Time Series Analysis Project

Electricity production from renewable sources &
Greenhouse gas emissions

Andrea Mauro mat.[0222400914]

1 Electricity production from renewable sources

The International Energy Agency (IEA) maintains a large dataset on electricity production from different source and country. In this project the aim is to focus and analyse on the French production of electricity from source renewables(solar, wind, hydropower, geothermal, biomass, marine). The dataset is updated regularly and provides detailed information on the total net generation result from the aggregate of generation from each renewable source.

It represents a valuable resource for researchers, policymakers, and analysts interested to use of green energy. The dataset can be used to track trends over time, compare the performance of different countries, and identify areas for improvement in the use of renewable source as a source of electricity.

1.1 EDA

The dataset used includes data from different parts of the world. However, we have chosen to focus our attention on an area limited to specific European country, particularly decied to consider France, one of major player in the economic world, then look at the transalpine contribution to one of the most crucial challenges of the 20th century: the green transition.

The time series has monthly frequency and composite of 162 observations, whose information refers to the period January 2010 to June 2023. The values reported are measured in GWh(GigaWatt Hours). The original historical series can be found at https://www.iea.org/data-and-statistics/data-product/monthly-electricity-statistics.

The assumption that our time series is a realization of a stationary process is clearly fundamental in time series analysis, so we must first determine whether the series can be considered a realization of a stationary process. In this stage a very useful tool is the graph of the series, many of the patterns in the data can be seen visually.

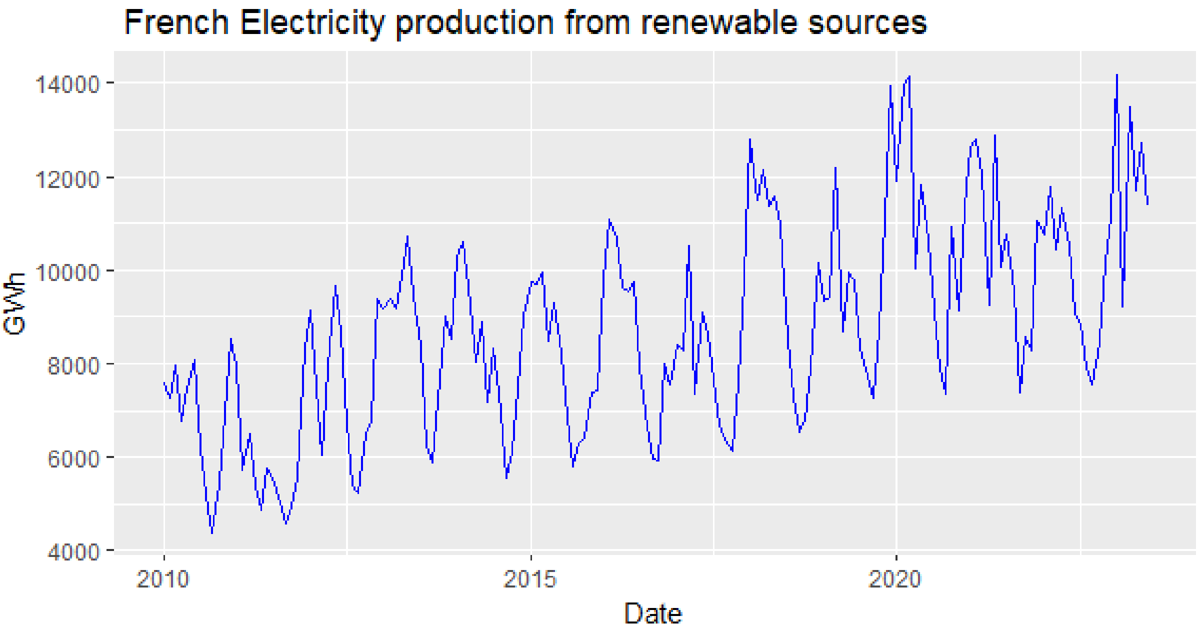
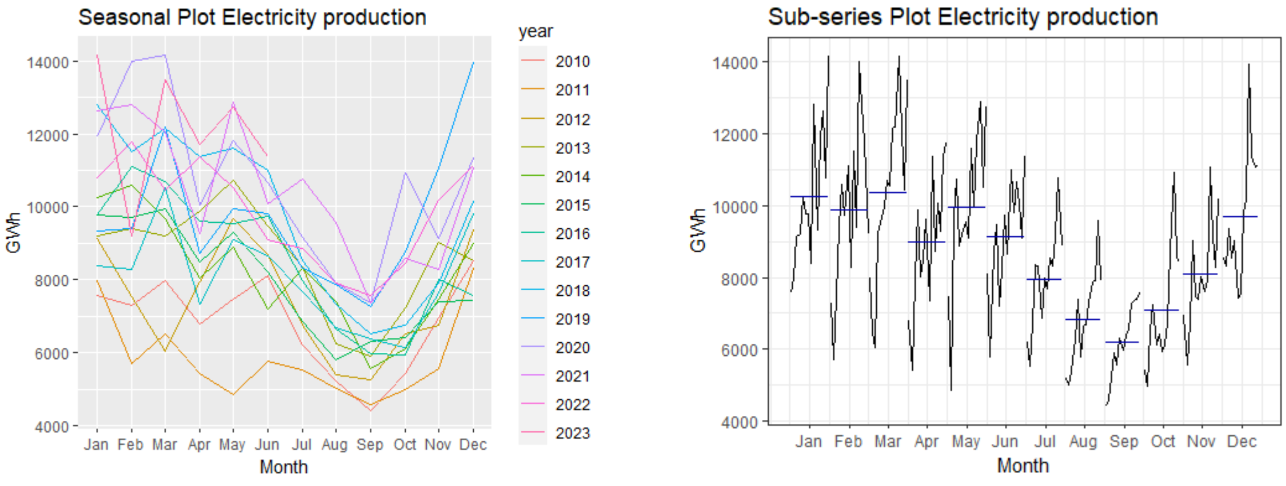


Figure 1: Time plot

In the time series plot (refer to Figure 1) we observe the values related to the production of electricity generated by source renewables, in particular values recorded during the period 2010-2023. The unit of measurement adopted to represent this production is the GWh (GigaWatt Hour).

From the figure we can observe that the energy production has increasing trend, moreover the presence of peaks with regular frequency would seem to suggest a seasonal component. To confirm the presence of the latter we are going to use suitable graphical tools, such as: Seasonal Plot, Subseries plot and Acf.



(a)

(b)

Figure 2: Seasonal Plot(a) and Subseries Plot(b).

In the seasonal plot (refer to Figure 2a) the data are plotted against the individual ”seasons” in which the data were observed making a curve for each year. In the Seasonal subseries plot (refer to Figure 2b) the data for each season are collected together in separate mini time plots and the horizontal lines, indicate the means for each season.

Both graphs show seasonal fluctuations vary proportionally with the level of the series, suggesting an increasing trend, systematic decreases in production are observed during the summer months, this could be related to increase temperatures and the consequent decrease in energy demand needed for domestic heating.

The seasonal assumptions suggested by the previous plots are confirmed in the graph below regarding the autocorrelations of the series lags(ACF).

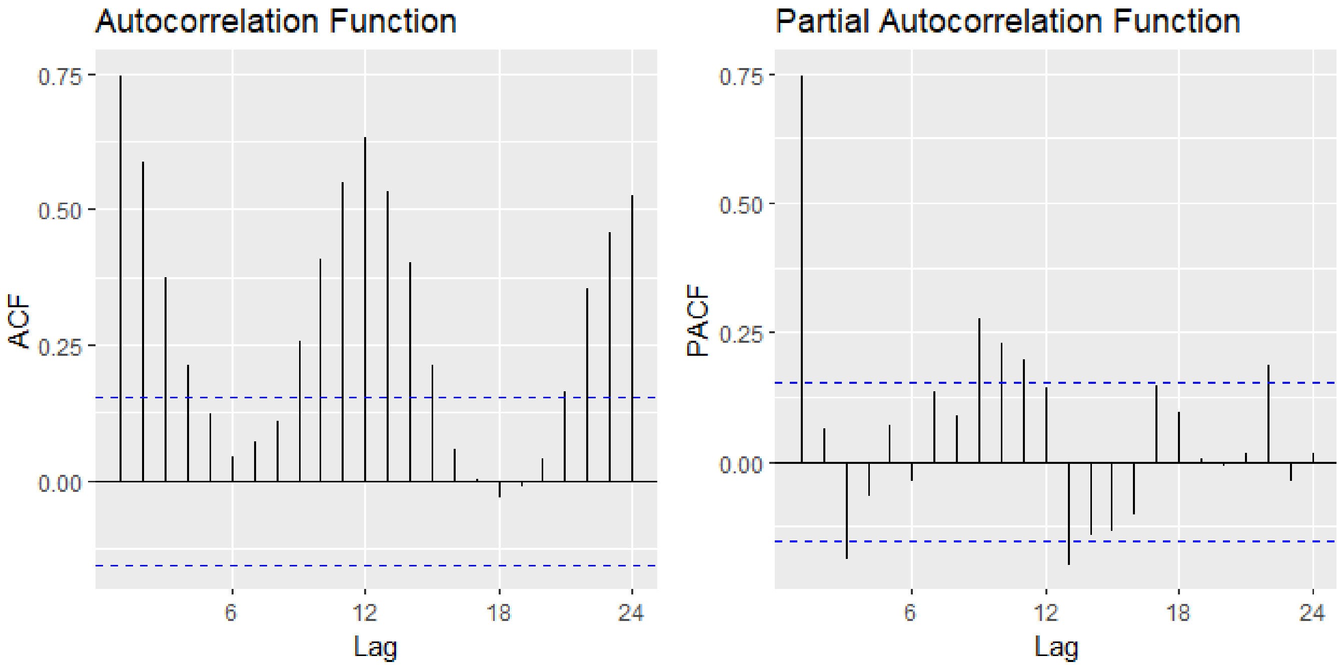


Figure 3: ACF and PACF

Clearly we note from the ACF (refer to Figure 3) how the peaks exhibit a periodicity, in particular we observe how the autocorrelations repeat the same behavior every 12 delays. Furthermore from the ACF we can observe how the autocorrelations, with increasing k, have a very slow decay to zero suggesting the presence of trend. From the analyses so far, it is clear that the time series cannot be considered as realizations of stationary processes.

1.2 Decomposition

To better understand the components and behavior of the series we can use decomposition, that is a procedure which split the original time series into component series:

- *Trend-cycle component* (T_t)
- *Seasonal component* (S_t)
- *Remainder component* (R_t)

Each decomposition methodology seeks to capture these three aspects of a time series, so as to provide an interpretation and to use this information in other analyses. In this study we used the following decomposition methodologies:

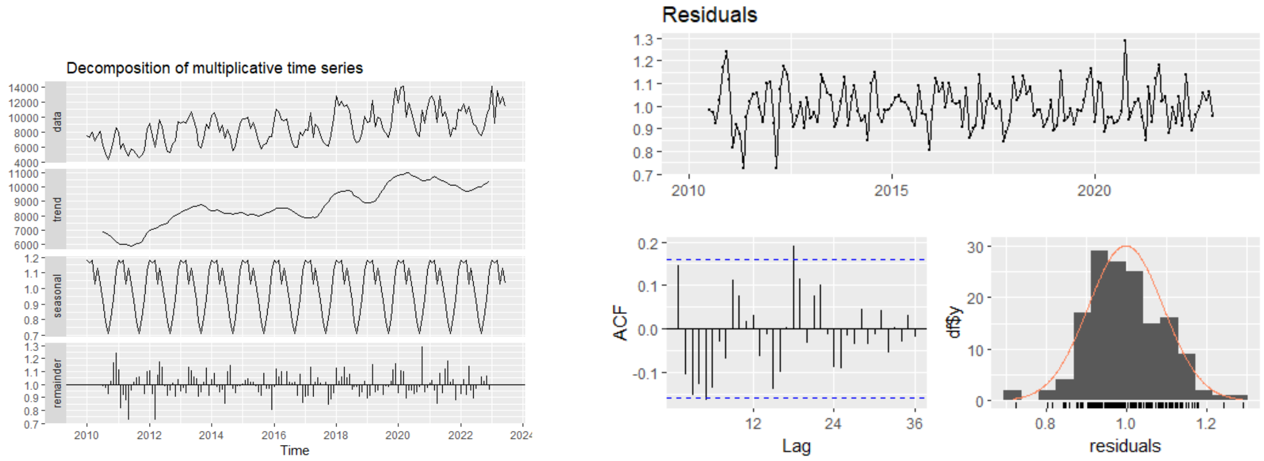
- *Multiplicative Decomposition*
- *STL Decomposition*

1.2.1 Multiplicative Decomposition

The multiplicative decomposition is called ”multiplicative” because the original time series is approximated by the product of these three components:

$$X_t = S_t \times T_t \times R_t$$

This type of decomposition is a classical approach that is easy to apply but have some disadvantages, for example, using moving averages for trend estimation, it is not possible to have values (of T_t and S_t) for the first and last $m/2$ observations.

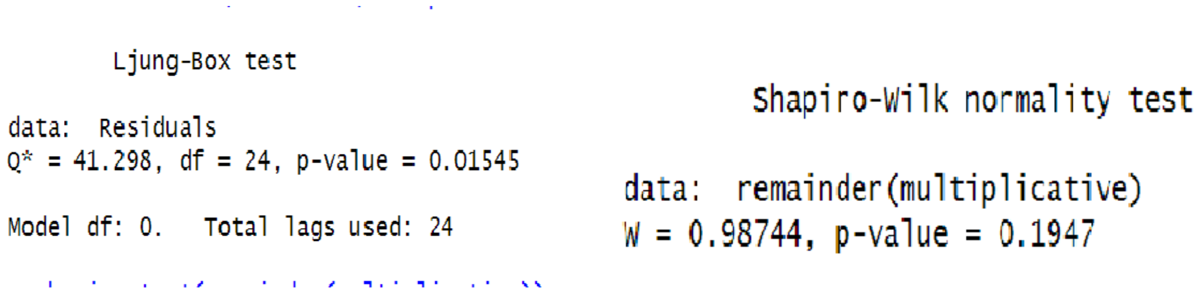


(a)

(b)

Figure 4: Component multiplicative dec.(a) and Residuals Analysis (b).

1. In general, the Trend component(Figure 4a-second sub-plot) confirms what was stated earlier: presence of an increasing trend.
2. Seasonality (Figure 4a-third sub-plot) was captured by decomposition although not in full, in fact we note how the seasonal component does not change in the years of the series (expected result using the moving average to estimate the component).
3. The residuals (Figure 4b) do not have a well-defined structure, and the distribution appears to be approximately normal. This hypothesis is not rejected by the Shapiro-Wilk test of normailty (Figure 5b) with level of significance of test α equal to 5%.
4. The Acf of the residuals shows the presence of some peaks that are different from 0. With the Ljung-Box test the hypothesis that the autocorrelations are jointly equal from 0 is rejected with level of significance of test α equal to 5%).



```
        Ljung-Box test

data:  Residuals
Q* = 41.298, df = 24, p-value = 0.01545

Model df: 0.   Total lags used: 24
```

```
        Shapiro-Wilk normality test

data:  remainder(multiplicative)
W = 0.98744, p-value = 0.1947
```

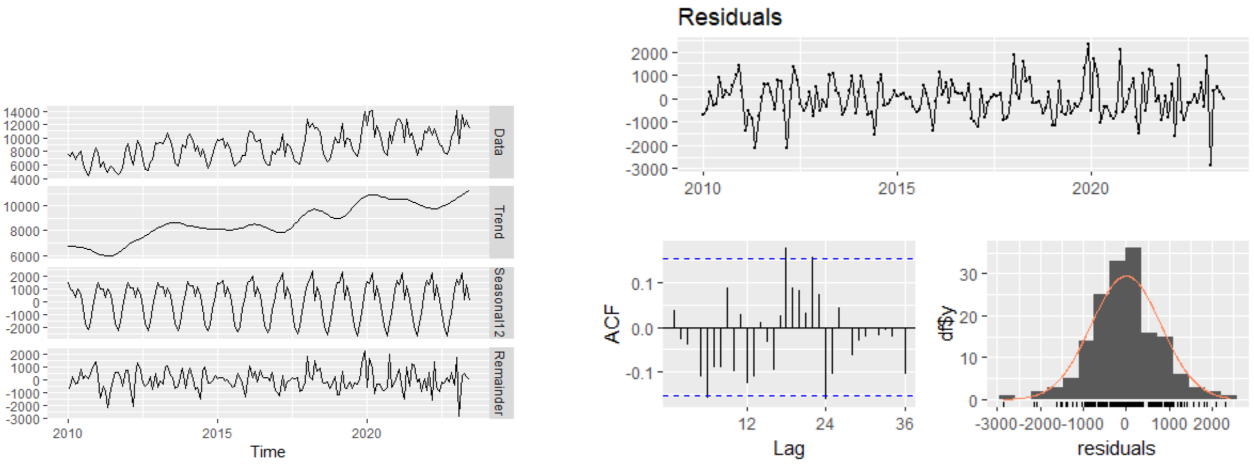
(a) (b)

Figure 5: Test of Ljung-Box(a) and Test of Shapiro-Wilks(b).

1.2.2 STL Decomposition

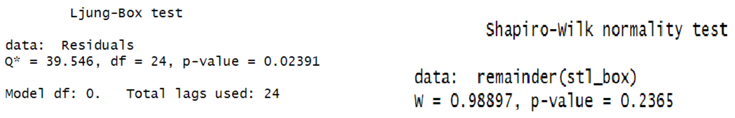
The Seasonal Trend decomposition using Lowess (STL) is one of the most widely-used decomposition methods that can accommodate whatever trend and whatever periodicity. STL decomposition runs two embedded loops:

- the inner loop iterates a user-specified number of iterations in which seasonal smoothing followed by trend smoothing respectively update the seasonal and trend components
- in outer loop iteration, a complete run of the inner loop is followed by the computation of robustness weights. These robustness weights are used to control for aberrant behavior.



(a) (b)

Figure 6: Component STL dec.(a) and Residuals Analysis (b).



```
        Ljung-Box test

data:  Residuals
Q* = 39.546, df = 24, p-value = 0.02391

Model df: 0.   Total lags used: 24
```

```
        Shapiro-Wilk normality test

data:  remainder(stl_box)
W = 0.98897, p-value = 0.2365
```

(a) (b)

Figure 7: Test of Ljung-Box(a) and Test of Shapiro-Wilks(b).

1. The trend component (Figure 6a-second sub-plot) turns out to be similar to that identified through multiplicative decomposition, although in the STL decomposition it is more smooth.

2. Seasonality (Figure 6a-third sub-plot) was captured by decomposition in fact we note how the seasonal component does change throughout development in the years of the series.

3. Like the previous decomposition, the residuals (Figure 11) do not have a well-defined structure, the distribution appears to be approximately normal and they seem to be correlated for some lags in the series. Assumption of normailty has been confirmed with Shapiro-Wilk test due to null hypothesis is not rejected with level of significance of test α level equal to 5%. Further by the Ljung-Box test it is checking assumption that autocorrelations of the first 24 lags are jonitly equal to 0,from the figure 7a, we observe how the null hypothesis is rejected at α level equal 5%, so there is at least one non-zero autocorrelation.

In light of the above analyses, we can state that the time series considered is a realization of non-stationary process.

1.3 ARIMA

The assumption that our time series is a realization of a stationary process is clearly fundamental in time series analysis. The Box-Jenkins methodology requires that the $ARIMA$ process to be used in describing the process generator of data to be both stationary and invertible. Thus, in order to construct an ARIMA model, we must fall under the hypothesis that our time series can be considered a realization of a stationary process. In our case, we have just checked in previous phase that the time series does not fall under this assumption so we must transform the time series in order to get the stationarity.

Detecting Stationary in variance

Fluctuations in the series that increase with time suggests a variance that is not time invariant(heteroschedasticity), which leads to problems in model estimation. In this case we have therefore considered a parametric transformation of the data that allows us to stabilize the variance of the series, the Box-Cox transform, defined as follows:

$$w_t = \begin{cases} log(x_t) & \text{if } \lambda = 0 \\ \frac{(x_t^\lambda - 1)}{\lambda} & \text{if } \lambda \neq 0 \end{cases} \tag{1}$$

The problem is to find λ^* such that the series is as homoschedastic as possible. The choice of the optimal λ value can be made according to different criteria. Among them, we selected the one that makes the standard deviation as invariant as possible with respect to the mean, relative to several subgroups of observations.

This can be observed in the Standard Deviation-Mean plot. In this kind of plot, it is essential to consider the scale on which the standard deviation is graphed. For example, using the range of the s.d. as the scale would make it impossible to visualize the actual dispersion of this from the mean, since the points would be "scattered" in any case. For this reason as the scale for each plot we adopted half the range of the series over which s.d. and mean.

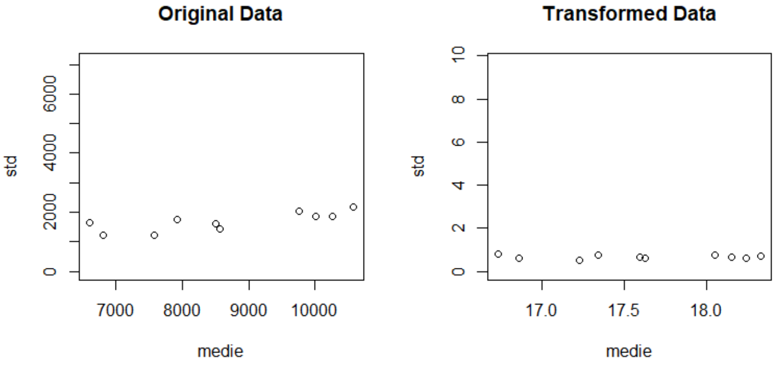


Figure 8: Std-Mean Plot

The λ that optimizes this relationship is the one identified by the method of Guerrero (1993). In this case we have $\lambda* = 0.1342286$.

The series after Box-Cox's transformation is shown below:

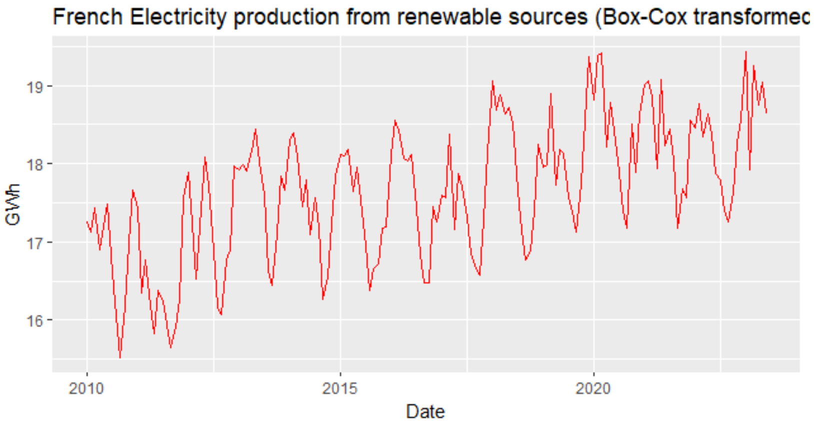


Figure 9: Plot Transformed Time Series

It can be seen that in the Box-Cox transformed series the fluctuations are approximately constant, meaning that the variance is invariant over time. On the basis of what has been asserted, therefore, we can establish that the time series as a result of the transformation performed is stationary in variance.

Detecting Stationary in mean

Based on what was asserted in the previous sections, given the presence of the trend and seasonal component, the time series considered w_t(following the Box-Cox transformation) is non-stationary on mean, differentiation can help stabilise the mean of a time series by removing changes in the level of a time series, and therefore eliminating (or reducing) trend and seasonality.

For this reason, we considered 3 types of differentiations of the series:

- *First-difference*: $\Delta w_t = (1 - L)w_t$
- *Seasonal-difference*: $\Delta_{12} w_t = (1 - L^{12})w_t$
- *Combination between Non-seasonal and seasonal differences*: $\Delta\Delta_{12} w_t = (1 - L)(1 - L^{12})w_t$

The differentiations on the time series w_t are shown below:

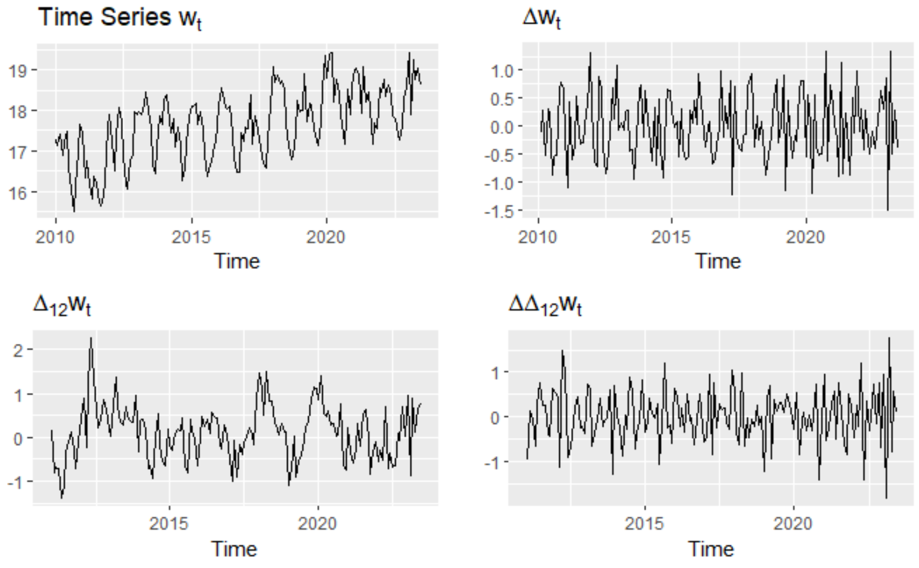


Figure 10: Differentiated Time series

At first look the plot of the produced series we note how there are series that seem to be stable around an average and others for which the same cannot be stated like $\Delta_{12} w_t$. To selection differentiation we exploited two criteria,one more objective,looking at the differentiation that minimizes variance, another one more subjective, which consists in selecting the differentiation that produces an ACF as close as possible to a stationary process.

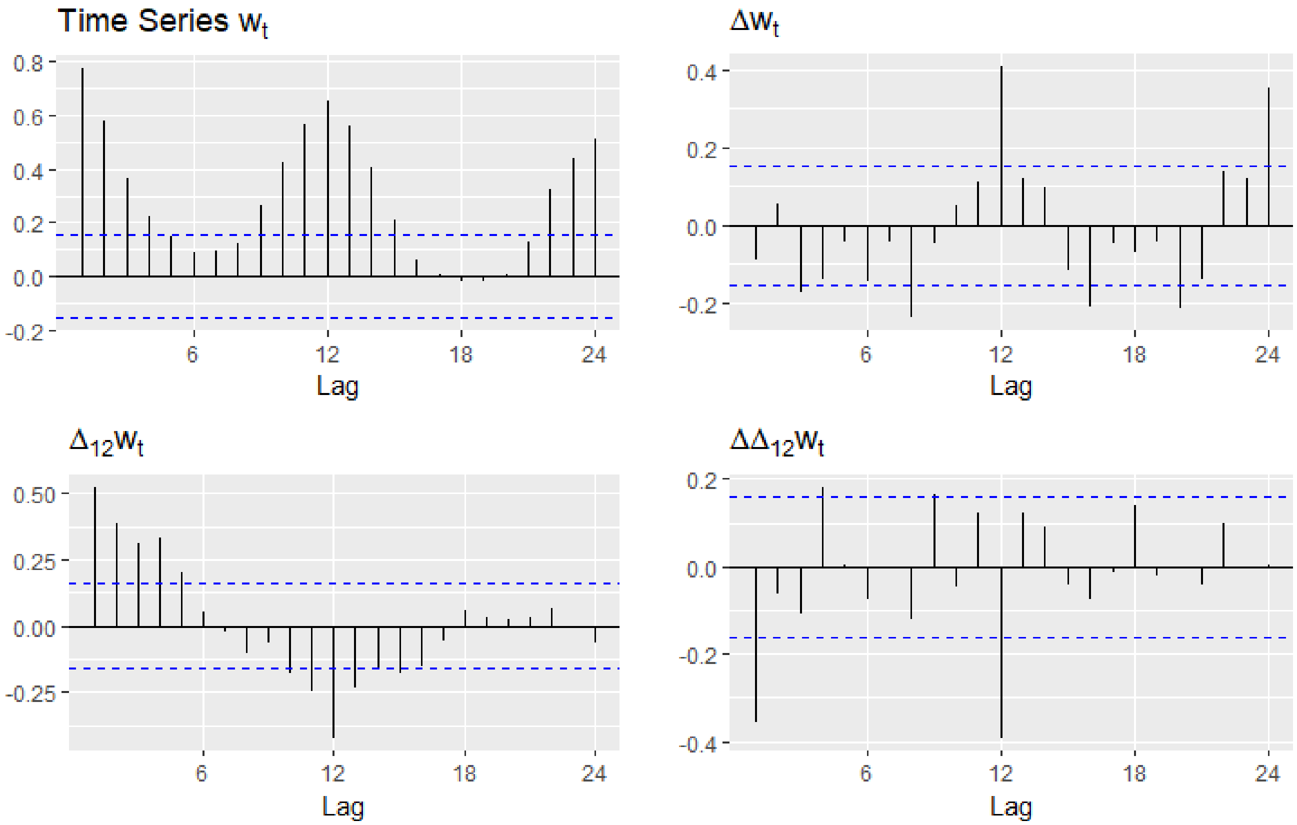


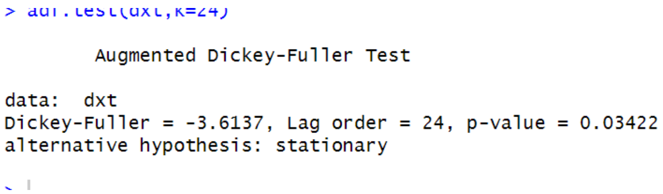
Figure 11: Acf of differentiated Time series

Focus on the Figure 11, we can observe that the autocorrelations of the $\Delta\Delta_{12}w_t$ series and Δw_t series differences appear to be similar those of a stationary process. So from an initial graphical analysis of ACF the choice falls on this types of differentiation.

Series	Variance	Fraction of Variance
w_t	0.75138 5	-
Δw_t	0.3384933	0.4504955
$\Delta_{12}w_t$	0.3894565	0.5183217
$\Delta\Delta_{12}w_t$	0.3694595	0.491708

Looking the objective method, based on the selection of the differentiation which minimizes the variance of the series, the choice falls on Δw_t.

Given the results shown above, both criteria for selecting the differentiation to apply to the w_t series lead to the choice of the First Difference (Δw_t).



```
> adf.test(dxt,k=24)

        Augmented Dickey-Fuller Test

data:  dxt
Dickey-Fuller = -3.6137, Lag order = 24, p-value = 0.03422
alternative hypothesis: stationary
```

Figure 12: Augmented Dickey-Fuller test on Δw_t

To validate the selection of the chosen differentiation, the Augmented Dickey-Fuller (ADF) test is used, in our case, the null hypothesis is rejected for α equal to 5%, so Δw_t is a stationary series.

In the images below there are the time plot for it.

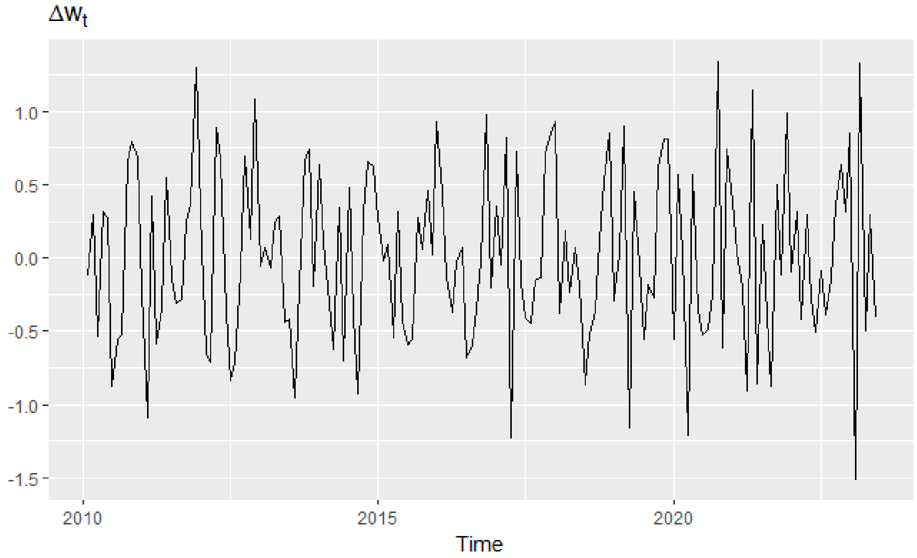


Figure 13: Plot of Δw_t series

Take into account the results show, after a box-cox transformation and apply a differentiation, the time series Δw_t turns out to be stationary in mean and variance.

1.4 Model identification

Given the assumption that our time series $\Delta\ w_t$ is the realization of a stationary process, the next step is model identification. Model Identification involves determining the order of the model required in order to capture the salient dynamic features of the data.
Due to presence of seasonality, we research to identify order of multiplicative seasonal $ARIMA$ model, that could be write like $ARIMA(p,d,q)\times(P,D,Q)_s$, on the basis of what was stated in the previous section the parameter of d(non-seasonal difference) is set to 1, D(seasonal difference) is set to equal to 0 and s set it equal to 12, this is a time span of repeating seasonal pattern.
Therefore our problem of model identification is focus on the multiplicative seasonal $ARIMA(p,1,q)\times(P,0,Q)_{12}$ with mean zero(from now on, due to only models with zero mean were taken into account in the project, it will no longer be specified next to each model 'with zero mean') that it can be written as: $\phi_p(L)\Phi_P(L^s)\Delta w_t = \theta_q(L)\Theta_Q(L^s)\epsilon_t$, where:

- $\phi_p(L)$ = 1- $\phi_1 L$-...-$\phi_p L^p$
- $\Phi_P(L)$ = 1- $\Phi_1 L$-...-$\Phi_P L^{Ps}$
- Δ = (1-L)
- $\theta_q(L)$ = 1- $\theta_1 L$-...-$\theta_q L^q$
- $\Theta_Q(L)$ = 1- $\Theta_1 L$-...-$\Theta_Q L^{Qs}$

Therefore set parameter of seasonal and non-seasonal difference, the problem of model identification focus on parameter of AR and MA for both component seasonal and non-seasonal, to achieve this aim, we use two different way:

- A procedure considered more subjective, based on the ACF and the PACF of the differential series(Δwt), trying to identify any exponential decays or cut-offs in the autocorrelations.
- A procedure considered more objective, based on the estimation of different SARIMA models, whose parameters are different possible combinations of p,q,P,Q.

Graphical method

As mentioned above, this procedure is based on identification of model's parameter by looking the correlogram and the partial partial correlogram of time series Δw_t.

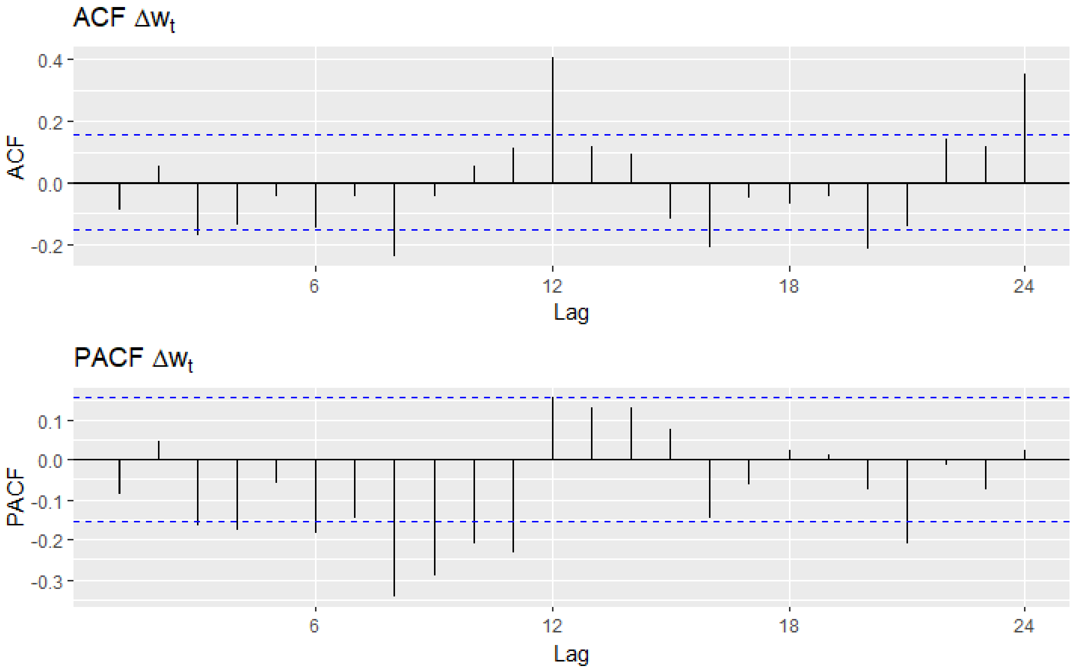


Figure 14: Residuals.

The ACF plot shows the absence of a non-seasonal moving average component ($MA(q)$) due to from the first lags the correlations are not different from 0, so we identify $q = 0$.
Different is the case of seasonal moving average component ($MA(Q)_s$), where a cut-off is observed at the first peak of the seasonal lag, so we identify the parameter $Q = 1$.
About the seasonal and nonseasonal autoregressive components ($AR(P)_s$ and $AR(p)$), looking the PACF, we do not notice cut-offs in the partial autocorrelations at any lag, so we consider the model parameters P and p equal to 0.

Given the above considerations, by the graphical method, we identified the model $ARIMA(0, 1, 0) \times (0, 0, 1)_{12}$.

Combination of parameter

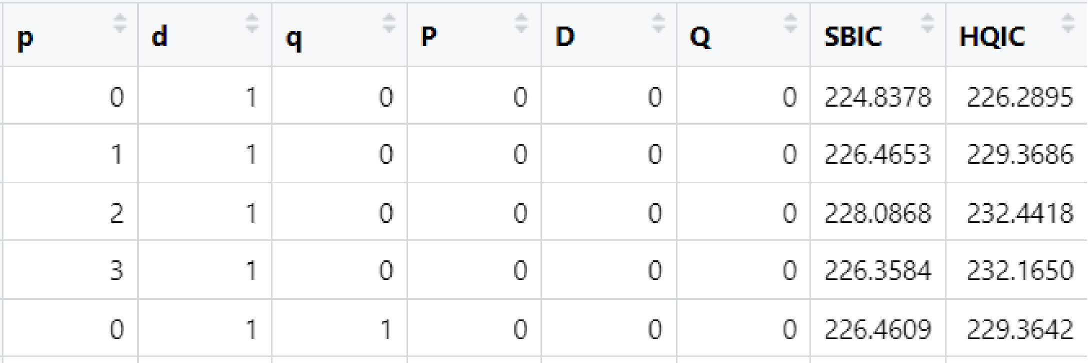
By using this procedure we select parameters of seasonal multiplicative ARIMA taking different combination of them. In particular for the choice of the model it has been chosen an interval of values for p,q,P,Q:

$$
\begin{gathered}
0 \leq p \leq 3 \\
0 \leq q \leq 3 \\
0 \leq P \leq 3 \\
0 \leq Q \leq 3
\end{gathered}
$$

Thereby 256 models were estimated and selection of model has done by looking and minimising two information criteria:

- $The\ Schwarz\ bayesian\ information\ criterion\ (SBIC)$
- $The\ Hannan - Quinn\ information\ criterion\ (HQIC)$

Next figure show different combination of model's parameters $ARIMA(p, 1, q) \times (P, 0, Q)_{12}$ (for brevity is shown first 5 of 256 rows), in particular each row represent a particular combination of parameters (p,P,q,Q), additionaly, last two columns show the value of SBIC and HQIC for each model.



p	d	q	P	D	Q	SBIC	HQIC
0	1	0	0	0	0	224.8378	226.2895
1	1	0	0	0	0	226.4653	229.3686
2	1	0	0	0	0	228.0868	232.4418
3	1	0	0	0	0	226.3584	232.1650
0	1	1	0	0	0	226.4609	229.3642

Figure 15: Table with model's parameter and IC.

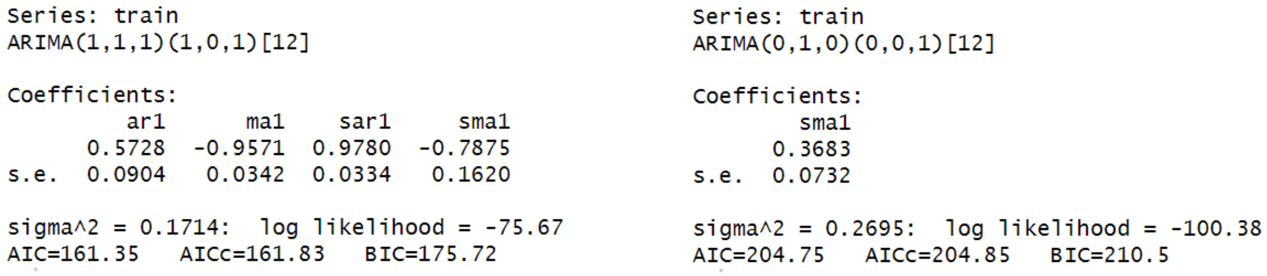
Comparing the two information criteria for all 256 models estimate, the model that turns out to be optimal for both criteria is $ARIMA(1, 1, 1) \times (1, 0, 1)_{12}$ with a $SBIC$=175.7608 and $HQIC$=167.204.

So at the identification step, the models identified are:

- $ARIMA(0, 1, 0) \times (0, 0, 1)_{12}$ (identified by using ACF/PACF)
- $ARIMA(1, 1, 1) \times (1, 0, 1)_{12}$ (identified by using Information Criteria)

1.5 Model Estimation & Checking

Once the models are established, the parameters are estimated using least square method.
The next two figures represent information about the estimates of the two models.



```
Series: train
ARIMA(1,1,1)(1,0,1)[12]

Coefficients:
         ar1      ma1     sar1     sma1
      0.5728  -0.9571   0.9780  -0.7875
s.e.  0.0904   0.0342   0.0334   0.1620

sigma^2 = 0.1714:  log likelihood = -75.67
AIC=161.35   AICc=161.83   BIC=175.72
```

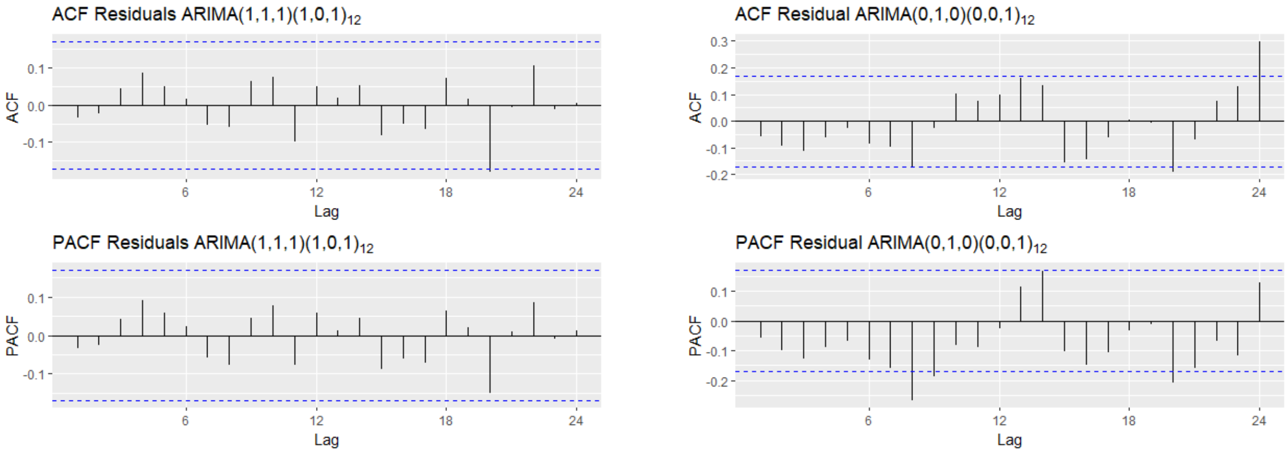
(a)

```
Series: train
ARIMA(0,1,0)(0,0,1)[12]

Coefficients:
        sma1
      0.3683
s.e.  0.0732

sigma^2 = 0.2695:  log likelihood = -100.38
AIC=204.75   AICc=204.85   BIC=210.5
```

(b)

Figure 16: Estimate models

The figure show information about estimate of the $ARIMA(1, 1, 1) \times (1, 0, 1)_{12}$ (refer to Figure 16a) and the model $ARIMA(0, 1, 0) \times (0, 0, 1)_{12}$ (refer to Figure 16b), we observe the coefficients for each estimated parameter of the model components and their standard deviations. In addition, other information such as the estimated residual variance, value of the log-likelihood and information criteria are shown.
After estimation, the next step is to validate the estimated models, in particular we focus on the models's residuals: analyzing the residual ACF and PACF and verifying that they have a white noise structure.



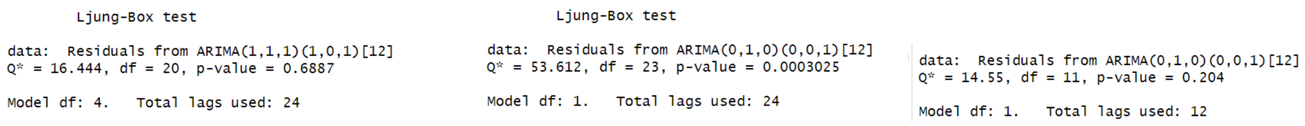
(a)

(b)

Figure 17: Acf and Pacf of model's residuals

Examining the autocorrelation function plot and partial autocorrelation function plot of the residuals from the $ARIMA(1,1,1)(1,0,1)_{12}$ (refer to Figure 17a), they show no one autocorrelations significantly different from zero. This observation lends support to the notion that the residuals follow a white noise distribution.

To further validate the proposition of white noise structure in the residuals of model $ARIMA(1,1,1)(1,0,1)_{12}$, the Ljung-Box test is conducted. The results state that the first 24 lags are jointly equal from 0, for α at the 5% level(refer to Figure 18a), providing additional evidence to the presence of white noise characteristics.



```
        Ljung-Box test

data:  Residuals from ARIMA(1,1,1)(1,0,1)[12]
Q* = 16.444, df = 20, p-value = 0.6887

Model df: 4.   Total lags used: 24
```

(a)

```
        Ljung-Box test

data:  Residuals from ARIMA(0,1,0)(0,0,1)[12]
Q* = 53.612, df = 23, p-value = 0.0003025

Model df: 1.   Total lags used: 24
```

(b)

```
data:  Residuals from ARIMA(0,1,0)(0,0,1)[12]
Q* = 14.55, df = 11, p-value = 0.204

Model df: 1.   Total lags used: 12
```

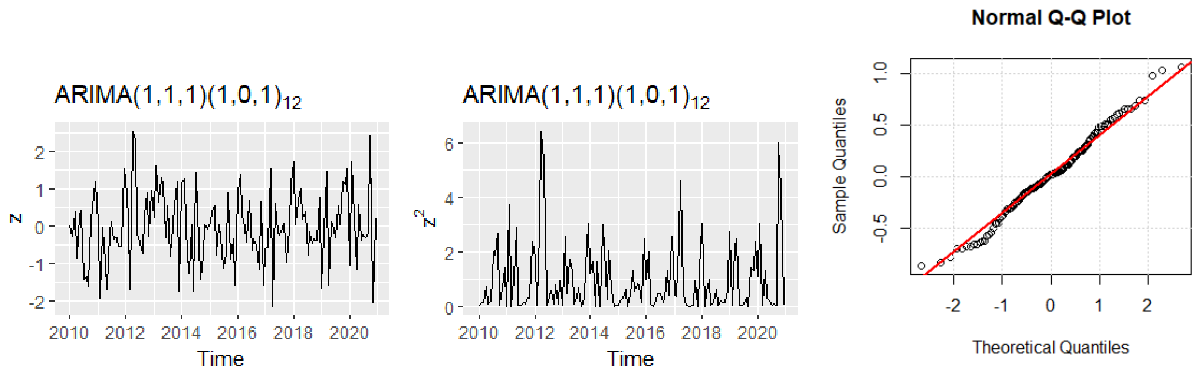
(c)

Figure 18: the Ljung-Box test

By analyzing the graphs of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the model's residual $ARIMA(0,1,0) \times (0,0,1)_{12}$ (refer to Figure 17b),they show some peaks significantly different from zero for high lags.
To investigate the structure of the autocorrelations in the residuals more specifically, the Ljung-Box test was conducted. The results indicate that the assumption of jointly zero autocorrelations for the first 24 lags was rejected for α at the 5% level (refer to Figure 18b). This rejection contradicts the assumption of white noise structure in the residuals, providing evidence against the absence of autocorrelation in the model's residual.

The test of Ljung-Box was also performed considering a lag number lower,precisely 12 (refer to Figur 18c), in this case, the resulting is consistent with the assumption that model's residuals following the White Noise distribution, for α at the level 5%. Consequently, it was decide to not exclude the $ARIMA(0,1,0) \times (0,0,1)_{12}$ model from the analysis that follow.

Next, graphicals analysis about residuals are shown, specifically the plot of standardized residuals,squared standardized residuals, and a plot that comparing the cumulative distribution of the model's residuals with the cumulative distribution of the normal (QQ-plot),

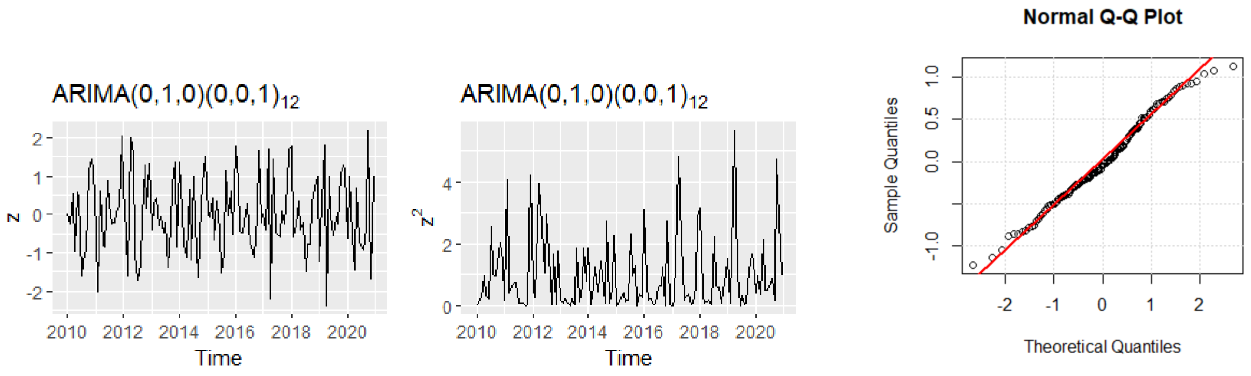


(a) (b)

Figure 19: Residual Analisys of model $ARIMA(1,1,1)(1,0,1)_{12}$

The residuals of the model $ARIMA(1,1,1) \times (1,0,1)_{12}$ appear to be approximately normal; indeed, the standardized residuals mostly fall within the range of [-2, 2] (refer to Figure 19a). By comparing the cumulative distribution of residuals to the cumulative distribution of normal (refer to Figure 19b), clear similarities emerge, affirming the Gaussian nature of the residuals. To provide a comprehensive assessment, the Jarque-Bera test is conducted. The null hypothesis is not rejected at level significance of test α equal 5%, further confirmation about the normality distribution of residuals.

Additionally, the squared standardized residuals (refer to Figure 19a) are examined to detect any potential patterns. As illustrated in the figure, no distinct patterns or clusters are observed, indicating that the model is correctly specified.



(a) (b)

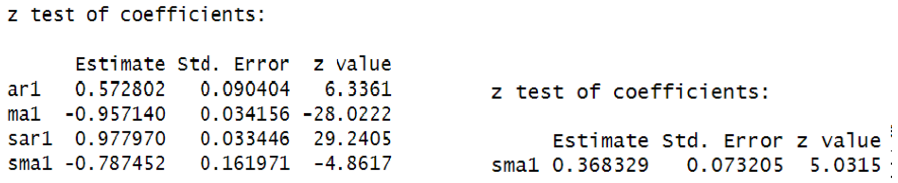
Figure 20: Residual Analisys of model $ARIMA(0,1,0)(0,0,1)_{12}$

The residuals of the model $ARIMA(0,1,0) \times (0,0,1)_{12}$ appear to be approximately normal; indeed, the standardized residuals mostly fall within the range of [-2, 2] (refer to Figure 20a). By comparing the cumulative distribution of residuals to the cumulative distribution normal (refer to Figure 20b), clear similarities emerge, stating the Gaussian nature of the residuals. To provide a comprehensive assessment, the Jarque-Bera test is conducted. The null hypothesis is not rejected at level α equal 5%, further confirmation of normality assumption about model's residuals.

Additionally, the squared standardized residuals (refer to Figure 20b) are examined, to detect any potential patterns. As showed in the figure, no distinct patterns or clusters are observed, indicating that the model is correctly specified.

Being under the assumption that residuals of model $ARIMA(1,1,1)(1,0,1)_{12}$ and $ARIMA(0,1,0)(0,0,1)_{12}$ are normal, we can check the significance of model's parameters.

To assess the significance of the parameters,we calculated the test statistic T for each model parameter. Since T is distributes as a student t with n - (p + q + P+ Q + 1) degrees of freedom, we declare significant at the 5% level the parameters for which $|T_{Coef}| > t_{(n-(p+q+P+Q+1),\frac{\alpha}{2})}$ approximately.



```
z test of coefficients:

     Estimate Std. Error  z value
ar1   0.572802   0.090404   6.3361
ma1  -0.957140   0.034156 -28.0222
sar1  0.977970   0.033446  29.2405
sma1 -0.787452   0.161971  -4.8617
```

(a)

```
z test of coefficients:

     Estimate Std. Error z value
sma1 0.368329   0.073205  5.0315
```

(b)

Figure 21: Test of significance of the coefficients

The test at level $\alpha = 5\%$ confirms the significance of the coefficients for all parameters of model $ARIMA(1,1,1)(1,0,1)_{12}$ and $ARIMA(0,1,0)(0,0,1)_{12}$.

Thus performed all the steps of the Box-Jenkins methodology we can say that the identified and estimated models are adequate for the considered time series, so in the next section through these two models forecast will be developed.

1.6 Forecasting

The dataset has been divided into train(from January 2010 to December 2020 $\sim$ 80% of observation) and test set(from January 2021 to June 2023 $\sim$ 20% of observation). On train set the parameters of the model are estimated,test set is used to valuate the prediction obtained with the model estimated on train.

Before evaluating the predictive accuracy of the model, it is necessary to back transform the forecasts to the original scale of the series and correct them to take into account the bias introduced in the forecasting process. The metrics used to evaluate models are:

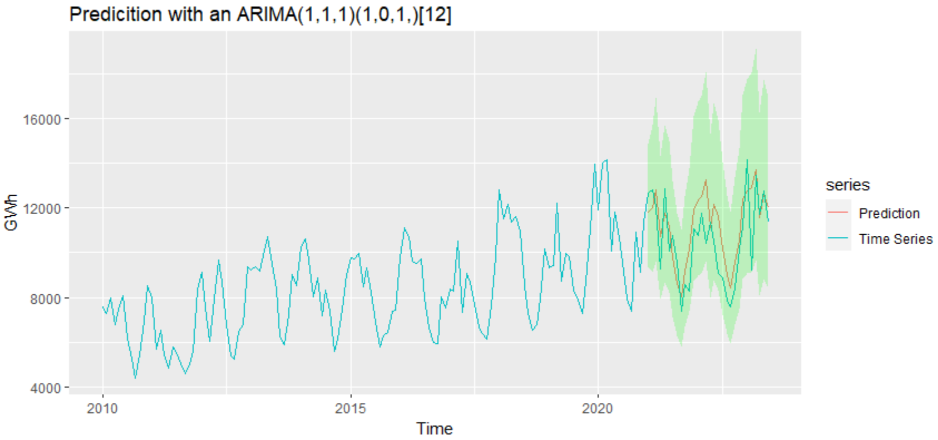
- RMSE
- MAPE
- MAE

Model	**RMSE**	**MAE**	**MAPE**
$ARIMA(1,1,1) \times (1,0,1)_{12}$ **with zero mean**	1442.92	1181.02	11.98
$ARIMA(0,1,0) \times (0,0,1)_{12}$ **with zero mean**	2493.20	1998.107	21.33

Table 1: Metrics

As we can see from the table of metrics, the two models present markedly different performances on the test: in particular, we observe how the model $ARIMA(0,1,0) \times (0,0,1)$ under each metric considered turns out to be worse, if we look at MAPE we can clearly see how the predictions developed by this model deviate on average approximately 21% from the test observation.

Better results are registered with $ARIMA(1,1,1) \times (1,0,1)$, the model develops on average forecasts that deviate by 12% from the test observation, and in general even when considering Root Mean Squared Error and Mean Absolute Error the model record a better forecasting performance.



(a)

Figure 22: Forecast of model $ARIMA(1,1,1)(1,0,1)_{12}$

The forecasting model succeeds in capturing all the main aspects identified in the exploratory stages of analysis. Indeed, the time series is characterized by a strong seasonal component and a slight upward trend, which we see reflected in the forecast as well. In the plot also illustrates the forecast intervals at the level of confidence 95%.