Contents

“Electricity prices predictions”

1. Introduction

In today’s fast-paced economy, the price of electricity is one of the most popular topics worldwide, which is evident due to the fact that it hits the headlines whenever it goes up. However, are these price fluctuations really arbitrary? Changes in prices generally reflect variations in electricity demand, availability of generation sources, fuel costs and power plant availability. Prices are usually highest in the summer when total demand is high because more expensive generation sources are added to meet the increased demand. In this report we will attempt to make a prediction regarding the electricity prices for the first half of 2019 and identify the expected monthly price change.

2. Data preparation

2.1. Description of the dataset

The data that we had in our disposal, contained information about the daily price of the kilowatt hours (kWh) for the period 2010-2018 (21/07/2010 – 31/12/2018), which corresponded to 3,086 different prices. Our aim was to use this data in order to predict the average monthly prices for the first half of 2019 and determine whether the average price during each month will increase or not.

2.2. Data cleansing

Our original dataset is a data frame and at first glance we observed that there were four different date formats used. After setting a unified date format (e.g., 2018-12-31), we saw that there were not any duplicate rows in the data or any infinite values. However, there was one price value that was missing for the "2013-03-11" date. Considering the fact that we will aggregate our data on a monthly basis later on, we would not face any problems if we simply omitted this value, but due to the fact that we have time series data we decided to replace it using the median electricity price of March for the 2013 year. A visualization of all the available daily prices is given in the figure below.

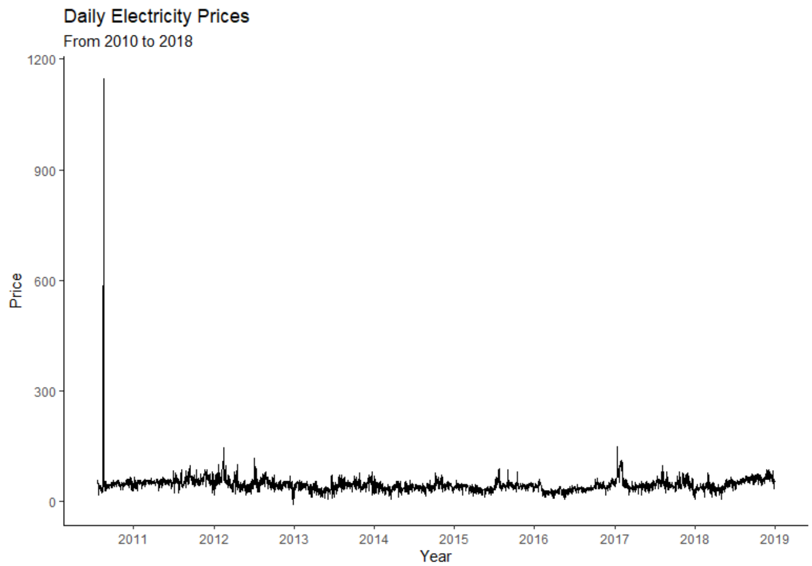

Figure 1: Time series visualization of all the available daily prices in the data

From Figure 1 we observed that during a day in August 2010, the electricity price skyrocketed to over 1,100$ which does not seem to be correct as the next higher price was equal to 150$. In our analysis later on, we will not include any values before 2010, so we do not care about this value. However, for completeness reasons we decided to replace it with the median price of that year's month. We also noticed that there were two values with negative prices during December 2012, but when demand for electricity is low but production is high, electricity prices can go negative, which is why we decided to not take any further actions regarding these values.

As was previously mentioned, the focus of our analysis was to accurately predict the prices for the upcoming first six months of 2019. This is the reason, we decided to aggregate the price from daily price to monthly. The resulted dataset, now contained 102 observations with a year-month format (e.g., 2010-07) and the price value ranged from 25$ to 81$ with a mean price of 46$. In order to take a representative price for each month, we tried some different ways which we will utilize in our models later on to see which one produces the best results (with the lowest RMSE). Specifically, we computed the mean, max, min and median price of each month along with the median price of the lower half of the data set (1st quantile) and the central point that lies between the median price and the highest number of the distribution (3rd quantile) of each month.

Considering that for 2010 the data were available only after September and based on the fact that we are interested for the prices of first half of 2019, we decided to exclude the values for this year. From Figure 2 we saw that there is positive skewness of the histogram for the prices (mean>median) and that the price was predominantly about 40$ with a high number of cases equal to approximately 35$ and 50$. We also noticed from Figure 3 that as years pass by, the prices become more stable, with fewer fluctuations. Specifically, from 2012 to 2015, there is a higher density with price equal to about 40$, while for 2011 there is higher density for the 52$ price. As for 2016 to 2018, there seems to be a more distributed range of prices. These figures, are presented below.

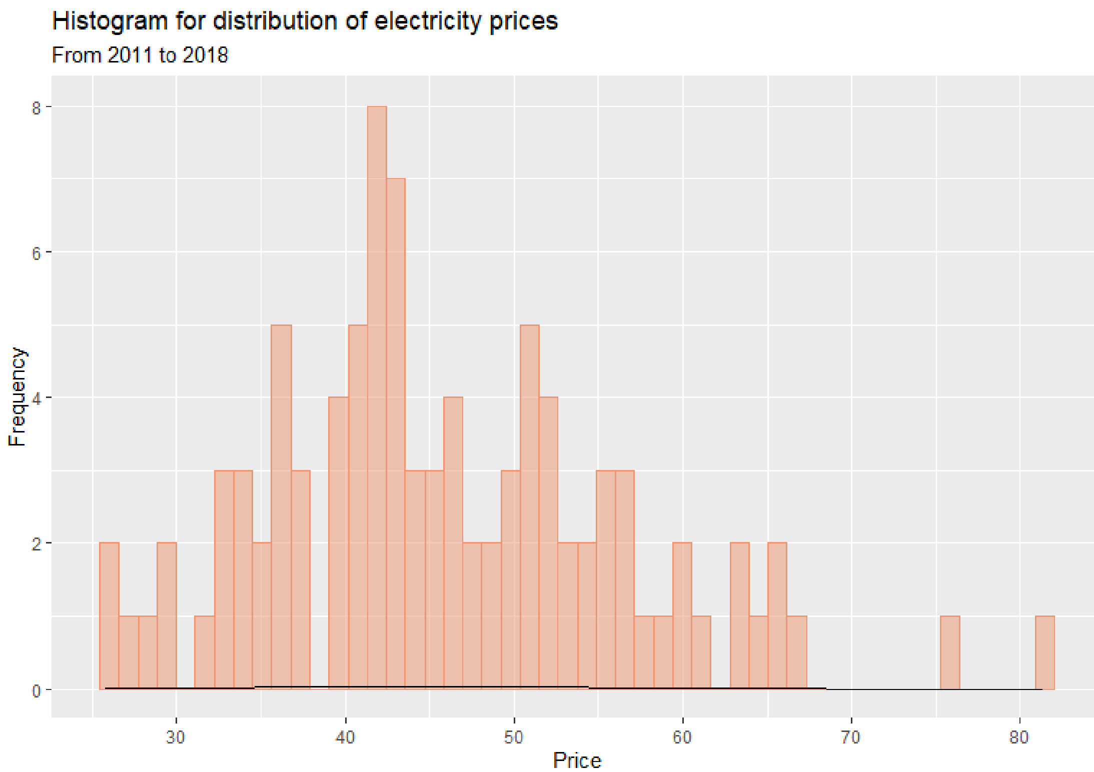

Figure 2: Distribution of electricity prices from 2011 till 2018

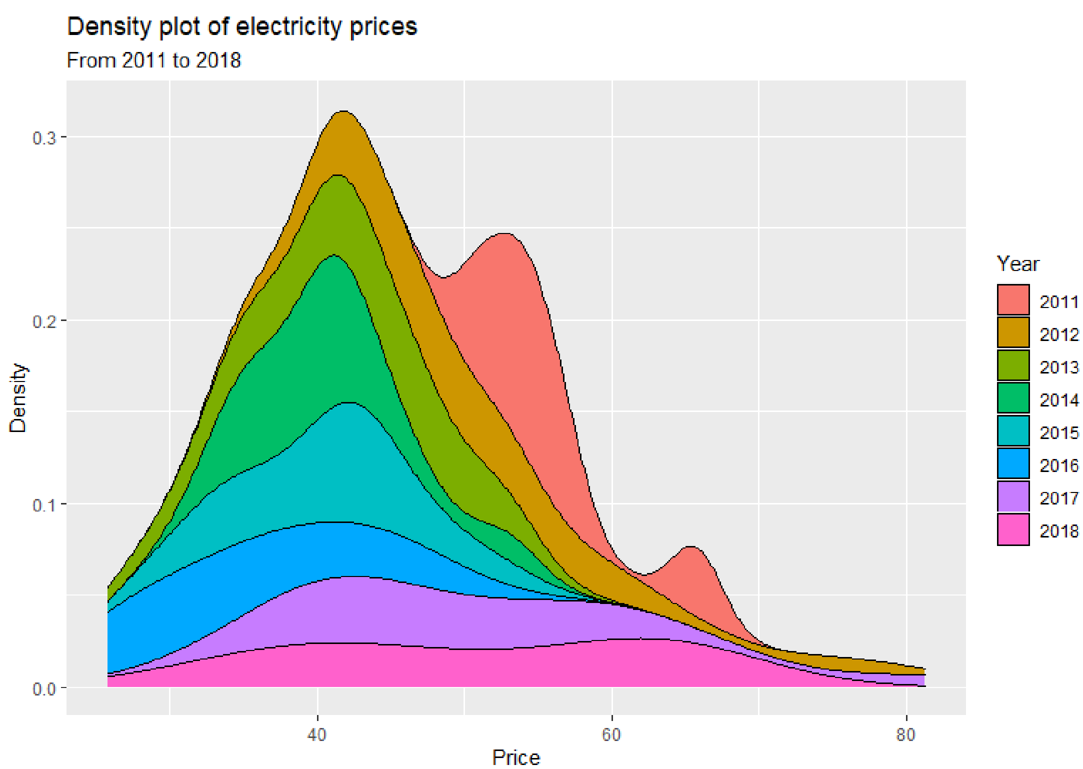

Figure 3: Density plot of electricity prices from 2011 till 2018

So, we saw that the distribution of prices for 2011 differs significantly from the rest of the available years, for which we observed that there is a continuous flattening of the curve (that describes the concentration of price values) as years pass. As a result we decided to only use the data after 2012 for our predictions.

3. Initial Time Series Analysis

3.1. Training / test split

Before beginning our analysis, we need to split the data in training and test datasets. That way we can make out of sample predictions and evaluate each model's performance effectively. We decided to split the data using approximately 70% of the observations as training set, and the remaining 30% as test set. So, the prices from 2012 to 2016 were used as the training set (60 months) and the ones from 2017 to 2018 as the test set (24 months). So, we will continue our analysis based on the training data. Note that, later on we also evaluated forecast accuracy using time series cross-validation.

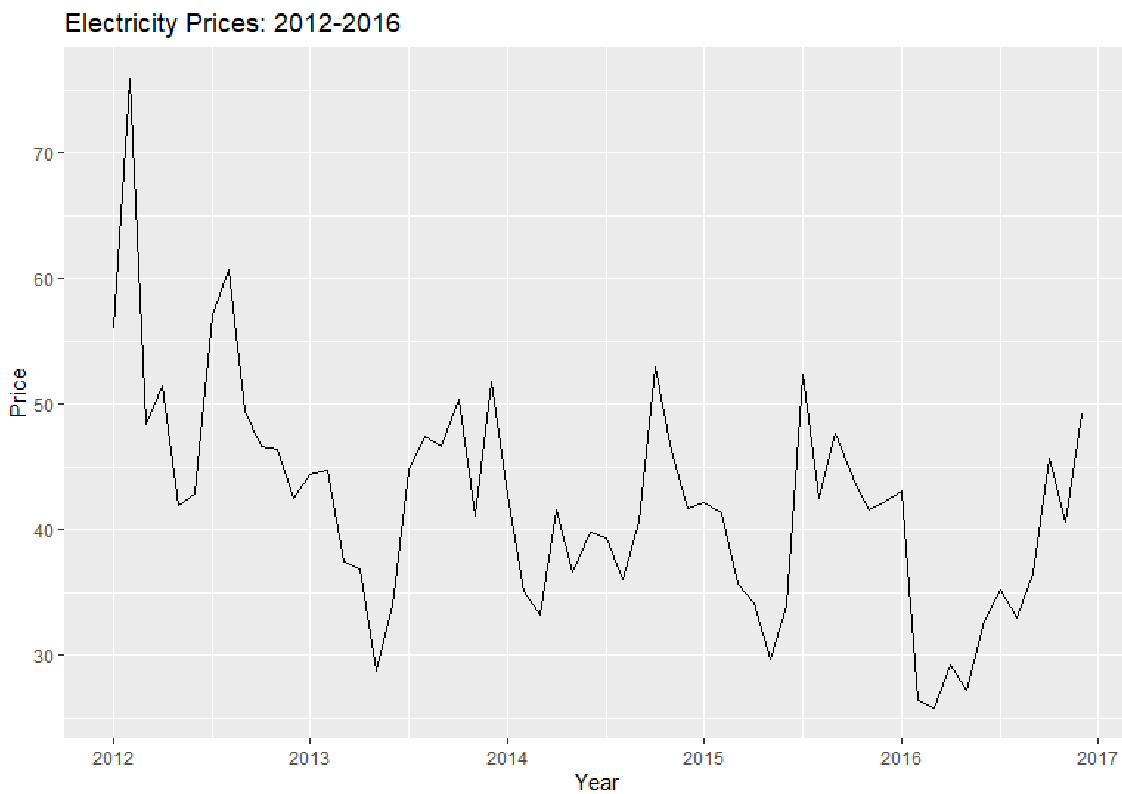

Figure 4: Time series visualization of monthly prices from 2012 till 2016

3.2. Time series exploration

From Figure 4, we see that there is a difference in prices based on the month of each year. Specifically, we observe that there might be some seasonal[1] effect in our time series for the monthly prices during the period of analysis (i.e. downward trend for the first 6 months and upward trend for the second half of the year). We also observed that during February 2012 the electricity price was a lot higher compared to the other prices which affects the trend[2] of the first three months of 2012. Note: We saw later on our analysis that these three months affected our forecasts accuracy, which is why we excluded them when creating our predictions.

In order to model a time series, we first need to be sure that it is stationary both in mean and variance. During some years the variance of the time series seemed to be stable, but for the whole time period it may be non-stationary as it has clear upwards and downwards trend. In order to be sure that the mean and variance has been stabilized for the whole period and due to the fact that the monthly prices we had in our disposal were non-negative, we tried using logarithms and taking differences to see what happens. After using the log-transformation, we did not observe any significant change as the logarithm did not seem to have any effect on the variance of the time series and due to the fact that a potential use of logs can be damaging for the forecast precision if a stable variance is not achieved, we decided to continue without applying this transformation.

So, we will take the differences for the time series without applying the log transformation. The difference plot in Figure 5, shows the difference of the each month with the previous one (e.g., the difference between the second and the first month, then the difference between the third and the second month, etc.). From this plot, we observe that the average price has not changed over time and that the variance of the series appears roughly constant over time. As a result, we can claim that the time series of first differences appears to be stationary in mean and variance.

[1] The seasonal pattern occurs when a time series is affected by seasonal factors such as the time of the year or the day of the week.
[2] The trend is the long-term increase or decrease in the data.

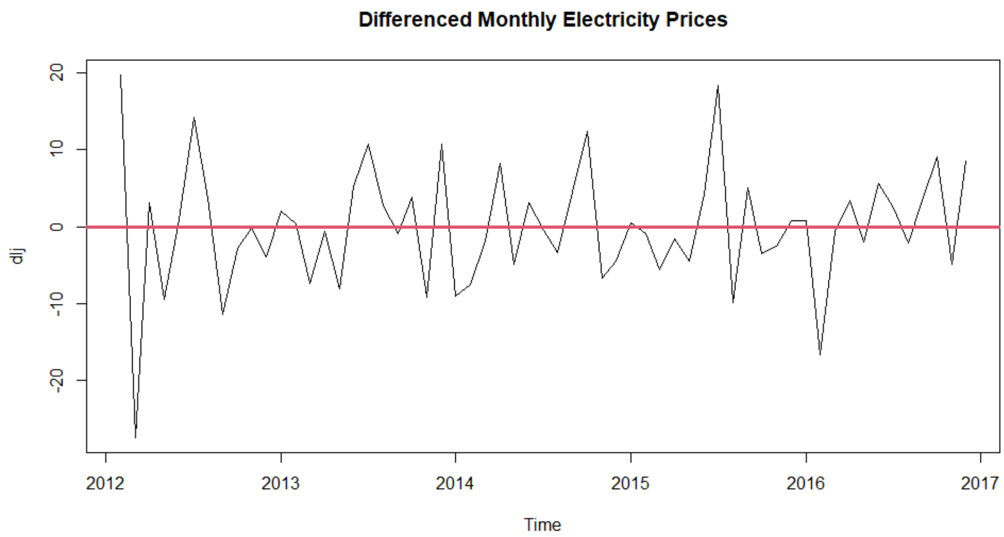

Figure 5: Difference plot for monthly prices from 2012 till 2016

After explaining why we decided to use the time series of first differences without applying the log transformation and presented the relevant figures, we have gathered the graphs that we based our reasoning and presented them in Figure 1 in APPENDIX for an easier comparison to be made.

In order to be sure about the stationarity of the time series, we also implemented the augmented dickey fuller test. From there, we saw that the null hypothesis was rejected (a=.05 > p-value = .01), thus our suspicions for the stationarity of the time series were confirmed. We also tested the normality for the time series of first differences, which appeared to be the case both from the figure 6 shown below and from the Shapiro-Wilk normality test where we did not reject the null hypothesis of normality (a=.05<p-value=.12).

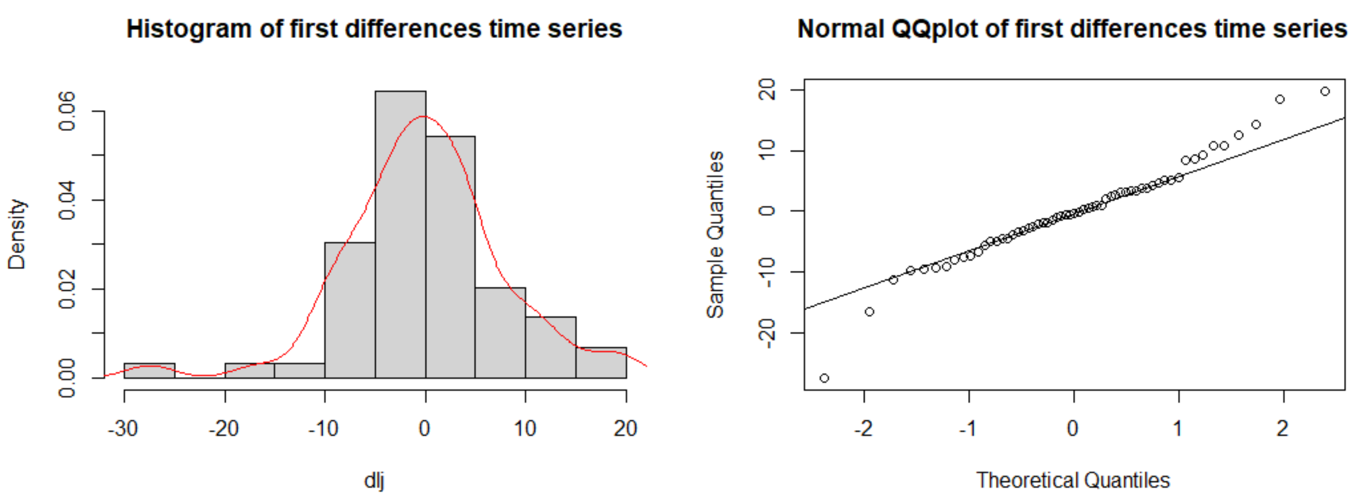

Figure 6: Histogram and normal QQplot for the time series using first differences

From Figure 7, we saw that during the first 6 months of a year there is usually a downward trend in prices, followed by an upward trend for the last 6 months. We also saw that as years pass by the prices are falling, leading to prices during 2016 to be the smallest over the whole time period for almost all year (except October and December). The unusual trend during the first three months of 2012 is clear in this plot also.

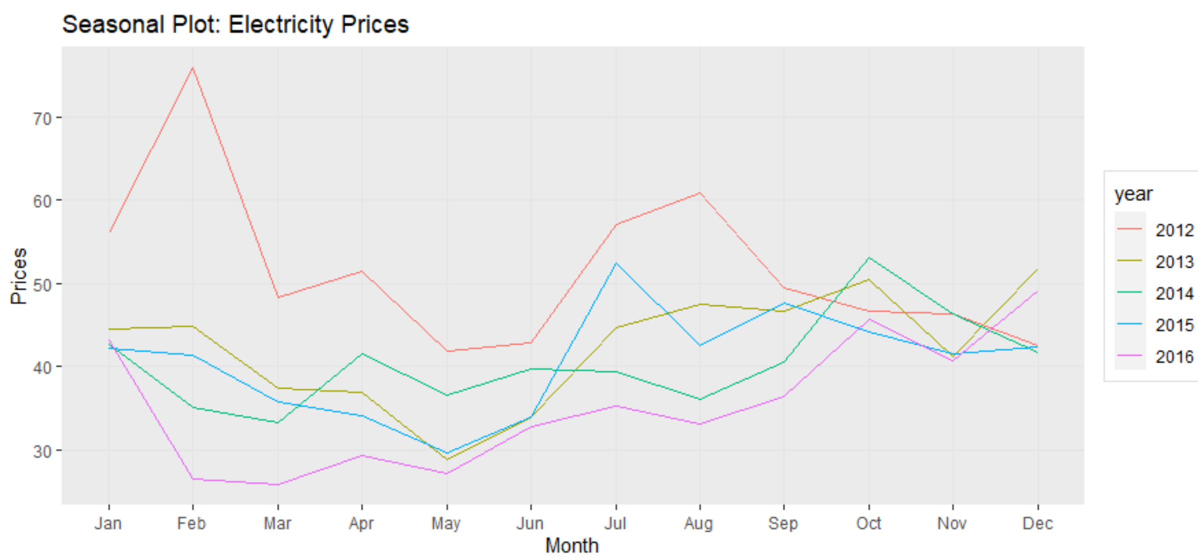

Figure 7: Seasonal plot for monthly electricity prices from 2012 till 2016

By looking at Figure 8, we were able to see the changes in seasonality over time and pinpoint changes in prices within particular time periods. Specifically, we saw a lot clearer compared to the previous graph that during the first 6 months there are generally lower prices compared to the second half of the year, especially from March till June which was a four month period with the lowest electricity prices. The peak period where the price is higher was during October (see blue horizontal line) and the period with the lowest price was during May.

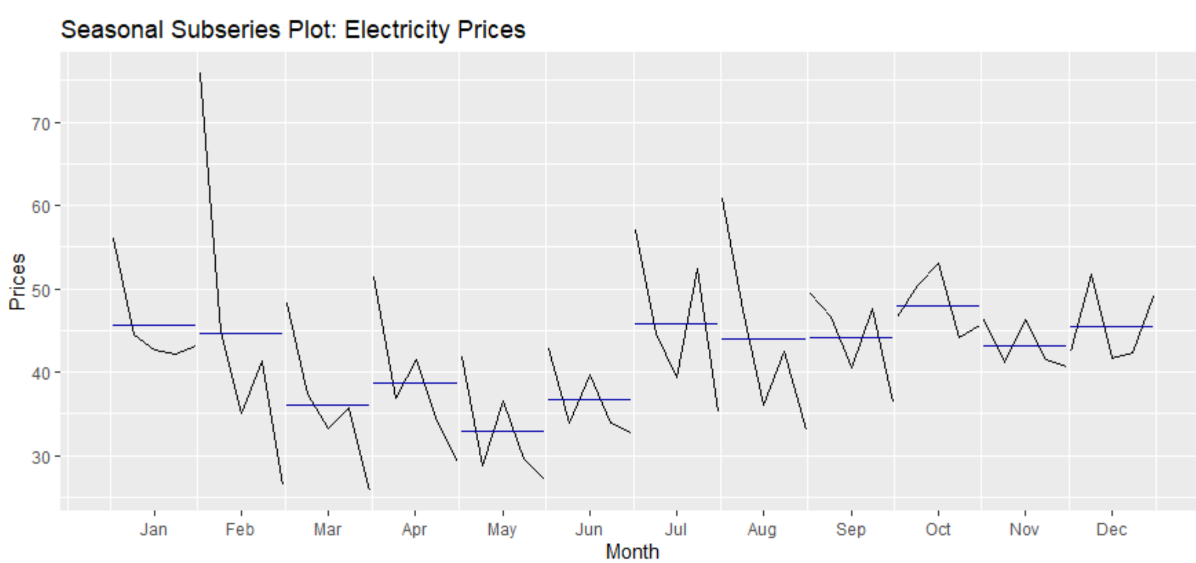

Figure 8: Seasonal subseries plot for monthly electricity prices from 2012 till 2016

We also created a plot of the autocorrelation and partial autocorrelation of the time series by lag as shown in Figure 9, where we saw that there was an autocorrelation problem with our time series as is, which needed to be reduce. For example, if an ARIMA model was to be used for this purpose, we have indications from these plots that we should start by using one AR and one MA terms (due to economy theory of the model, we prefer starting with the first lags that show autocorrelation out of bounds).

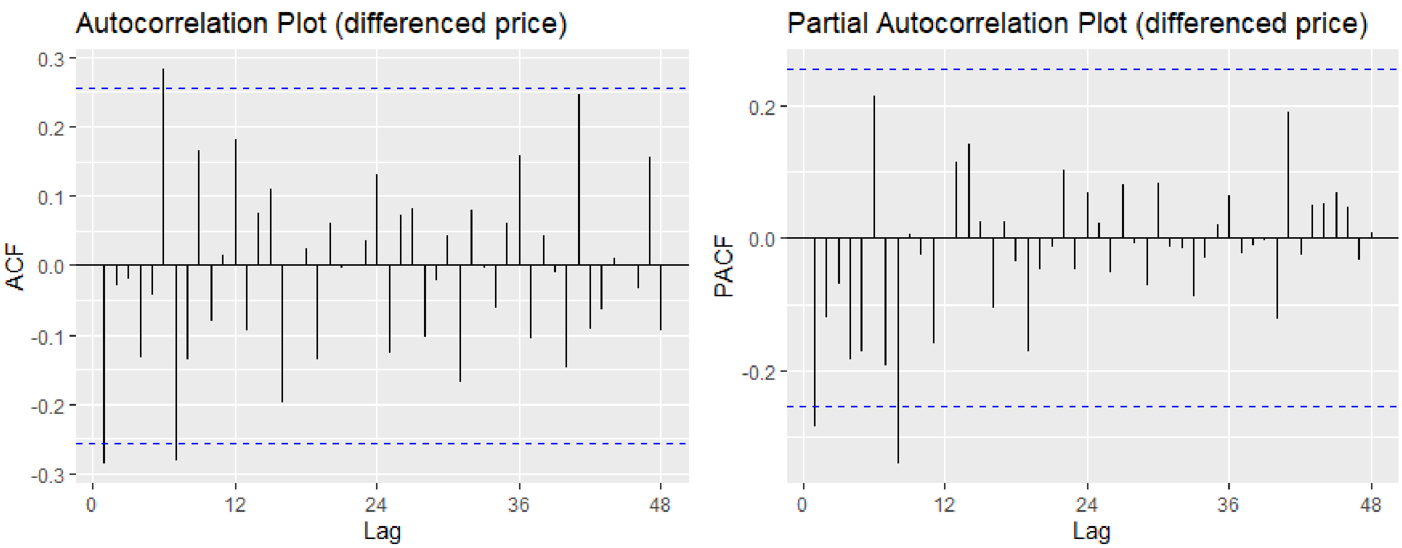

Figure 9: Autocorrelation and partial autocorrelation plots for the time series using first differences

In order to be more accurate, we also used the Ljung-Box test for the next 48 months, which tests whether any of a group of autocorrelations of a time series are different from zero. This test suggested that the data were significantly different from white noise (reject null hypothesis as a=.05>p-value=.003), thus this test confirmed the randomness of our time series data set. So, we saw that there is redundant autocorrelation in our time series which we have to model.

4. Time Series Modeling

Now, we will continue our analysis by modelling the time series in order to identify the best model for predicting the electricity prices in the next 24 months/2years (i.e. for the test dataset). The models' forecasting accuracy will be evaluated based on the RMSE metric (the smaller the value, the better). A good model would have similar RMSE score in the test and train sets, while a much bigger value in the test set is sign of overfitting while a much bigger value in the train set is sign of underfitting the data. For most problems, the value of this metric is expected to be lower for the training test. The forecasting methods we will use in this analysis refer to Autoregressive Integrated Moving Average Models (ARIMA) with and without regressors as well as to (Holt-Winters) Exponential Smoothing.

4.1. ARIMA Models

ARIMA models take into account the correlations in the data due to the fact that they include an explicit statistical model for the irregular component of a time series which allows the existence of non-zero autocorrelations in the irregular component.

These models are only defined for stationary time series, which is why we are expected to use the time series of first differences that we have previously saw transformed our time series into a stationary one. So, an ARIMA (p,1,q) model is probably appropriate for modelling the particular time series., where p indicates the order of the autoregression and q indicates the number of previous values we use for the moving average. We will now continue by identifying the best ARIMA model for this time series. This would be done by using the "auto.arima" function, but it could also be done by hand as well (i.e. using ACF/PACF plots to determine candidate models and AIC to identify the best one).

4.1.1. "Auto.arima" function

With "auto.arima" function, which uses a combination of unit root tests, for the minimization of the AIC and MLE to obtain an ARIMA model there is no need to use any stationarity transformations in advance (e.g. first difference time series), so we fitted the whole training set as is. Specifically, auto ARIMA takes into account the AIC and BIC values generated, are estimators to compare models, in order to determine the best combination of the model's parameters (p, d, q). The lower these values, the better is the model. The values of the parameters that this function had checked were up to five for each component (e.g. maximum value of AR and MA terms equal to 5).

The model recommended by the "auto.arima" function, is the one with the appropriate parameters that minimize the AICc. In our case, the function suggested an SARIMA model (ARIMA (1,1,1)(1,0,0)[12]) whose non-seasonal part contained one past value in our regression and one past error (i.e. one AR and one MA term) along with the past value for differencing (first differences) which was also seen from Figure 5, and the seasonal part contained only one past value (i.e. one AR term) of the same season.

Regarding the seasonality[3] of this model, the span of the periodic seasonal behavior was equal to 12 (frequency = 12), which was expected due to the fact that we used monthly data. The AICc of this model was equal to 404.95. We also noticed that the intercept was not included in the suggested model. Since the model suggests the use of a first difference time series, we expect that the average price will be very close to zero, so there is no reason to estimate the intercept. The mathematical form of the model is presented below:

ARIMA (1,1,1)(1,0,0)[12] model (where Y : Price)

$$\hat{Y}_t - Y_{t-1} = .45*(Y_{t-1} - Y_{t-2}) + .36*Y_{t-12} - .89*e_{t-1} \Rightarrow$$
$$\hat{Y}_t = Y_{t-1} + .45*(Y_{t-1} - Y_{t-2}) + .36*Y_{t-12} - .89*e_{t-1}$$

It is of pivotal importance to understand that the model proposed by “auto.arima” function is the one with the lowest AIC value, so the best model is considered to be the one with the best fit on the data, not the one with the best predictive ability.

4.1.2. Time series decomposition

To get an even clearer look at the trend and seasonal movements we mentioned before, we will go through a time series decomposition process, from which we will be able to see the actual data, the seasonal component (if it is a seasonal time series), the trend component and the remainder part. The sum of this decomposition together allows us to return to the actual data. From Figure 10, we observed that for the first difference time series the data seem to have slightly constant trend, seasonal fluctuations and random residuals.

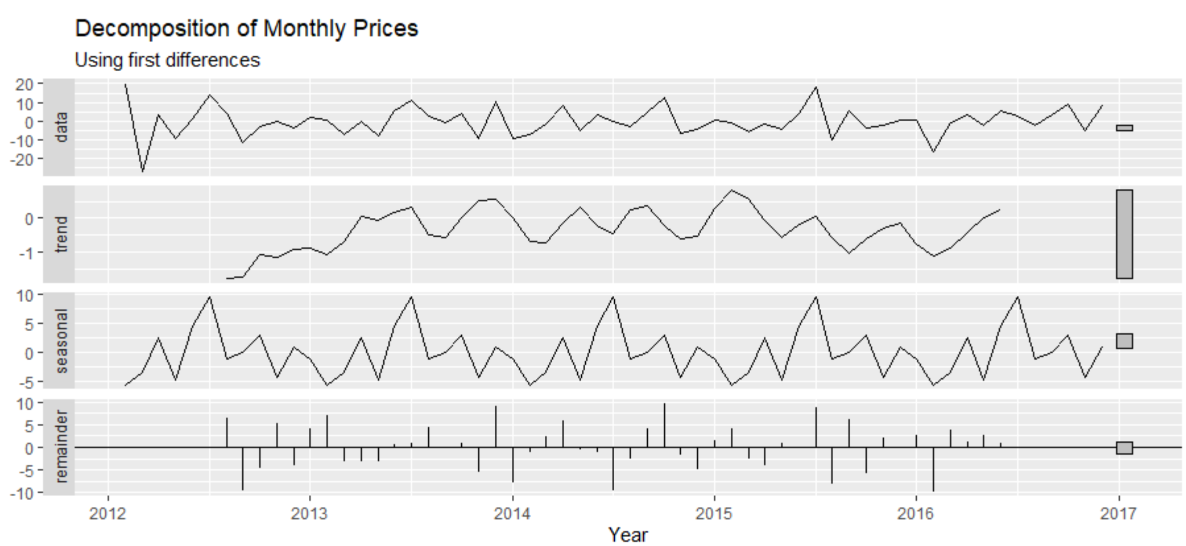

Figure 10: Time series decomposition (using first differences) of monthly prices

---

[3] Seasonality in a time series is a regular pattern of changes that repeats over S time periods, where S defines the number of time periods until the pattern repeats again.

4.1.3. Diagnostic checking

After estimating an identified model, we will now make a diagnostic checking for the residuals of the above model. We want the residuals to resemble a White Noise process (i.e. uncorrelated, homoscedastic and normally distributed). This is because a good forecasting method should yield residuals that are uncorrelated[4] and have zero mean[5]. Furthermore, it is useful (but not necessary) for the residuals to also have constant variance and be normally distributed. The last two properties also help in the calculation of prediction intervals.

4.1.3.1. Test autocorrelation of residuals

Regarding the autocorrelation of the residuals, we observed from Figure 11, that there was no significant correlation in the residuals time series (i.e. the residuals are behaving like white noise) as more than 95% of the autocorrelations were within the threshold limits, which is good because the lack of autocorrelation suggests that the forecasts of this model will be good.

In addition to looking at the ACF and PACF plots and in order to be more accurate, we also tested the autocorrelation of the residuals using the Lyung-Box test for autocorrelation, which also suggested that the residuals have no remaining autocorrelations, thus they are white noise (did not reject null hypothesis as a=.05<p-value=.27). So, we saw that the mean of the residuals is close to zero and there is no significant correlation in the residuals time series.

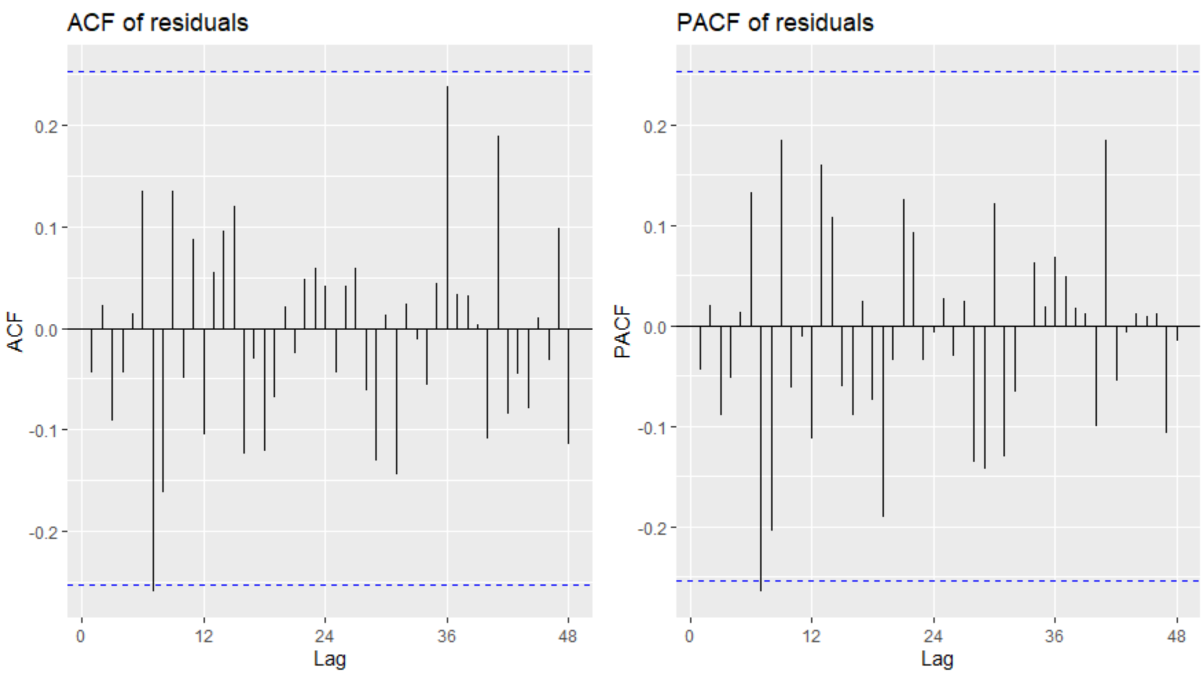

Figure 11: Autocorrelation and partial autocorrelation plots for the residuals of the model suggested by 'auto-arima' method

---

[4] If there are correlations between residuals, then there is information left in the residuals which should be used in computing forecasts

[5] If the residuals have a mean other than zero, then the forecasts are biased.

4.1.3.2. Test Heteroscedasticity of residuals

As for testing the heteroscedasticity of residuals saw from Figure 12 that there is no evidence of changing variance (i.e. all the spikes are inside the bounds) and therefore the residual variance could be treated as constant, which justifies our earlier reasoning that there is no need for using a Box-Cox transformation. This could also be seen by plotting the residuals of the time series shown in Figure 13 (where all the residuals seem to be inside a 'virtual' zone between [-7,7]) or from the histogram of the residuals shown in Figure 14.

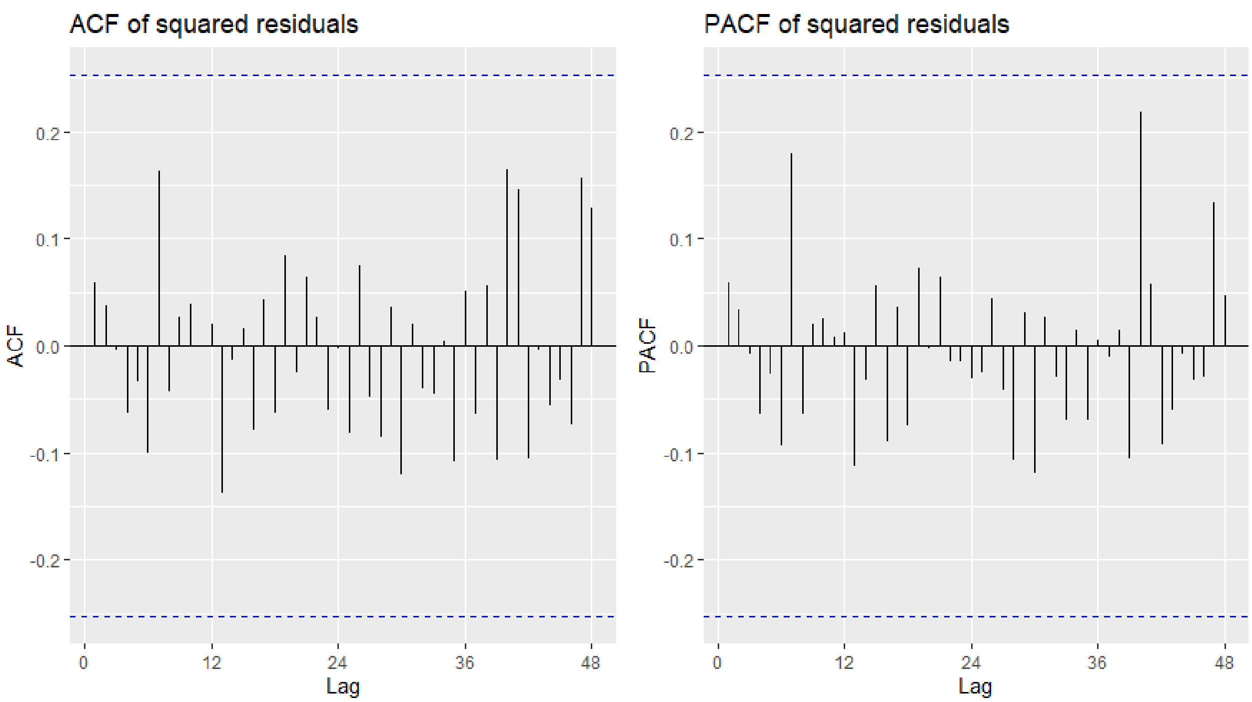

Figure 12: Autocorrelation and partial autocorrelation plots for the squared residuals of the model suggested by 'auto-arima' method

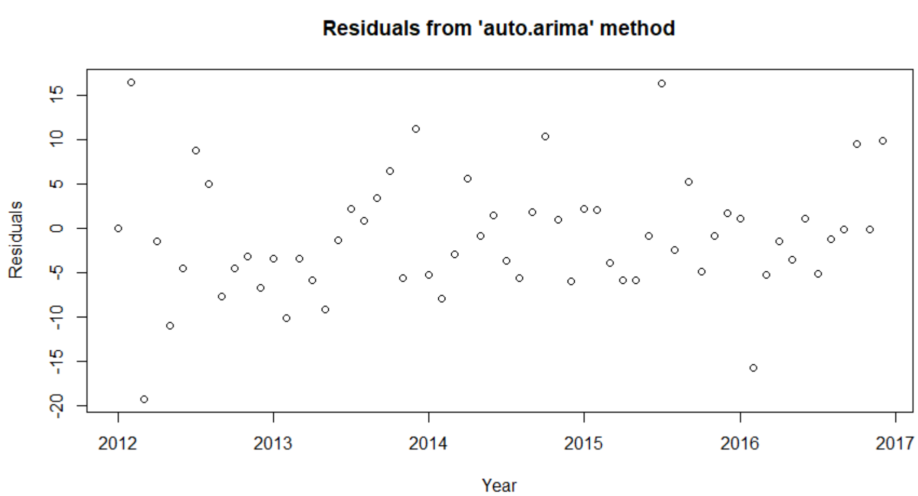

Figure 13: Residual plot of the model suggested by 'auto-arima' method

4.1.3.3. Test Normality of residuals

For testing the normality of the residuals, we created a qqplot and a Histogram as shown in Figure 14, from which it was suggested that the residuals seem to be normal.

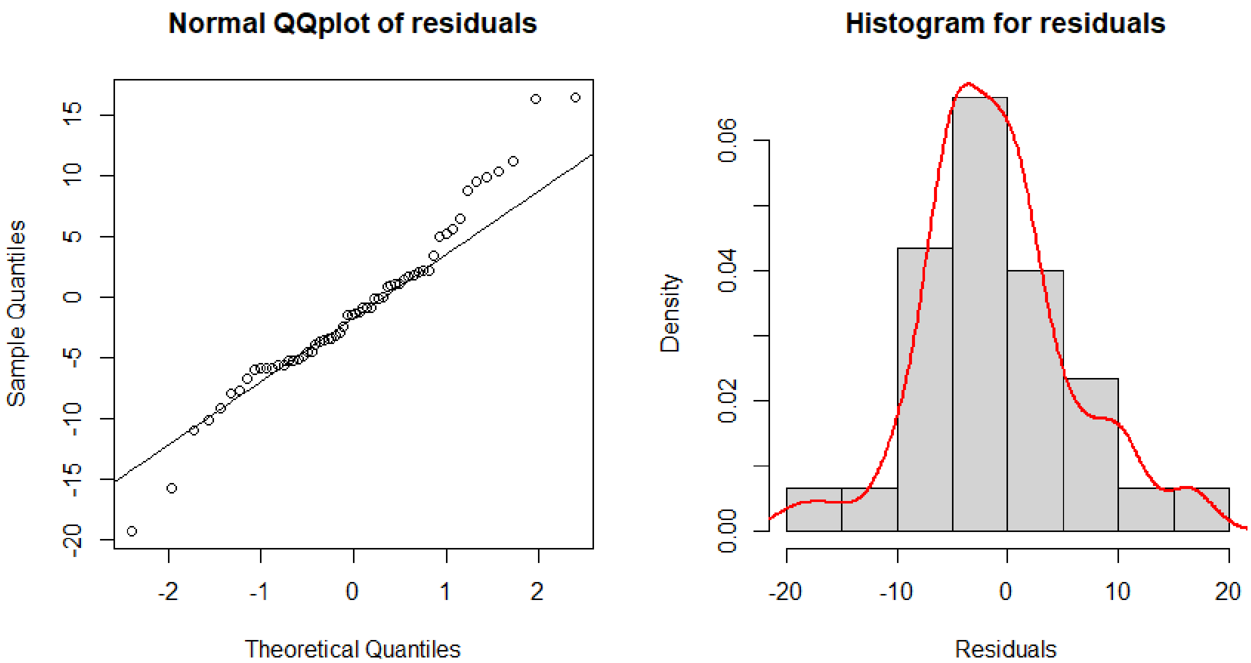

Figure 14: Histogram and normal QQplot for the residuals of the model suggested by 'auto-arima' method

So, we understood from the graphs that we have at our disposal a seasonal ARIMA model that passes all of the residual tests, thus the residuals of this model look like white noise. As a result, the forecasts produced using the model suggested by the "auto arima" method will probably be quite good and the prediction intervals that are computed would be accurate, as a normal distribution is assumed.

Note: We also saw that we did not have to account for heteroscedasticity problem of the residuals, that's why we decided that there was no need to try other models such as ARCH-GARCH that tend to fix this problem by modelling the variance.

4.1.4. Forecasting

We will create forecasts for the 2017 and 2018 years and compare them with the actual values in the test dataset. In "auto.arima" function the fitted values came from the training set, so the predictions will be in the scale of the actual electricity prices. In Figure 15 below, we present a plot of these forecasts with 1 standard error along with the prediction intervals. The RMSE value of the predictions in the test set was equal to 16.7.

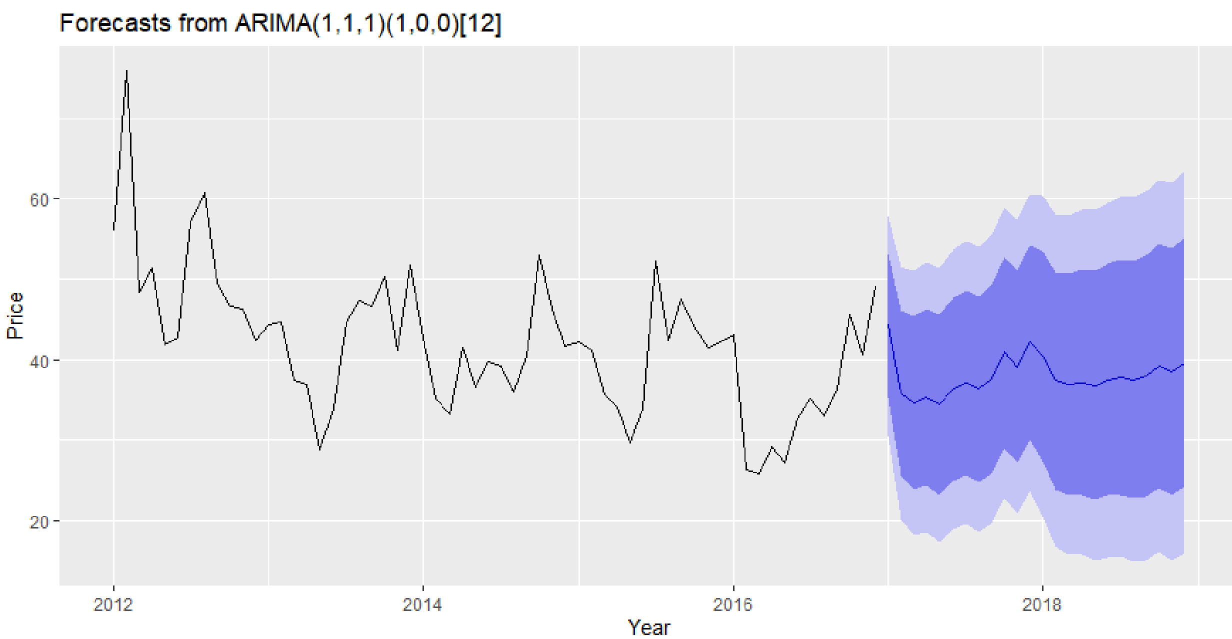

Figure 15: Time series plot for the forecasts (along with prediction intervals) created using the model suggested by 'auto-arima' method

4.2. ARIMA model with regressors

4.2.1. Interpretation and diagnostic testing

The ARIMA model can be extended to incorporate information provided by other exogenous variable, which can be done by adding one or more regressors to the ARIMA forecasting equation (dynamic regression models). In our analysis, we decided to use an indicator regarding the total primary energy consumption which was excluded from the monthly energy review of the U.S. Energy Information Administration and was measured in Quadrillion Btu. *Note: The choice of the regressor used was done for academic and completeness purposes and its addition in the ARIMA model may not be helpful.*

First of all, we saw that there were not any duplicate rows in the data or any infinite or missing values in the new data we added. After converting the monthly data into the same unified date format as we did with the electricity price dataset, we kept the data form 2012-01 till 2019-06. We observed that we also kept the values for the first half of 2019, which was done because we want to have this values in the model when predicting for the specific time period. The total primary energy consumption had a mean price of 8.1 quadrillion Btu and a negligible correlation with price.

After splitting the new dataset with the added regressors using the same proportions as before into training and test sets, we saw that the terms in the ARIMA model suggested by the "auto.arima" function were exactly the same as before, but their coefficients values are different and of course there is the part of the regressor added. The AICc value of the model was equal to 403.9.The mathematical form of the model is presented below:

<u>ARIMA (1,1,1)(1,0,0)[12] with regressor:</u>

$$\hat{Y}_t - Y_{t-1} = .41*(Y_{t-1} - Y_{t-2}) + .43*Y_{t-12} - .86*e_{t-1} + 4.8*(X_t - X_{t-1} - .41*(X_{t-1} - X_{t-2})) \Rightarrow$$

$$\hat{Y}_t = Y_{t-1} + .41*(Y_{t-1} - Y_{t-2}) + .43*Y_{t-12} - .86*e_{t-1} + 4.8*(X_t - 1.41*X_{t-1} - .41*X_{t-2}$$

Where:

- Y : Price
- X: Total primary energy consumption

So, we observed that the AR part of the model and the differencing transformation is applied to the regressor variable (X) in exactly the same way as it is applied to the price variable (Y) before the regressor (X) is multiplied by the regression coefficient. In other words, the ARIMA (1,1,1)(1,0,0)[12] model is fitted to the errors of the regression of price (Y) on regressor (X).

Similarly, as in the model without the regressor, this seasonal ARIMA model passes all of the residual tests, thus the residuals of this model look like white noise and the model can is appropriate for forecasting. The validity of the model, can be seen in Figure 16. We also used a Lyung-Box test for autocorrelation, which just like the ACF and PACF plots suggested that the residuals are behaving like white noise. Note that due to the fact that we now have regressors in our model, we can also test the heteroscedasticity of residuals using a residuals plot against the fitted values as is shown in Figure 17, where their homoscedasticity is also clear.

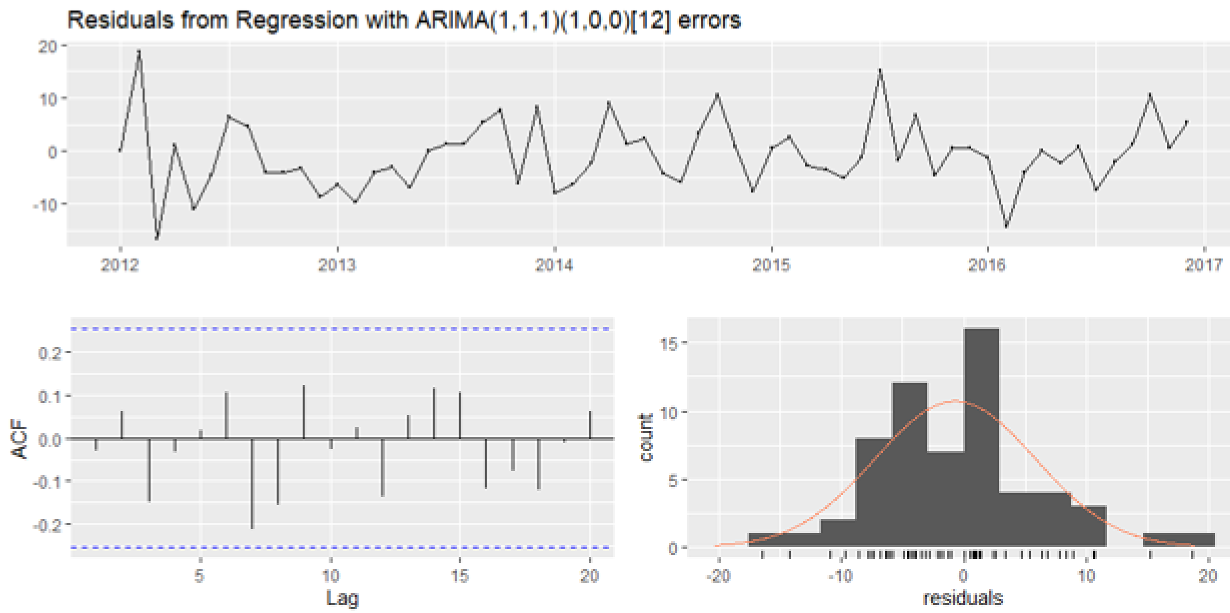

Figure 16: Basic residual plots for testing the properties regarding the residuals of the suggested by 'auto-arima' method along with other regressor

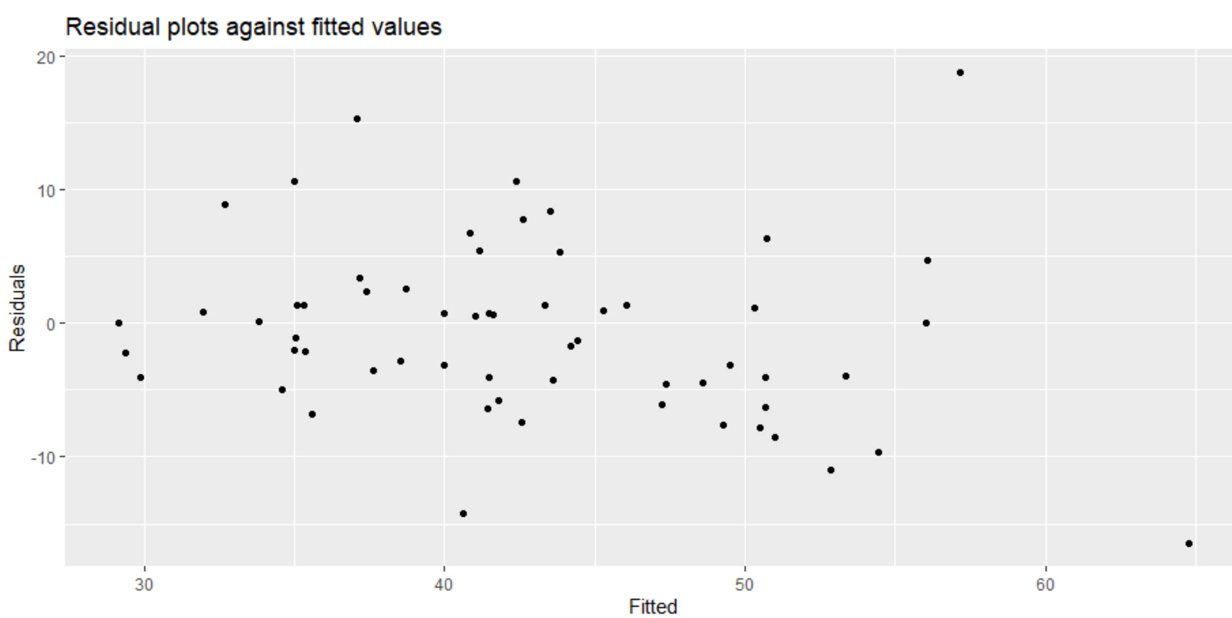

Figure 17: Residual plots against fitted values for the model suggested by 'auto-arima' method along with other regressor

4.2.2. Forecasting

In order to generate forecasts for the next 24 months (i.e. 2017, 2018 years) we also need to supply the future values of the total primary energy consumption indicator for this time period (i.e. from 2017-01 until 2018-12). In Figure 18 below, a plot of these forecasts with 1 standard error along with the prediction intervals is presented. The RMSE value of the predictions in the test set was equal to 17.1.

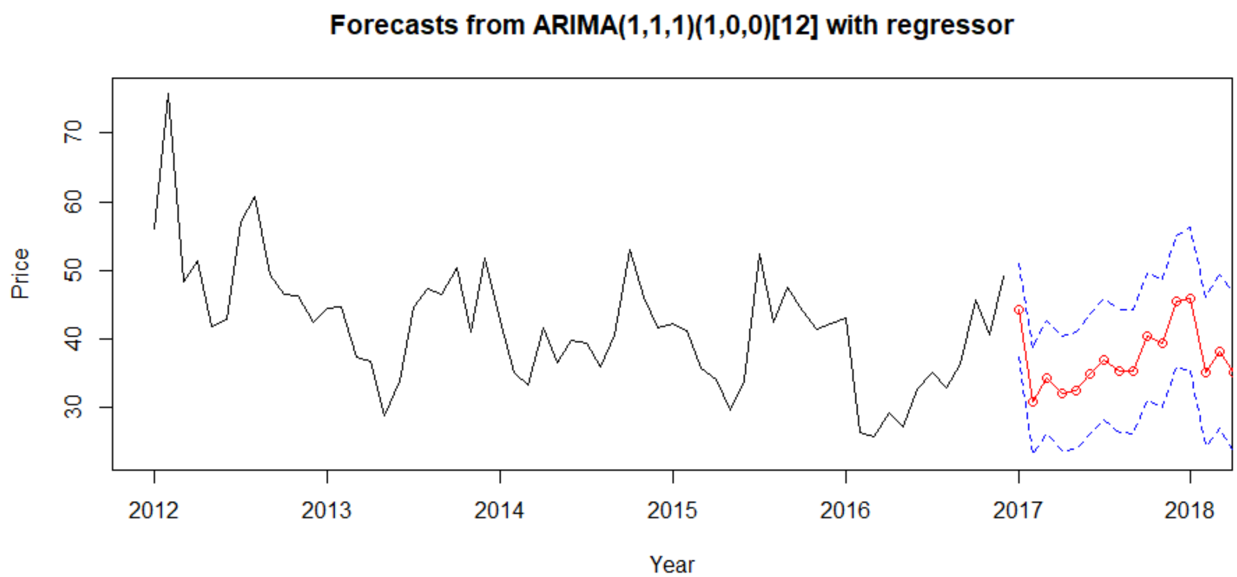

Figure 18: Time series plot for the forecasts (along with prediction intervals) created using the model suggested by 'auto-arima' method along with other regressor

4.3. Holt-Winters Exponential Smoothing

Exponential smoothing is another forecasting method which in contrast with the ARIMA models, does not make any assumptions about correlations between successive values of the time series. However, the construction of prediction intervals requires that the residuals are uncorrelated and are normally distributed with mean zero and constant variance.

Our data contained seasonality, so we cannot use smoothing methods such as simple exponential smoothing (SES) as they only work if there is no seasonality on the data. Instead of choosing a seasonal adjustment and implementing these methods, we decided to use the Holt-Winters exponential smoothing model which deals with both trends and seasonality. The selection of the Holt-Winters exponential smoothing for making short-term forecasts was also based on time series decomposition, from which we saw that our time series could be described using an additive model with (changing) trend and seasonality.

Holt-Winters method estimates the level, slope and seasonal component at the current time point, where the smoothing is controlled by the parameters alpha, beta, and gamma, for the estimates of the level, slope b of the trend component and the seasonal component, respectively, at the current time point. The parameters alpha, beta and gamma take values in the range of 0 and 1 (values that are close to 0 mean that relatively little weight is placed on the most recent observations when making forecasts of future values).

To deal with the trend, this method allows the smoothing parameters alpha and beta to minimize the sum of squared errors and in order to deal with seasonal changes it changes the average of gamma parameter to either average of zero (additive) and average of one (multiplicative). In our case, we chose to set it equal to zero (additive).

After implementing this method to the training data, we saw that the estimated values of alpha, beta and gamma smoothing parameters were 0.26, 0.04, and 0.56, respectively. The value of alpha (0. 26) is relatively low, indicating that the estimate of the level at the current time point is based upon both recent observations and some old observations. The value of beta is 0.04, is so negligible that essentially it indicates that the estimate of the slope b of the trend component is not updated over the time series (approximately it remains the same). As for the value of gamma (0.56) is medium, indicating that the estimate of the seasonal component at the current time point relies significantly on the very recent observations, but also takes into account some older observations.

In Figure 19 below, a plot of these forecasts with 1 standard error for the 2017, 2018 years along with the prediction intervals is presented. The prediction intervals were able to be constructed because we observed from Figures 20 & 21, that the residuals were white noise. The RMSE value of the predictions in the test set was equal to 18.99.

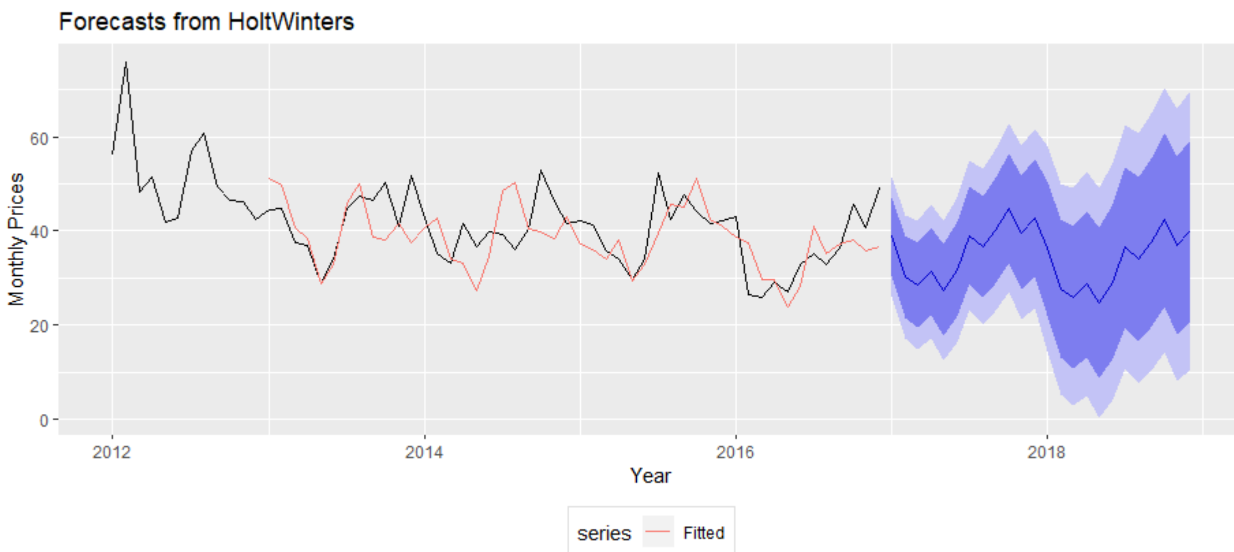

Figure 19: Time series plot for the forecasts (along with prediction intervals) created using the model suggested by Holt-Winters method

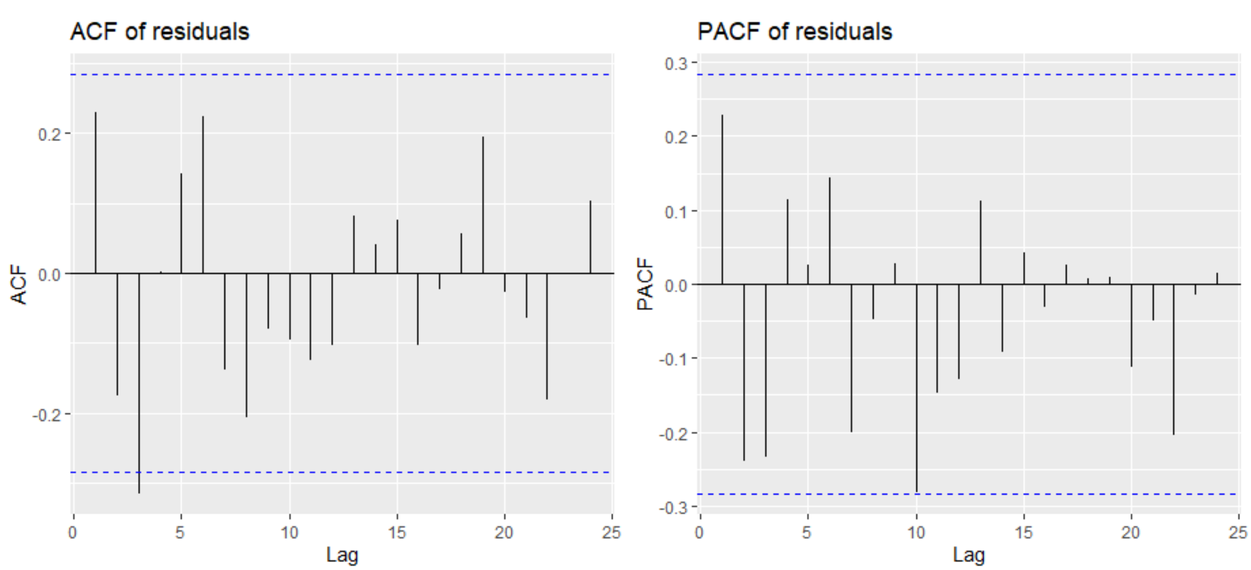

Figure 20: Autocorrelation and partial autocorrelation plots for the residuals of the model suggested by Holt-Winters method (used for testing the autocorrelation of residuals)

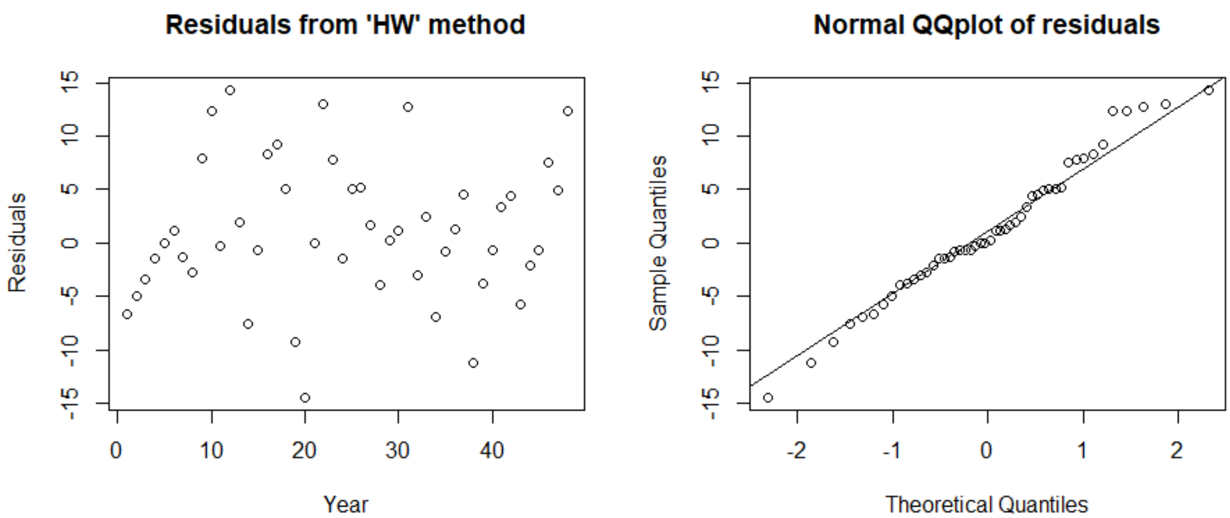

Figure 21: Residual plot and normal QQplot for the residuals of the model suggested by Holt-Winters method (used for testing the heteroscedasticity and normality of residuals)

5. Evaluating forecast accuracy

In order to evaluate the accuracy of the forecasts created and select the "best" model for predicting the prices for the first half of 2019, among the ones provided by "auto.arima" with and without regressor and from Holt-winters method, we compared their results using two different methods. The first one was based on training/test sets and the other on time series cross-validation. The comparisons between the results of the different methods implemented will be done using the Root Mean Squared Error (RMSE) which is a scale-dependent measures that, as the title suggests, is based on the squared errors and the best forecasting model will be the one with the smallest RMSE. Also, a forecast method that minimizes the RMSE will lead to forecasts of the mean, which was the target of our analysis as we wanted to predict the mean monthly prices, while other measures such as the MSE their minimization leads to forecasts of the median. The RMSE is computed as shown below:

$$\text{Root mean squared error: RMSE} = \sqrt{\text{mean}(e_t^2)}.$$

5.1. Training and test sets

As have already mentioned, we have split about 70% of the data into **training** set and the rest 30% as **test** set. We have used the training data to estimate the parameters of the implemented forecasting methods and the test data to evaluate their accuracy in terms of predicted ability. That way, the test data would provide an indication as to how well each model is likely to forecast on new (unseen) data.

In Table 1 below we have summarized the results of the different models used, which were fitted using data from January 2012 to December 2016 and the forecasts were created for the time period from January 2017 till December 2018. *Note: Later on our analysis we will exclude the first three months of 2012 from our data because we saw that they seriously affected the forecast accuracy.*

Table 1: RMSE values in the test set for various forecasting methods

Method	RMSE
Auto-arima	16.7
Auto-arima with regressors	17.1
Holt-Winters Exponential Smoothing	18.9

It seems like the model chosen with "auto.arima" function is considered the best model as it has the lowest RMSE value on the test set. All of these models have passed the residual tests, but due to the fact that our goal is forecasting we would select the best model (i.e. the one with the lower RMSE) even if its residuals did not behave like white noise.

5.2. Time series cross-validation

In time series cross-validation, there are a series of test sets each consisting of a single observation and the corresponding training set only consists observations that occurred prior to the observation that forms the test set. As a result, future observations cannot be used for constructing the forecast. Of course, a reliable forecast cannot be obtained based on a small training set which is why the earliest observations are not considered as test sets. The forecast accuracy is computed by averaging over the test sets.

In our analysis, we want to predict the prices for the first six months of 2019 (i.e. 6-step-ahead forecasts), which is why we used a modification of the cross-validation procedure in order to use multi-step errors.

For each one of the forecasting methods we have created a matrix containing the forecast errors[6], between the forecasted ahead value and the actual value, for a forecast horizon h equal to six (as the values that we want to predict). As mentioned, we have used the prices until December of 2016 as the training set, which corresponded to the 60th observation (last period of training data). So, the value in row 60 (last of training set) of the newly created matrix will contain in the first column the 1-step error made at time 60, for time period 61, in the second column the 2-step error made at time 60, for time period 62, etc.. It is clear, that the last h cells of the matrix (as the forecasted h-step ahead value will be NAs as there will not be any more future actual values available for the comparison to be made.

Considering that in the matrix we have calculated the forecast errors which are on the same scale as the data, the predicted value of price will be computed by subtracting the forecast error from the actual value. For example, a one-step forecast such as the one for the prediction made regarding January 2017, the actual value of January 2017 minus the one-step forecast error made at December 2016 for January 2017.

After comparing the results of these two forecasting evaluation methods, it was clear that when using time series cross-validation the RMSE from the residuals was smaller compared to the ones using the training/test split,. This which was expected due to the fact that the corresponding “forecasts” were based on a model fitted to the entire data set, rather than being true forecasts.

[6] A forecast error is the difference between an observed value and its forecast.

6. Best combination for forecasting

So, we will evaluate the implemented forecasting models using time series cross validation. While constructing our predictions, we decided to exclude the first three months of 2012 from the data. This was because the price during February 2012 far exceeded the rest, leading to a significant increase in January and a significant decrease in March, which affected the accuracy of the forecasts. This trend can be seen in Figure 3 we have already created. Also, through trial and error we observed that the best results for the Holt-Winters method were achieved using the data from January 2014, which are the ones we will use for this method from now on.

An example that highlights the difference between these two approaches regarding data usage, can be seen in Table 2 below. From this table, it is clear that the predictions became better (smaller RMSE) without included the data from the first three months of 2012.

Table 2: RMSE values compared for different time periods and various forecasting methods using the mean for aggregating daily prices to the month level

	RMSE for one-step forecast (h=1)	
Model	With first three months of 2012	Without first three months of 2012
Auto-arima	12.45	9.59
Auto-arima with regressors	12.37	9.36
Holt-Winters Exponential smoothing	12.76	10.02

The predictions so far, have been created by using the mean price as the representative price for each month when converting the data from daily to monthly. We are not sure if this is the "ideal" representative whose replacement leads to the best predictions. So, we decided to evaluate our models by using the mean, max, min and median price of each month along with the median price of the lower half of the data set (1st quantile) and the central point that lies between the median price and the highest number of the distribution (3rd quantile) of each month and see which one produced the best results in terms of predictive ability.

In Table 3, we presented the RMSE values for all the different representative prices of each months, for all the different methods and for the whole forecast horizon which refers to the first six months of 2019.

Table 3: RMSE values compared using different forecasting horizon for various forecasting methods and for different aggregation methods used for converting daily prices to monthly

		RMSE (forecast horizon h)					
Aggregation of each months' price	**Method**	**h=1**	**h=2**	**h=3**	**h=4**	**h=5**	**h=6**
Mean price	Auto-arima	9.59	12.29	12.78	12.95	13.2	13.22
	Auto-arima with regressors	9.36	12.08	12.89	13.18	13.30	13.10
	Holt-Winters Exponential Smoothing	10.02	12.37	12.31	13.07	14.02	14.51
Max price	Auto-arima	21.63	23.12	28.50	26.20	26.51	25.51
	Auto-arima with regressors	20.44	25.03	25.09	27.13	26.34	26.36
	Holt-Winters Exponential Smoothing	25.37	29.66	30.26	32.48	34.27	33.38
Min price	Auto-arima	11.23	11.35	12.06	12.32	11.79	11.93
	Auto-arima with regressors	11.42	11.99	12.24	12.24	12.08	11.88
	Holt-Winters Exponential Smoothing	11.49	13.35	14.28	14.86	15.44	16.53
Price in 1st quantile	Auto-arima	**7.22**	**9.36**	9.94	10.41	11.01	**11.06**
	Auto-arima with regressors	7.40	9.61	10.32	10.76	11.15	11.14
	Holt-Winters Exponential Smoothing	7.99	9.47	**9.72**	**10.18**	**10.86**	11.62
Price in 3rd quantile	Auto-arima	11.07	13.51	14.03	14.30	14.87	15.02
	Auto-arima with regressors	11.01	13.35	13.65	14.40	14.7	14.93
	Holt-Winters Exponential Smoothing	11.87	13.83	13.86	14.53	15.83	16.30
Median price	Auto-arima	9.53	12.16	12.63	12.67	12.69	12.93
	Auto-arima with regressors	9.69	12.13	12.64	12.64	12.56	12.52
	Holt-Winters Exponential Smoothing	9.58	11.33	11.45	12.26	13.24	14.00

So, we observed that the best (smaller RMSE values) were achieved using different methods for each month, but they all occurred when converted the daily data to monthly using the median price of the lower half of the data set (1st quantile). Furthermore, we observe that we have to use a combination of the implemented methods in order to achieve the smaller RMSE during each month. Specifically, for January, February and June of 2019 we will use the model suggested by the 'auto.arima' method and from March till May we will implement the Holt-Winters exponential smoothing method.

In brief, because we have already explained how the models work and give an interpretation of their values, we saw that the model suggested from the 'auto.arima' function for this data, was an autoregressive model with one AR term (with coefficient equal to .63) and non-zero mean (equal to 44.5). As for the Holt-Winters method that was applied to the data from January 2014 till December 2018, the estimated values of alpha, beta and gamma smoothing parameters were 0.42, 0, and 0.03, respectively. The value of alpha (0. 42) indicated that the estimate of the level at the current time point is based on both recent and old observations, the zero value of beta 0.04 indicates that the estimate of the slope b of the trend component remains the same and the value of gamma (0.03) shows that for the estimate of the seasonal component relatively little weight has been placed on the most recent observations when making forecasts.

7. Findings

As we have finished our analysis, we will now present the results of our research. In Table 4 and Figure 22, we have presented the mean monthly electricity prices for the first six months of 2019, along with the expected change (increase or decrease) compared to the previous month. Specifically, we observed that the mean price for 2019 will decrease during January (as the price during December 2012 was equal to 65$) and also during February. For March 2019, we expect to see an increase in the price, before beginning a downward trend which will last from April till June.

Table 4: Predictions of electricity prices for first half of 2019 (rounded to 3 decimal places) along with an indication of the price change from month to month

Month	Forecast	Price change from previous month
January	57.637	(-11.33%)
February	52.897	(-8.22%)
March	53.980	(+2.05%)
April	52.577	(-2.59%)
May	49.094	(-6.63%)
June	45.928	(-6.45%)

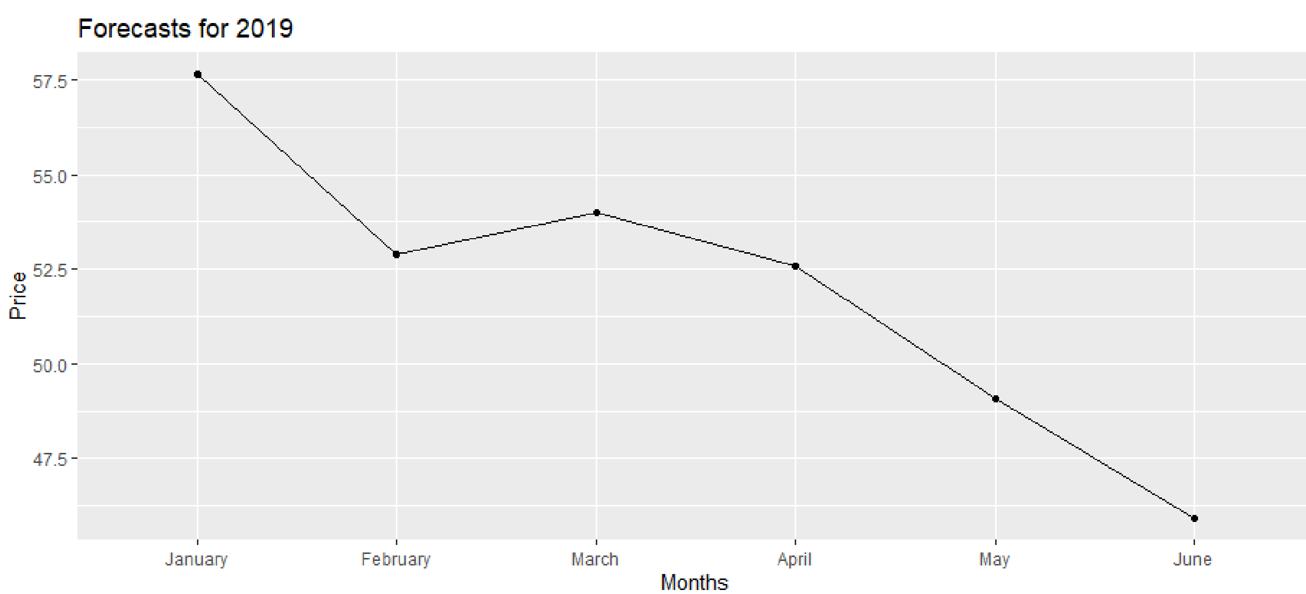

Figure 22: Time plot for the predicted prices of the first half of 2019 showing the price change from month to month

Bibliography

Hyndman R. J. (2014). *Forecasting: Principles & Practice*. University of Western Australia. 23-25 September 2014. Available at: https://robjhyndman.com/uwafiles/fpp-notes.pdf

Hyndman R. J. & Athanasopoulos G. (2018). *Forecasting: principles and practice*. 2nd Edition. OTexts. Available at: https://otexts.com/fpp2/

Long, J. D. & Teetor, P. (2019). *R Cookbook*. 2nd Edition. O'Reilly Media. Available at: https://rc2e.com/index.html

Nau R. (2020). *Statistical forecasting: notes on regression and time series analysis*. Available at: https://people.duke.edu/~rnau/411home.htm

APPENDIX

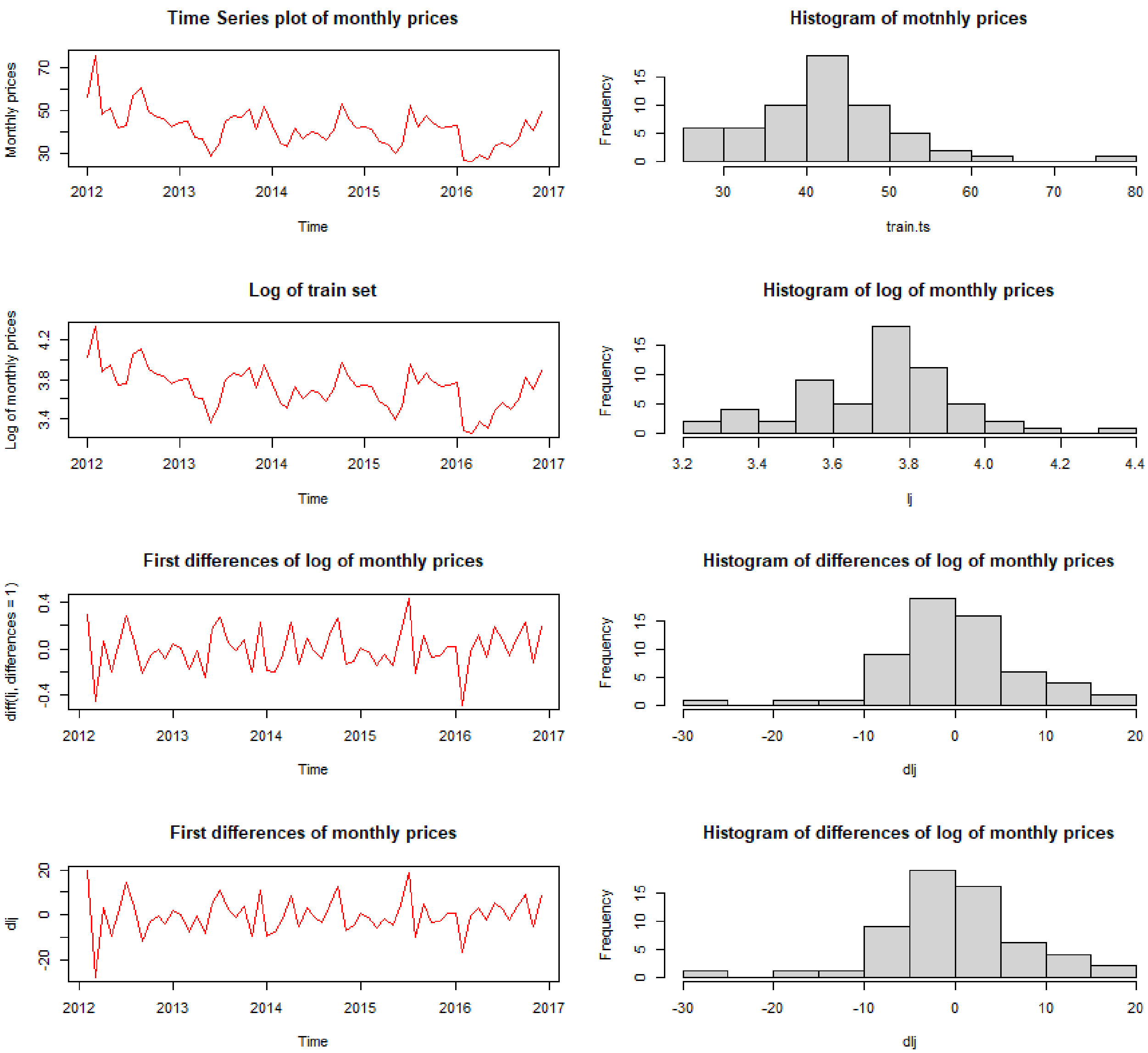

Figure 1: Series of plots used for deciding the appropriate transformations of the time series referring to monthly electricity prices