D2             0.0610      0.040      1.518      0.130      -0.018       0.140
D3             0.0590      0.040      1.460      0.145      -0.021       0.139
D4            -0.0360      0.067     -0.538      0.591      -0.168       0.096
D5            -0.1172      0.079     -1.481      0.140      -0.273       0.039
D6            -0.1376      0.100     -1.380      0.169      -0.334       0.059
D7            -0.1170      0.104     -1.128      0.260      -0.321       0.087
D8            -0.1311      0.123     -1.064      0.288      -0.374       0.112
D9            -0.1457      0.098     -1.485      0.139      -0.339       0.047
D10           -0.0759      0.072     -1.049      0.295      -0.218       0.066
D11            0.0100      0.043      0.234      0.815      -0.074       0.094
==============================================================================
Omnibus:                        1.912   Durbin-Watson:                   0.680
Prob(Omnibus):                  0.384   Jarque-Bera (JB):                1.616
Skew:                           0.145   Prob(JB):                        0.446
Kurtosis:                       3.211   Cond. No.                     1.11e+06
```

Figure 3.3.2

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   MSTA   R-squared:                       0.624
Model:                            OLS   Adj. R-squared:                  0.604
Method:                 Least Squares   F-statistic:                     31.73
Date:                Sun, 20 Mar 2022   Prob (F-statistic):           3.69e-52
Time:                        22:11:41   Log-Likelihood:                 200.29
No. Observations:                 303   AIC:                            -368.6
Df Residuals:                     287   BIC:                            -309.2
Df Model:                          15
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -3.5331      1.639     -2.155      0.032      -6.760      -0.306
CH4            0.0024      0.001      2.684      0.008       0.001       0.004
GMAF        1.067e-05   1.55e-05      0.689      0.492   -1.98e-05    4.12e-05
ET12          -0.0125      0.009     -1.431      0.154      -0.030       0.005
D1             0.0297      0.037      0.807      0.420      -0.043       0.102
D2             0.0643      0.040      1.605      0.109      -0.015       0.143
D3             0.0663      0.040      1.640      0.102      -0.013       0.146
D4            -0.0105      0.068     -0.154      0.877      -0.144       0.123
D5            -0.0904      0.080     -1.130      0.259      -0.248       0.067
D6            -0.1025      0.101     -1.014      0.311      -0.301       0.096
D7            -0.0854      0.105     -0.817      0.415      -0.291       0.120
D8            -0.0765      0.126     -0.607      0.545      -0.325       0.172
D9            -0.1002      0.101     -0.995      0.321      -0.298       0.098
D10           -0.0388      0.075     -0.521      0.603      -0.186       0.108
D11            0.0214      0.043      0.496      0.620      -0.063       0.106
time           0.0006      0.000      1.888      0.060   -2.47e-05       0.001
==============================================================================
Omnibus:                        3.303   Durbin-Watson:                   0.687
Prob(Omnibus):                  0.192   Jarque-Bera (JB):                3.092
Skew:                           0.175   Prob(JB):                        0.213
Kurtosis:                       3.351   Cond. No.                     1.34e+06
```

Figure 3.3.3

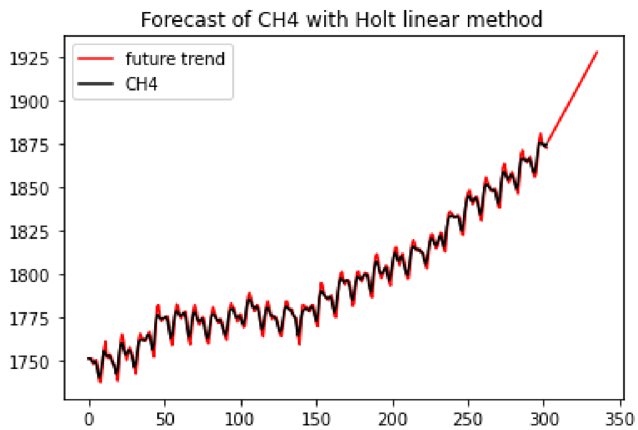


Figure 3.4.1

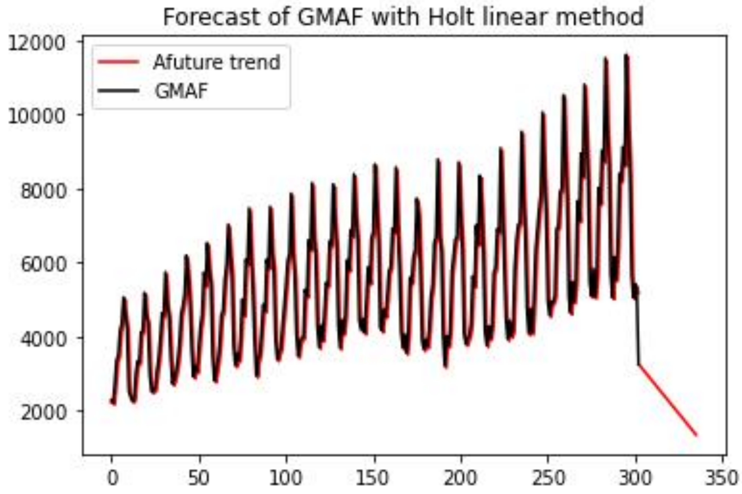


Figure 3.4.2

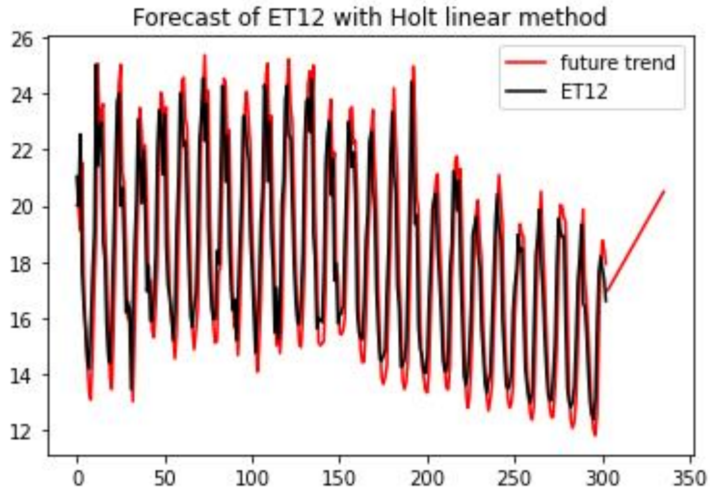


Figure 3.4.3

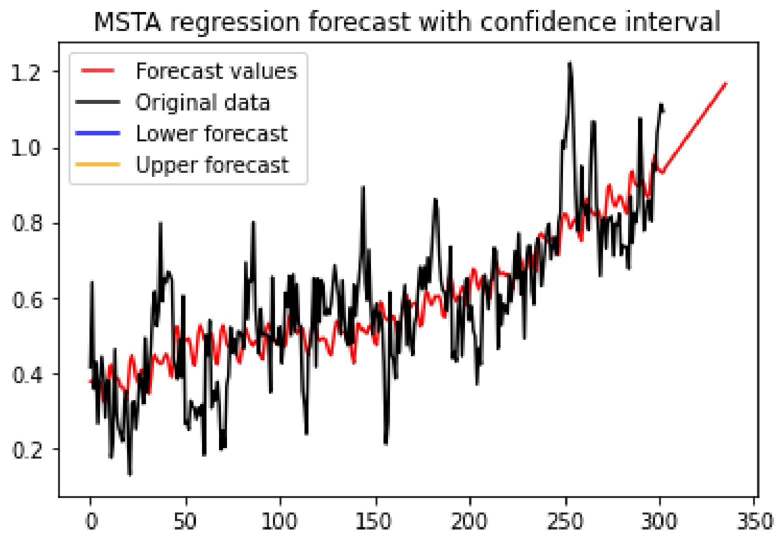


Figure 3.4.4