Europa-Universität Flensburg

“Fundamental Analysis of the German Prompt Power Market in Dec/Jan 2018/2019”

„Energy- and Environmental Management“

Lecture: Trading Energy
Lecturers: Jörn Higgen, Dario Kocevar, Larissa Leienbach and Johannes Viehmann

Presented by :

Daniel José Camacho Thielepape (550093)

Mohammed Abi Afthab (550031)

Manuel Fernando Valois (550039)

17 of February 2019

I. Table of contents:

II. Abstract

The following report presents the fundamental analysis of German day-ahead market and a procedure conducted for forecasting the hourly prices of the electricity in the German spot market for the period between the 21st and the 27th of January, as well as the results obtained. A comparison against the actual values is carried out; conclusions about the causes in the deviation in the forecast close the report.

Chapter 2 present an analysis of the German electricity market, the market participants and the electricity price evolvement. Chapter 3 gives a brief introduction to the different approaches that can be used for performing the forecast; based on the accuracy of the results and the complicity of their implementation, the Multiple Regression Analysis is selected as the most appropriated method for conducting the forecast task. Chapter 4 introduces the factors influencing the electricity price in the German power market; later in this chapter the variables selected as the exogenous variables needed for the implementation of the Multiple Regression Model are summarized. In Chapter 5 the collected historical values of the exogenous variables are analyzed; the impact of environmental conditions and remarkable situation like the occurrence of negative prices are studied in detailed. Chapter 6 introduces the concept of "Spike", its causes and the undesirable impact in the forecast results are explained; at the end of this chapter the thresholds used for limiting the impact of spikes in our model are summarized.

In chapter 7 the challenges faced for forecasting the demand, the wind and solar generation for the target week are introduced; similarly, the price forecasting model is tested by predicting the prices for the month of December, to later compare the predicted and the actual values; the results evidenced that the model is able to provide an adequate forecast. Coming to the end, in chapter 8 the results of the forecast for the target week are presented; the preliminary results shows that the average value of forecast which is 64.48 Euro/MWh is very closer to the actual average which is 63.54 Euro/MWh. However, there were few hours which showed with considerable deviation from the actual prices; a further analysis shows that expected wind generation for the target week used as exogenous variables in the mathematical model deviate considerably from the actual wind generation during the target week. To finish, the forecast was computed one more time, but this time using the actual wind generation, obtaining much accurate results.

III. List of figures

IV. List of tables

V. List of abbreviations

Abbreviation	*Meaning*
EPEX	European Power Exchange
GHG	Greenhouse Gases
OTC	Over the counter
SE	Societas Europaea
TSO	Transmission system operators
GFS	Global Forecast System
ECMWF	European Centre for Medium-Range Weather Forecast

1. Introduction

1.1 Purpose of the report

The objective of this report is part of the practical assignments of the lecture; “Trading Energy”. The subject matter is divided into two areas, first the fundamental theory of trading energy and the second part practical case study. This report aims to analyze the practical part, “Fundamental Analysis of the German Prompt Power Market in Dec/Jan 2018/2019”.

1.2 Process of the report

After attending the three days of the seminar, acquired the information, and source recommended by the lecturer, the next step is to proceed to analyze and conclude the practical part. The most relevant information has been extracted to set a general understanding of this topic, following the purpose of covering all the recommended content concluding with own comments.

2. Business Understanding

Before focusing on the analysis of the mentioned case study, it is recommended to interconnect all the terms in an overall picture, describing the context. This point of view gives a better understanding of interfering in a forecast of electricity prices. Down below, a description of the market, the participants and the price evolvement is given.

2.1 Electricity Markets

At the core, the electricity power sector is regulated in many different ways across Europe. Each country has its own regulatory model, and also political changes in national regimes make this sector unstable. For over a decade, Europe has been moving towards a liberalized system for power generation and energy sales. This has supported development of power trading markets and cross-border activity between European power and utility companies. Regulators are demanding more and more efficiency, and they are pushing for fair prices. The German regulatory model focuses on driving costs down by benchmarking similar operator against each other, to fully eliminate inefficient costs (EY, 2013).

In Europe, there are two prevailing electricity markets: exchange market and over the counter (OTC). In the exchange market, the trading parties stay anonymous but the transparency on prices and power plants availability for example, is the key challenge and framework to trade, moreover, the products traded are standardized. In the OTC, also known as bilateral, suppliers trade directly with customers, and the traded prices stay anonymous, therefore a less formal context allows more customized products. Although a major amount of electricity is traded in the OTC market, the exchange market plays a central role in determining the electricity prices. In both markets, long- and short-term contracts are possible.

Our study case exclusively is about the Day-Ahead prices at the German Power Exchange, which takes place at the EPEX SPOT market. EPEX SPOT SE is a company under European law, based in Paris, the power spot markets at the heart of Europe, where Day-Ahead and Intraday markets are its engine.

Not only power generators and power consumers trade at the market, but also the Transmission System Operators (TSO), who are responsible of the bulk transmission of electric power, making the trading energy physically possible and who provide grid access to the electricity market players. It plays an important role in the electricity prices, not only because they manage the distribution and transition capacity of the grid, but also have to offer customers a secure and stable power system. For this, power generators offer its contribution of primary, secondary and tertiary reserve as a valuable and traded product to ensure in unexpected occasions a balanced 50Hz grid. In this manner, the contracting balancing capacity goes parallel to the volume trading, being another way to take advantage for power generation units.

2.2 Trading participants

At the EPEX SPOT market various players are found, from banks, energy trading companies, industrial consumers, produces, suppliers to TSOs.

On the generation side, the share of the energy sources in gross German power production in 2018 was dominated by lignite with 22.5%, then wind onshore 14.5% plus wind offshore 3%, follows natural gas with 12.8%, same share of hard coal with 12.8%, then nuclear 11.7%, solar 7.1%, biomass 7%, hydro power 2.6% and other sources like waste, pump storage or geothermal with very small shares (CLEW, 2018). As shown later, in the analysis of the determinants of the electricity price, the wind plays an important role.

On the consumption side, dividing it by consumer groups, the major participant is the industry sector with a 46.9%, follows by the trade and services sector 26.6%, households in third place with 24.3% and finally the transport sector 2.3% (mainly trains) (CLEW, 2017). This huge part of the consumption of business activities means that on holidays or weekends, consumption drops to almost half that of working days.

Within Europe, Germany is the largest economy and power market in terms of electricity consumption. Its geographical position has a large number of neighbors, which allows commercial exchanges with Sweden, Denmark, Poland, Netherlands, Belgium, Czech Republic, France, Austria and Switzerland. It must be noted that currently the resulting physical flows between countries are strongly limited by the capacity of the grid connections.

2.3 Power price evolvement

After the general conditions of the market for electrical energy were examined in the last sections, the content of this chapter is to elaborate the components of the electricity price, identifying the significant cost drivers. There follows from the theoretical cost to the price evolvement in the auction mechanism of the Day-Ahead exchange.

In most industrialized countries, electric power has been provided by large central generating facilities that serve a large number of customers, often located far from the point of consumption. The economics of central station generation is largely a matter of costing, which currently encourage a transformation by the branching out to renewable technologies. As with any other production technology, generation entails fixed and variable cost. The fixed costs of power generation are essentially capital costs and land. It includes the cost of building the station and costs of obtaining permits, environmental approvals, insurances, and so on. Variable costs or operating costs or mostly referred to as the “marginal costs”, include fuel, CO2 emitting certificates, labor and maintenance costs, and depend on how much electricity the plant produces, cost per unit each MWh.

In day ahead trading, demand and supply match through an auction mechanism. In the simple structure, for the following day, the market agents submit offers to buy and to sell, composed by quantity-price pairs for each trading interval. After the gate closure at 12:00, a merit order is established ranking offers to buy by decreasing prices and offers to sell by increasing prices. In such a way, aggregate demand and supply curves are built and their intersection determines market equilibrium (price and volumes). All the bids with a price lower than the equilibrium or clearing price are accepted.

This market uses marginal cost based price, in other words, ordering generation technologies in increasing order of their marginal costs. The price is set by marginal costs of the last producer needed to cover all load, usually a fossil fuel power plant. For renewables, fuel is generally free with the exception of biomass in some scenarios, and the fuel costs for nuclear are actually very low. For these types of power plants, labor and maintenance costs dominate marginal costs. Due to these varying costs structures each type of power technology will run to cover base- or peak-loads, but nevertheless renewables are set out by the lay EEG and are always feed-in to the grid, except in extreme conditions or emergencies.

Such a scheme works well as long as there is some predictability in the system, but as it will be highlighted later, it is real challenge. However, if fossil fueled plants are increasingly replaced by renewable sources of energy, the system price is likely to decrease. The most appropriate costing methodology for electricity markets and the future electricity price development assuming the major coverage of the consumption by the new renewable energy production, are topics of debate at present.

3. Analytic Approach

This section is dedicated to describe various approaches to analyze the electricity markets and to forecast the prices. It is important to understand various approaches and models in order to do price forecasting. Following section gives a broad overview of such models. Although there are wide variety of approaches available, it is essential to mention that so far there is no single model which is considered as superior for all markets and market players.

3.1 Overview of forecasting models

- Production Cost Models

Production cost models were used for medium and long term horizons in which the existing and planned generation units are stacked up in the order of their operating costs to meet the demand estimates. One of the limitations of this model is that this does not consider the strategic bidding practices and is not suitable for the deregulated competitive markets where there are lot of uncertainties and gaming. However, strategic producing cost model (SPCM) is available which considers the strategic bidding practices (Weron, 2006, p. 102).

- Game Theory Models

Like SPCM, game theory models can be considered as a modification of production cost models with inclusion of strategies of market players. In game theory models, the strategies of different market players are modelled to find a solution. These models can predict expected price levels in markets without much price history. The known supply and market concentration will be used to predict the price. However, this model involves many assumptions and a wide range of results are possible with the change in assumptions (Weron, 2006, p. 102).

- Fundamental Models

Fundamental models try to build the exact model of the demand and supply side. In this model, the actual dispatch is simulated. The functional relationship between fundamental drivers (weather, load etc.) is stated and each of these drivers are independently modelled using statistical, econometric or non-parametric techniques. Many fundamental models are proprietary in-house products and details are not disclosed. However, the fundamental models require extensive data (Weron, 2006).

- Artificial Intelligence Models

These models can deal with complexity and non-linearity and suitable for short term price forecasting. In these models, the input-output relationships are mapped without exploring the underlying process to construct the model. Artificial Neural Networks (ANN) have been the most popular one. However, the artificial models are criticized not to be intuitive and too complex to be interpreted by the modeler (Weron, 2006, p. 106) (Duffner, 2012, p. 30).

- Statistical Models

In statistical models, the price is mathematically modelled as a function of different factors. Time series models like multiple regression models, autoregressive models (AR), moving average models (MA), autoregressive moving average models (ARMA), autoregressive integrated moving average models (ARIMA) and generalized autoregressive conditional heteroeskedsticity (GARCH) are some of the statistical models widely used (Weron, 2006) (Duffner, 2012). In this study, multiple regression model has been used for price forecasting.

3.2 Multiple Regression and Ordinary Least Squares

In multiple regression models, a dependent factor is mathematically modelled as a linear function of various other independent factors and an error term. Here, the price is the dependent factor and the factors like demand, wind feed-in, solar feed-in, fuel price etc. are the independent factors.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_K X_K + e$$

Equation 1 Multiple regression linear equation (Duffner, 2012, p. 81)

In the equation, Y is the dependent factor β_0, β_1 … β_K are the intercept and coefficients and e is the error term.

The ordinary least squares method is used to solve this equation in order to find the intercepts and coefficients. When the estimators for the coefficients and intercept are obtained, it means that the error term (e) ideally should be zero. Under this assumption, in order to solve the equation, the square of the error term which is called the least square function (L) is minimized with respect to β_0, β_1 … β_K will give k+1 equations with which the estimators of β0, β1 … βK are obtained.

3.3 Multiple Regression Method: Preconditions

There are some assumptions which have to be fulfilled before doing multiple regression analysis. They are the following.

- There is a linear relationship between the dependent variable and the independent variables.
- There is no perfect linear relationship between the independent variables. That means there should not be multi-collinearity between the independent variables. If there is such a linear relationship, any one variable has to be eliminated from the model. Inclusion of the two variables will result in high standard errors and less significant calculation of the estimators for the coefficient.

However, minor deviations from these assumptions will not affect the results much.

3.4 Autoregressive model

A simpler method for conducting a forecast is the use of an autoregressive model. As explained in (Jason Brownlee, 2017), an autoregressive model based its estimations on the previous data of the variable we are trying to forecast. It is also said that this method should be used only when the historical data of variable to be forecasted shows an autocorrelation, trend or seasonal variation; it is specially used for short term forecasting or as a reference point for more elaborated forecasting methods.

By using this method for forecasting the spot price of the electricity in the German Power Market we would be assuming that price can be represented as a linear combination of previous prices; which can be considered as a risky approach if taking into account the high volatility of important factors for the German power system, like wind speed and global solar irradiation, which can change drastically from one week/day to another, possibly occasioning a considerable deviation of the results. Thus, the autoregressive model was ruled out as a reliable method for conducting our task.

$$L_t = \sum_{i=1}^{p} \theta_i L_{t-1} + \varepsilon_t$$

Equation 2 General Formula autoregression

The order of the model tells how many lagged past values are included.

L_t is the forecast value

L_{t-1} are the lagged past value

ε_t are a prediction error

θ_1,..., θ_p are the unknown coefficients

4. Data Requirement

The accuracy of the forecasting results not only depend on the algorithm used, but also on the variables used for the modeling and their quality.

4.1 Factor influencing the forecast

The spot price in the Day-Ahead power market, not only in Germany but in any market around the world is influenced be the technical characteristics of such markets and the external events inducing changes in their performances. As mentioned in (Weron, 2006, p. 106), some drivers of the electricity price are the time of the year, the day of the week and the hour of the day (special social events and holidays are examples of events that have a considerable impact in the demand, and subsequently in the electricity price), forecast loads, plant availability, grid traffic, weather (Temperature, Humidity, Wind Speed, Cloudiness, Solar Irradiation), fuels prices.

When talking about electricity price forecasting in the long term, it is also imperative to consider possible changes in the grid extension and new generation plants, as well as the national targets, new laws and policies.

It is important to highlight that the factors to be chosen and their degree of influence in the electricity price will depend on the specific characteristics of the market under study; the generation technologies, the share of industrial, commercial and domestic consumption are example of how electricity markets can differ from each other. It is significant owing to the fact that the driving factors to be used in the forecasting must be relevant for the that specific market.

4.2 Exogenous variables selected and data collection

Taken into account the characteristics of the German power system, its energy mix and the analysis of the historical data, the factors with the highest impact in the electricity price, and which will be later used for the forecasting exercise are presented in the table below.

Variable	**Source**	**Variable**	**Source**
Price	SMART.de	Coal Price	API: market quote*
Demand	SMART.de	Neutral Gas Price Index	TTF: market quote*
Hydropower Infeed	SMART.de	NetConnect Germany Gas Market	NCG: market quote*
Wind Infeed	SMART.de	Currency Exchange Rate US-EU	FXUSEU*
Solar Infeed	SMART.de	Availability Nuclear Power Plants	EEX transparency*
Import	SMART.de	Availability Lignite Power Plants	EEX transparency*
Exports	SMART.de	Availability Gas Power Plants	EEX transparency*
Demand Forecast	Wattsight*	Availability Coal Power Plants	EEX transparency*
Wind Feed-in Forecast	Wattsight*	CO_2 Certificates	EUA: market quote*
Solar Feed-in Forecast	Wattsight*		

*Table 1 Exogenous variables historic value and forecast sources. (*source provided by the Lectures)*

The forecast values of these variables for the target week (Jan 21-27,2019) were not available online. A first trial for solving this issue was the use of an autoregressive model for obtaining the value of the exogenous variables, nonetheless as explained in section 3.4 this approach has certain limitations and was ruled out. Finally, the forecast values of demand and other exogenous variables were provided by the Lectures. The sources are shown in the above table. Since the forecast of coal price, currency exchange rate and emission price were not available, latest values of Jan 17, 2019 were used.

Towards the end, the historical and forecast values of the exogenous variables listed in Table 1 were available for conducting the ultimate goal of forecasting the day ahead price of the German spot market from the 21st January 2019 until 27th January 2019.

5. Data Understanding

Once a reliable historical data has been acquired for the purpose, it continues by analyzing in depth the behaviors of these variables and their critical points. By understanding the price evolvement in the past, a modelling tool to forecast will have a more solid base to run.

5.1 Natural trends

There are many variables that influence the electricity price, but these relationships are not visible in a simple view, therefore it is advisable to start with the most remarkable drivers, such as consumption and renewable energies. By Solar Energy the influence is not as strong as the wind energy because the hours of sun coincide with hours of high consumption. In the website EPEX SPOT under Day-Ahead auction, different block-prices are found; its names already say something about main drivers and peculiarities of different periods of time during the day. For example, Off-peak 1 (01-08), Early Morning (05-08), Business (09-16), Sun Peak (11-16), etc. Patterns or similar trends are repeated each day.

5.2 Extreme/remarkable situations

Analyzing specific cases can draw conclusions that clarify the most typical trends. A high price case and a very low price case have been examined. The 27th of February 2018 at 7:00, the load almost came to its highest level; no sun was there, almost no wind and low temperatures. This caused the price to shoot up to 79 Euro/MWh in a course of one hour because of the high dependence on conventional power plants.

Another extreme situation which can explain the behavior of the power market was the first of January at 7:00, almost the opposite of the previous example, there was very low load because it was a holiday, a storm which means much wind energy, even the sun began to rise and the electricity price fell so much that it reaches a negative number, -71 Euro/MWh. This situation of a negative prices lasted three hours long. The load was 80% covered by renewables nevertheless still an overproduction ran by conventional, which has been exported.

5.3 Negative prices

Negative prices are a price signal to reduce production that occurs when a high inflexible power generation meets low demand. Inflexible power sources can't be shut down and restarted in a quick and cost-efficient manner. They have to compare their costs of stopping and restarting their plants with the costs of selling their energy at a negative price, which means paying instead of receiving money.

5.4 Fuel and emission markets

Conventional power plants will play a fundamental role in the energy transition, because its function as back-up ensure and support the system. In periods of rapid variation of the demand, during peak-loads or during the absence of renewable sources, fossil fuel based power plants determine the price level of the electricity. Here the marginal costs owing to fuel and CO_2 emissions costs drive on the price of electricity on the Day-Ahead market.

5.5 Weather impact

It was already mentioned how the wind and sun are influencing the electricity price, but the temperature is also another important factor. When the temperatures fall below zero, or when they generally fall sharply signaling the arrival of winter, the need of heating is reflected in the energy system.

The Combined-Cycle Gas Turbines offer a bi-product which is a double profit or a reduction in the profit margin, nevertheless hence an increase in the necessity of flexible operations to avoid injection of power to the grid when spot prices drop below marginal costs.

5.6 Neighboring electricity

As mentioned in the chapter "electricity markets", in Europe the electricity power sector is regulated in many different ways, almost each country has its own regulatory model. These differences give rise to attractive prices to import or export electricity. Whenever the prices are different, the capacity of the grid-connections will be used in their entirety. However, there is limitation after a certain extent of power flow since this may cause grid congestion. Ideally, if all the power supply in Europe is integrated into a single copper plate and then distributed, the European power market will have good flexibility to enable power flow between countries.

6. Data preparation

In any exercise which consider the use of large amount of data, one of the processes to be conducted in the preparation of the data before used is the mathematical model. The data downloaded from open sourced web pages usually include missing or wrong data which need to be corrected.

6.1 Spikes

In the particular case of electricity price forecasting, when analyzing its historical data, the abrupt and generally short-lived and unanticipated extreme changes due to extreme load fluctuations caused by severe weather conditions often in combination with generation outages or transmission failures, needs to be pre-processed before conducting the forecast (Weron, 2006); as explained by the author in the previously referred source, the causes of such events are random events and it is highly probable that they will not occur again in the future at the same point in time; nonetheless, these events are an intrinsic characteristic of the power market and should not be totally eliminated. "*Instead of excluding them, we can limit their severity or damp all observations above a certain threshold"* (Weron, 2006, p. 107).

In (Weron, 2006) it is stated that the bidding strategies used by the players can also cause the appearance of spikes in the electricity prices. More specifically the author claims that.

> *Since electricity is an essential commodity for many market participants, some are willing to pay almost any price to secure a sufficient and continuous supply of power. As a result, some agents place, on a regular basis, bids at the maximum allowed level for the amount of electric power they anticipate to need for that hour. Recall that in uniform-price auction markets the spot price is what a buyer has to pay for each unit of power independent of what he or she did bid initially as long as the bid was above (or equal to) the spot price. Hence, with this type of strategy, the worst-case scenario is that a buyer has to stick with the high prices for a maximum of 24 hours. After this period, he or she is free to try to get power cheaper from alternative sources. With this type of bidding strategies, there will always be some buyers that are willing to pay a considerable amount in order to cover their need of electricity. And since the suppliers are aware of these strategies, they place their bids accordingly, to maximize their profits.*
>
> (Weron, 2006, p. 32)

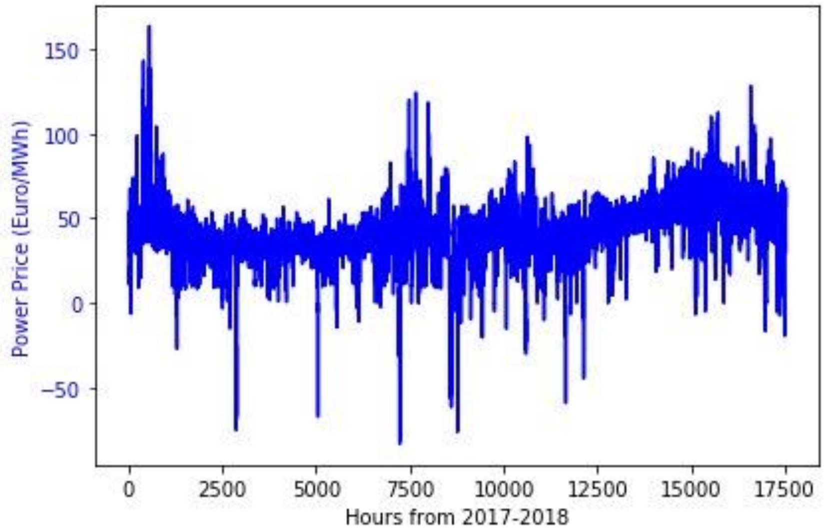

Figure 1 Day Ahead Prices in the German Power Marker Autocorrelation from 1st January 2017 until 31th December 2018

In Figure 1 some spikes prices can be observed, both, abrupt increases and abrupt decreases in the price. Negatives prices are related to a very high increase on the wind feed-in, while abrupt increases are more related to the causes exposed in the paragraph above. It is neither “possible” nor necessary to understand the reason of all these unexpected changes if they represent a clear deviation of the natural trend of the price.

To limit the influence of these events in our forecast, thresholds of maximum 120 Euros/MWh and minimum -30 Euros/MWh were established; in other words, all historical values out of this range were limited to these values.

7. Modeling

For implementing the selected mathematical models to be used for forecasting the demand, wind and solar feed-in, and electricity prices, a programming approach in python was chosen. The reason for such decision was the flexibility for manipulating large amount of data and the rapid implementation of different forecasting approaches offered by this programming platform.

7.1 Demand Forecast

The autoregressive model described in section 3.4 was implemented in python for carrying out this task. Before implementing the method, the correlation of the historical data from the 1st of January 2016 until the 14th of January 2019 was evaluated. The results are shown in the following graphs.

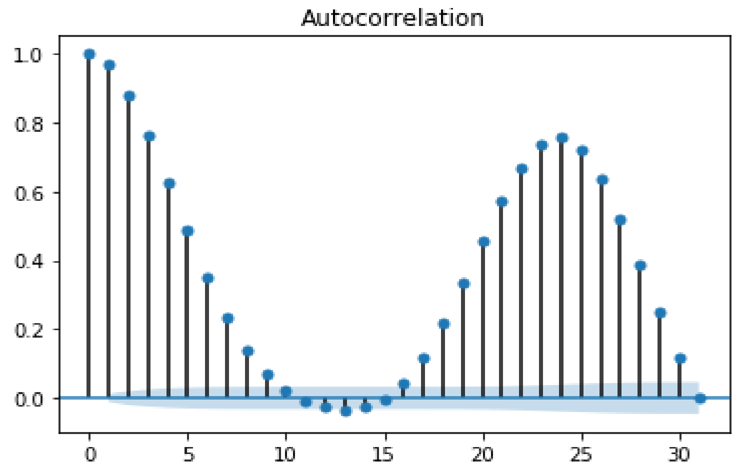

Figure 2 Autocorrelation Analysis of the Hourly Demand in the German power system- 32 Hours Lagged

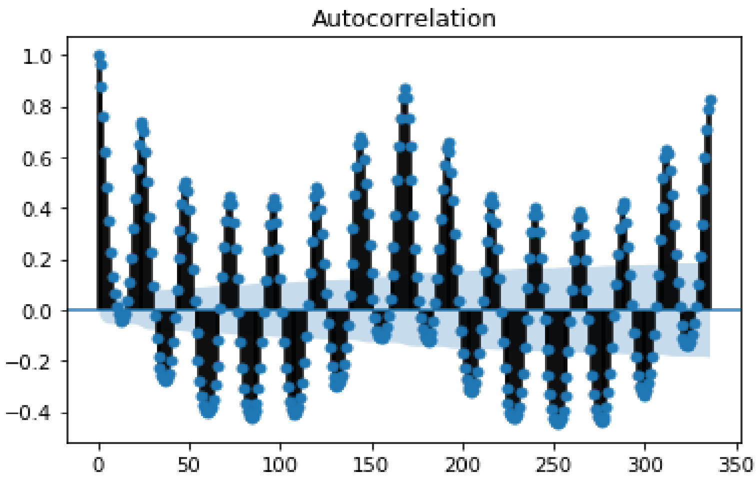

Figure 3 Autocorrelation Analysis of the Hourly Demand in the German power system- Two Weeks Lagged

From these results it can be concluded that there exists a high correlation between the demand a certain hour and the demand in the previous three hours. Similarly, the demand is highly correlated with the demand in the previous weeks, same day, same hour. Once the correlation among the historical data was proved, an autoregressive model was used for forecasting the demand for the target week (168 values). The programming code presented in (Jason Brownlee, 2017) was used for conducting this task. For a more detailed description of the code refer to the Appendix A.

The first step was to obtain the optimal lagged past values necessary for autoregressive model and their associated coefficients. The optimal lagged past values were estimated in 48, and their coefficients are shown as follows.

```
Lag: 48
Coefficients: [ 1.15997503e+03  1.73258236e+00 -1.02178006e+00  2.84441298e-01
  1.20696948e-02 -4.36837602e-02  5.03785126e-02 -2.71978035e-02
  3.46691254e-03  4.40442184e-03 -4.24632586e-02  5.02668382e-02
 -5.09293631e-02  8.32546673e-02 -1.54560735e-01  7.70661713e-02
 -2.68875812e-02  3.31837797e-02 -1.69498534e-03  5.25207277e-02
 -5.70586682e-02  1.78771487e-02 -3.71820460e-02  1.09518187e-01
  5.41952780e-01 -1.01627105e+00  5.19703813e-01 -5.96841246e-02
 -8.69682260e-02  5.69910832e-02 -9.14650301e-02  3.55896146e-02
  2.21442557e-02 -6.79174807e-03 -1.25513702e-03  5.99778196e-02
 -7.67878147e-02  6.19087887e-02 -3.75710376e-02  6.31685904e-02
 -4.68744864e-02 -3.30254241e-02  3.40872603e-02 -6.43292709e-03
 -9.31022573e-03  4.53464505e-02 -5.53700108e-02  2.56427340e-02
 -6.86565708e-03]
```

The forecasted demand using the autoregressive model was compared against the real demand and the forecasted demand from SMARD.de, and the forecasted demand provided by the Lectures (from wattsight). The results are presented in the following graphs.

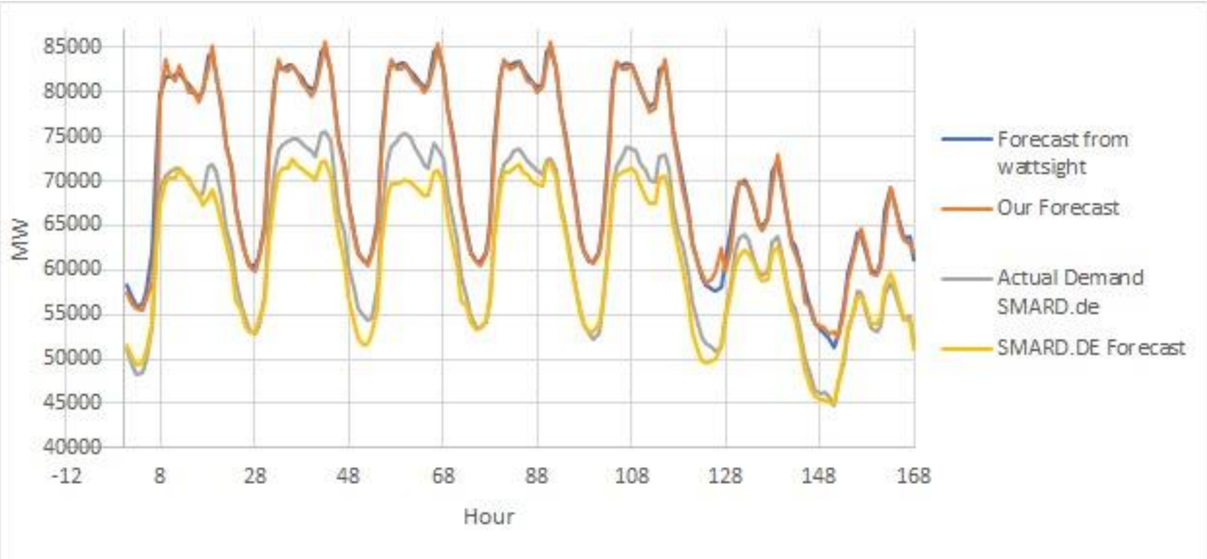

Figure 4 Actual and Forecast Demand Comparison – Target Week

The results' comparison shows a similar shape during the entire week, but an almost constant difference in the between the values from SMARD.de values, and the values forecasted by us and the ones provided by the Lecturer (from wattsight).

The explanation to these discrepancies in the result can be the timeframe of the forecasts; while our forecast used actual values until the 14th of January, completing the period until the 20th of January with the forecasted values provided by the Lecturer (from wattsight), the forecast developed by SMARD.de probably used actual values until the 19th or 20th of January, improving this way the accuracy of the results.

7.2 Forecasting Wind and Solar

One of the biggest challenges of forecasting electricity prices is the weather uncertainty, especially in countries with large amount of renewable generation or temperature sensitive demand. This also implies that weather is one if not the main driver prices.

The first effort was to use a linear regression analysis with the wind speed as the independent variable. Before starting the limitations were clear; wind speed average for the whole country, no wind direction and geographical variations considered or even the height of the wind turbines. The next figure 5 shows the results for one week, predicted wind feed-in.

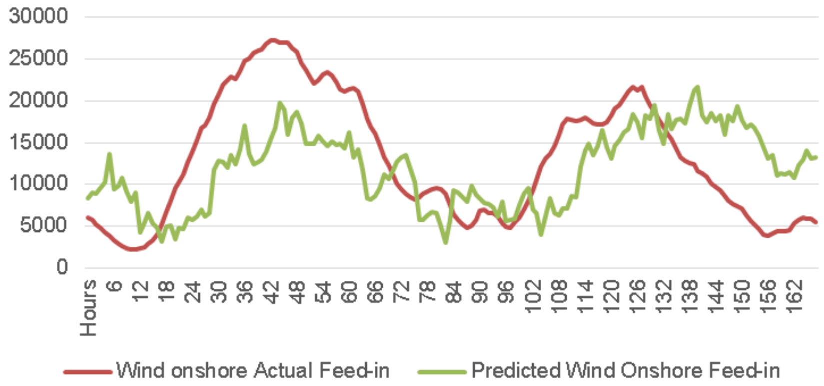

Figure 5 Forecast of solar feed-in for last week of 2016 using 2015 and 2016 data. (Own developed in excel from results)

Due to the due to the low quality of the results and the unreliability of the model, for this study case, the data source is from Wattsight provided by the lecturers. This task has its own specialists, meteorologists responsible for a large amount of data, to be able to delivery results that help forecasting. The weather forecast has a reasonable accuracy three to four days ahead, after that everything becomes very blurry.

The location of the wind and solar installed capacity and its own weather conditions has to be considered by grids, see below figure 6. The aggregation is done with a nearest neighbor algorithm.

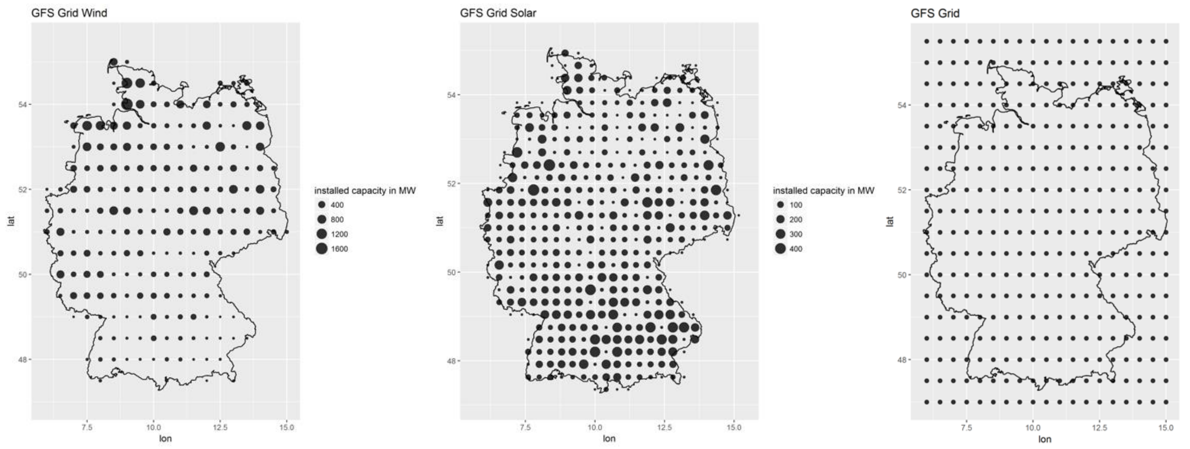

Figure 6 large scale gfs models (GFS)

7.3 Price forecast using multiple regression analysis

As discussed earlier, this study has used multiple regression analysis for price forecasting. The target is to forecast the day-ahead prices for the week of Jan 21-Jan 27, 2019. This means electricity prices of 24x7=168 hours of this week is to be predicted. The historic data used in this study is from Jan 1 2017 to Jan 14 2019. In this study, two approaches will be used for forecasting the prices of these 168 hours. They are

- One regression model to predict all hours of the week. One equation will be developed for all the hours of the week using the historic data (Python code in Appendix 1 10.3).

- 168 regression equations to predict each hour of the week independently. Since some hours of a day are more predictable than others, this will give better results (Python code in Appendix 1 10.1).

Before predicting the target week prices, the two approaches will be tested for predicting the prices of December 2018. In order to predict prices of December 2018, the actual values of fundamental and exogenous variables of the forecasting week will be used. However, in real case, when predicting the target week (Jan 21-Jan 27, 2019), the forecasted values of the fundamental and exogenous variables will be used.

7.4 Checking the preconditions of multiple regression

The linear relationship between the price and other independent variables like demand, wind feed-in, solar feed-in, fuel prices, imports and exports, emission prices and plant availability were checked. These were found to have good correlation with historic prices with some exceptions like that of plant availability (Figure 7). Plant availability does not show big variations throughout the two years and the price does not fluctuate much with change in plant availability. However, the plant availability is used in the modelling for learning purpose.

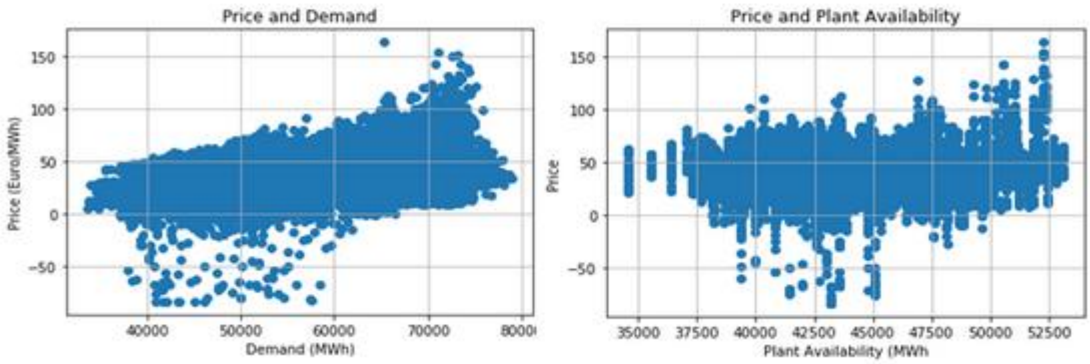

Figure 7 Plots of historic price with demand and plant availability (Own developed in python)

The multi-collinearity between all the fundamental and exogenous variables were checked. The coal prices and gas prices were found to be linearly correlated (Figure 8). This means that either coal price or gas price has to be eliminated from modelling.

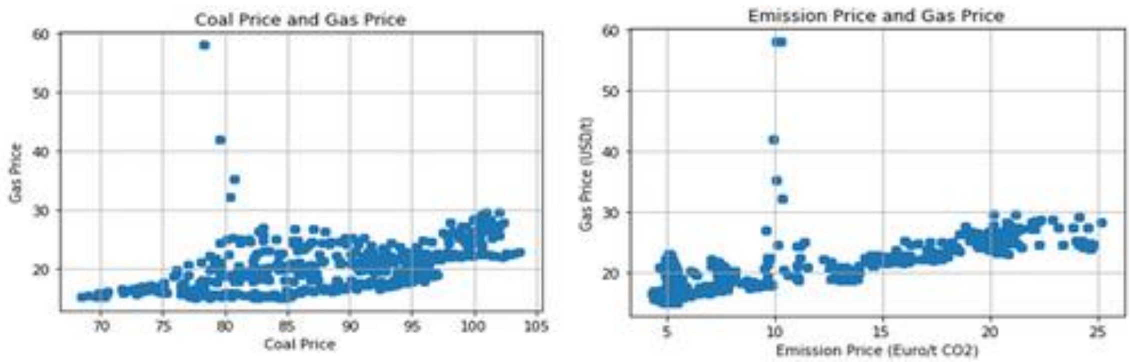

Figure 8 Linear relationship of gas price with coal price and emission price (Own developed in python)

It was also observed that gas price is linearly correlated with emission price. Hence, only coal price is used as the factor of fuel price in the modelling. Other variables were not showing any linear relationship between them.

7.5 Modelling and Evaluation

Finally, the price was modelled using the demand, wind feed-in, solar feed-in, coal price, emission price and plant availability as the independent variables using the two approaches: 168 equations approach and single equation approach. Figure 9 illustrates the forecasted prices using 168 equations approach and actual prices for December 2018.

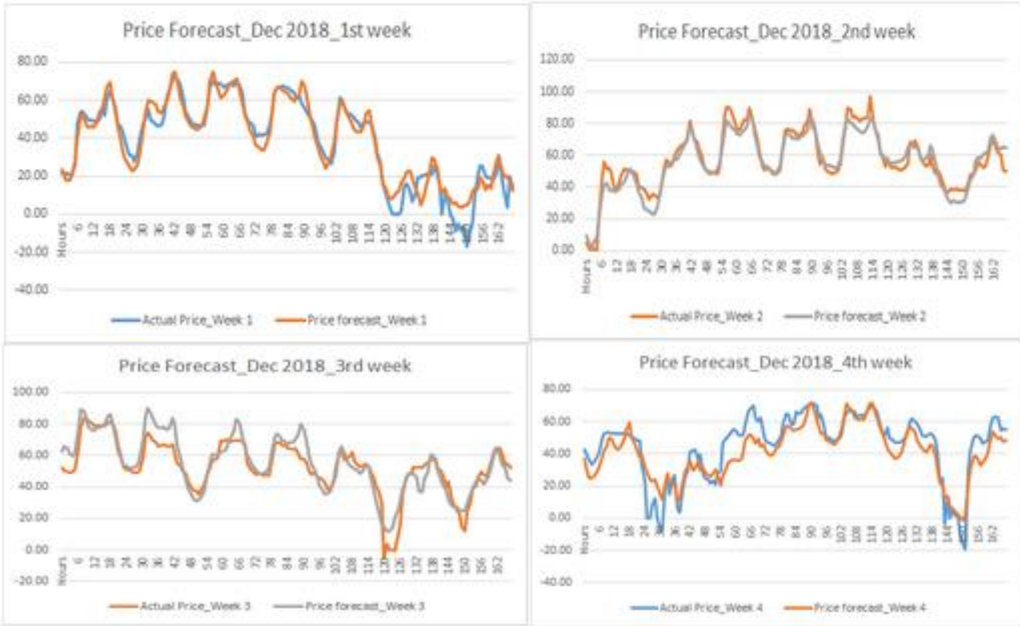

Figure 9 Comparison of actual and forecasted prices with 168 regression equations for Dec 2018 (Own developed in excel from results)

This model was able to forecast the prices in December closer to the actual prices. However, the model was unable to predict the negative prices as these happen due to extreme situations. The mean absolute error for four weeks which is calculated by the Equation 2 is found to vary between 1.5 euros and 14.3 euros with an average of 5.8 euros (Figure 8).

$$MAE = T^{-1} \sum_{t=1}^{T} \left| P_{t,fc} - P_t \right|$$

Equation 3 Mean absolute error (Duffner, 2012, p. 35)

Here T is 4 since there are 4 weeks in a month. $P_{t,f\,c}$ is the forecasted price and Pt is the actual price.

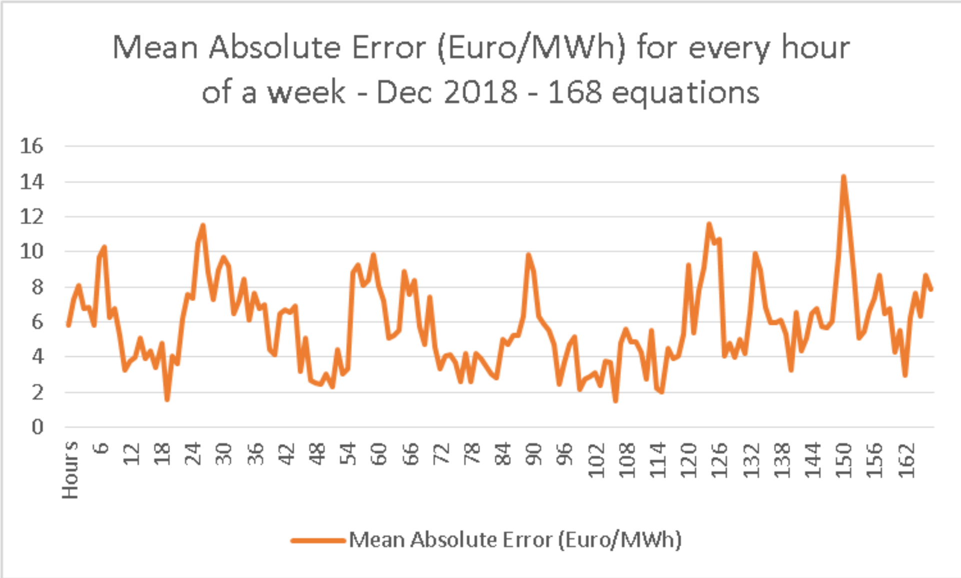

Figure 10 Mean Absolute Error for December 2018- 168 equations approach. (Own developed in excel from results)

Next, the prices of December were forecasted using single equation approach where just one regression equation was used to forecast the prices. The forecasted prices could be seen in Figure 11.

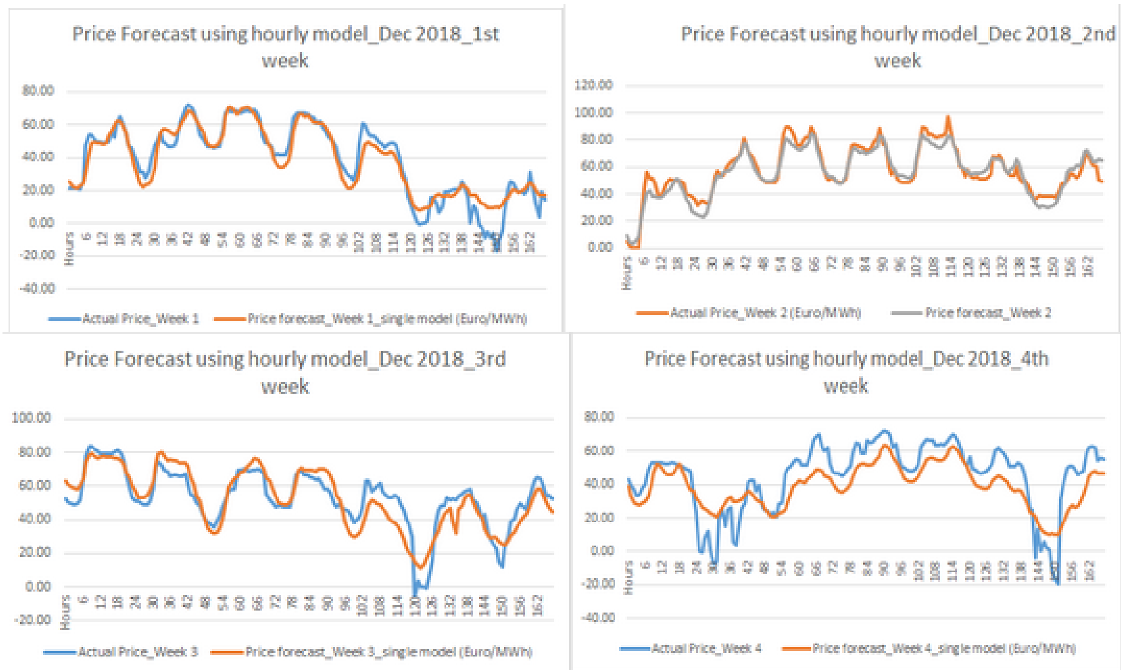

Figure 11 Comparison of actual prices and forecasted prices with single equation for Dec 2018. (Own developed in excel from results)

It is clear that the forecast using 168 regression equations (Figure 9) gives better forecasts. The mean absolute error for four weeks using this method is found to vary between 1.9 euros and 18 euros with an average of 7.3 euros (Figure 12).

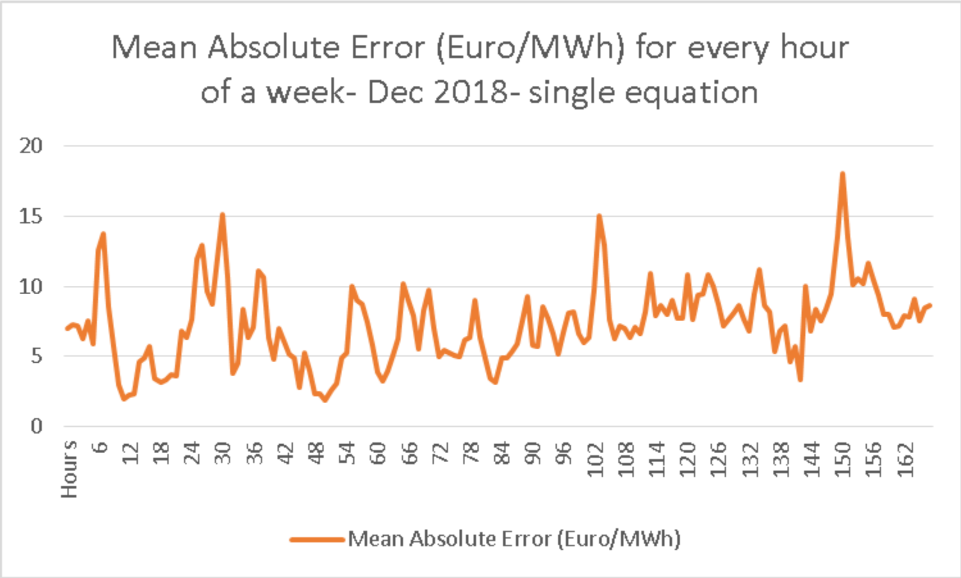

Figure 12 Mean Absolute Error for December 2018- single equation model. (Own developed in excel from results)

8. Modelling and Results for the target week

Finally, the 168 equations approach only was used to forecast the day-ahead prices of the Jan 21-27, 2019 week since this gives better results (python code in Appendix 1 10.1). The data used was from 2017 till Jan 13, 2019. Contrary to the earlier forecast of December 2018, the forecasted values of independent variables were used. Since the forecasted values of coal prices, USD-Euro conversion rate and emission prices were not available, the latest values of Jan 17, 2019 were used.

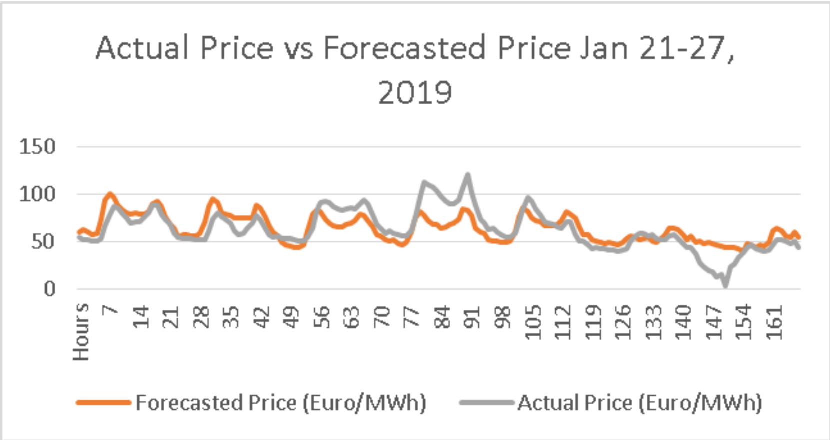

Figure 13 Comparison of actual and forecasted prices for target week Jan 21-27 2019. (Own developed in excel from results in python)

The average for the week was forecasted to be 64.48 Euro/MWh which came very close to the actual average which was 63.54 Euro/MWh. The Figure 13 shows the comparison of the forecasted and actual values. The model was unable to forecast the peak prices which underlines the limitation of multiple regression analysis in predicting the extreme situations. The forecast had an average absolute error of 11.36 Euro/Mh which was mainly due to the large errors during the peak positive and negative price situations. However, for most of the hours of this week, the forecasted prices came close to the actual price. Since this forecast also depends on the forecast of fundamental and exogenous variables like demand, wind feed-in, solar feed-in etc, the inaccuracy in predicting these variables affects the forecast. This is clearly evident in this forecast. Figure 14 shows the comparison of actual wind feed-in of Jan 21-27 and the wind feed-in forecast. We used the average of wind feed-in forecasts from wattsight which are made using ECMWF and GFS weather models. The inaccuracy in the forecast is clearly evident in the figure. Also, the larger errors in the price forecast is occurring when there are large errors in the wind feed-in forecast. There was higher expectation of wind feed-in which forecasted lower price between hours 56 and 91. However, the low actual wind feed-in resulted in a higher price than the forecasted price.

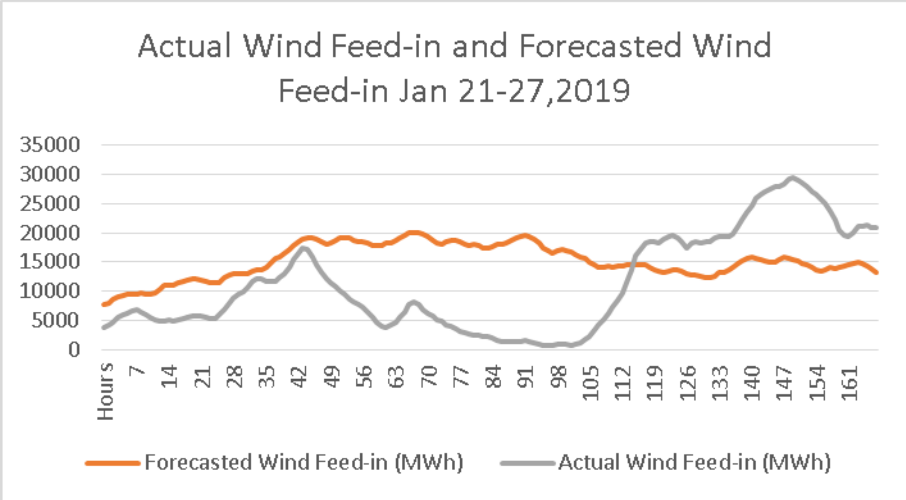

Figure 14 Comparison of actual wind feed-in and forecasted wind feed-in for Jan 21-27. (Own developed in excel from results in python)

Then, instead of the forecasted independent variables, the real values of the independent variables were used in the model to forecast the price. The comparison of this forecast and the actual prices are shown in Figure 15. This forecast is very close to the actual prices of the target week. Even though in a real scenario such a forecast is not possible, this shows that the independent variables chosen for modelling are appropriate. This also indicates the dependency of the price forecast especially on the weather data. The wind feed-in and solar feed-in forecast depends on the forecast of wind speeds and solar irradiation. Hence, challenges in the weather forecast undermines the accuracy of the price forecast.

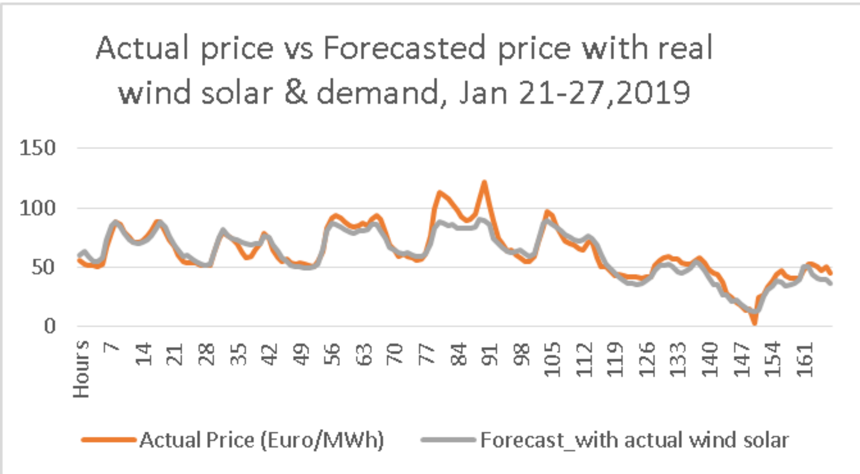

Figure 15 Comparison of actual price and forecasted prices using real wind and solar data. (Own developed in excel from results in python)

8.1 Regression Coefficients

The python simulation results give 168 sets of regression coefficients for 168 hours of the target week. These along with intercepts and forecasted prices have been listed in the appendix 1 10.2. Table 2 shows the calculated coefficients of the independent variables for Monday (Jan 21) during various times of the day. Regression coefficients indicate the change in the power price when there is a unit change in the independent variable. For example, the power price at 12:00 pm will change by 0.9029 Euros if the emission price changes by 1 Euro/t CO2. Therefore, the coefficients can tell how much impact a unit change in the independent variable can make on the power price. Here, at midnight 1 am the coefficient for solar feed-in is approximately 0 and it is almost close to the coefficient for wind feed-in at 12 pm in the noon. This is realistic since solar will play a significant role at noon and it has no effect on power price at midnight.

Time	Demand Coeff	Wind Coeff	Solar Coeff	Coal Price Coeff	Emission Price Coeff	Plant Availability Coeff
01:00	0.0011	-0.0017	5.7E-14	0.1406	1.1391	0.0004
12:00	0.0014	-0.0015	-0.0014	0.3545	0.9029	-0.0003

Table 2 Coefficients of independent variables at various time of the Monday of the target week (Jan 21). (Own results from modelling in python

8.2 Statistical Parameters

We have used the following statistical parameters in order to evaluate the significance of the coefficients predicted.

- R squared

It is the square of the regression coefficient. This indicates the correlation between the dependent and independent variables. Since we have many independent variables, there is adjusted R squared value should be ideally closer to one to have good correlation.

- Standard Error

This indicates how far are the data points away from the fitted line on an average in dependent variable units (here in Euros).

- p value (probability)

For the forecast of coefficients to have significance within 95 percent confidence interval, the p value should be less than 0.05. Higher p value indicates multicollinearity or lack of correlation with price.

Using statsmodel library of python, these values were calculated for every hour. Figure 16 shows an example for hour 164 of the target week. Here the adjusted R-square is 0.832 which shows good correlation between dependent and independent variables at this hour. P-value is less than 0.05 for all independent variables except for plant availability. The p-value of 0.1 for plant availability is due to lack of correlation between availability and price which can be seen in Figure 7.

```
164
                            OLS Regression Results
==============================================================================
Dep. Variable:         Filtered Price   R-squared:                       0.842
Model:                            OLS   Adj. R-squared:                  0.832
Method:                 Least Squares   F-statistic:                     88.80
Date:                Wed, 06 Feb 2019   Prob (F-statistic):           8.17e-38
Time:                        11:43:26   Log-Likelihood:                -343.44
No. Observations:                 107   AIC:                             700.9
Df Residuals:                     100   BIC:                             719.6
Df Model:                           6
Covariance Type:            nonrobust
======================================================================================
                        coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------
const               -10.3013     13.693     -0.752      0.454     -37.467      16.864
Demand                0.0010      0.000      4.264      0.000       0.001       0.001
Wind                 -0.0014   8.17e-05    -16.746      0.000      -0.002      -0.001
Photovoltaics        -0.0025      0.001     -2.971      0.004      -0.004      -0.001
Coal Price            0.4283      0.098      4.391      0.000       0.235       0.622
CO2 Certificates      0.9363      0.156      5.990      0.000       0.626       1.246
Availability (MWh)   -0.0006      0.000     -1.661      0.100      -0.001       0.000
==============================================================================
Omnibus:                        0.818   Durbin-Watson:                   1.385
Prob(Omnibus):                  0.664   Jarque-Bera (JB):                0.472
Skew:                           0.144   Prob(JB):                        0.790
Kurtosis:                       3.151   Cond. No.                     1.60e+06
==============================================================================
```

Figure 16 The OLS regression results summary showing Adj.R-squared,p value, standard errors etc (Results from python modelling)

9. Conclusions

The study could perform a fundamental analysis of German Day-ahead electricity market. It is understood that the along with the fundamental driver demand, the exogenous variables like wind feed-in, solar feed-in, fuel prices, emission prices etc. also play a major role in determining the day-ahead market price. However, for most of the time, the demand and availability of renewable power sources like wind causes a lower or higher price in German day-ahead market. The study was unable to capture a particular trend in the electricity price according to changes in neighboring electricity markets.

Various forecasting methods available were studied and a price forecasting model using multiple regression analysis in python was developed in order to estimate the day-ahead electricity prices for the week of January 21-27, 2019. Two approaches for modelling were developed: single regression equation model and 168 regression equations model. Later was found to be having better accuracy since each hour is predicted independently. This model was able to forecast the prices with a reasonable accuracy. The forecasted average of the week was 64.48 Euro/MWh which came closer to the real average of 63.54 Euro/MWh. However, the model had its own limitations. The model was not able to capture the peak prices like negative prices. This underlines the limitation of multiple regression analysis which relies on the historical data for forecasting. Therefore, the unusual events which may cause huge price variations cannot be forecasted. Also, the developed model relies heavily on the forecast of independent variables like demand, wind feed-in, solar feed-in etc. Wind feed-in and solar feed-in forecast in turn depends on weather forecasts. Therefore, the challenges in weather forecast modelling undermines the accuracy of price forecasting. The inaccuracy of wind feed-in and solar feed-in forecast was clearly reflected in the forecasting model where the hours with major errors in price forecasting coincided with the hours with major errors in wind feed-in and solar feed-in forecast.

10. Appendix 1. Model Programming Code & Results

10.1 Python code for 168 equations approach (Final Model)

This code was developed based on the python code for performing regression analysis in (*Example of Multiple Linear Regression in Python*, 2019).

This code is to be run with along with the input csv files ‘Historic Data’ and ‘futurevalues_targetweek_1’. Running the code gives an output csv file ‘finalfile_targetweek’ with the intercepts, coefficients and forecasted prices for 168 hours.

```
#This block is to import important libraries to be used later for data manipulation
import pandas as pd
from pandas import DataFrame
import numpy as np

#This line is to import data from the excel file with the historic values of all dependent and independent variables
combineddataset = pd.read_csv(r'C:/Users/ASUS/.spyder-py3/Historic Data.csv')

#This line tells python about the columns available in the csv file
combineddataset=DataFrame(combineddataset,columns=['SerialNumber','Year','Hours','Filtered Price','Demand','Hydropower','Wind','Photovoltaics', 'Import','Export','Coal Price','Neutral Gas Price Index','NetConnect Germany Gas Market', 'CO2 Certificates','Currency US-EU','Availability'])
print(combineddataset)

#This line is to make python read the values only upto a certain serial number in the excel file
combineddatsetyear=combineddataset[combineddataset.SerialNumber<17833]

#Even though the dataset contains many independent variables, after a few analysis only following variables were used for modelling
useddataset=combineddatsetyear[['SerialNumber','Year','Hours','Filtered Price','Demand','Wind','Photovoltaics','Coal Price','CO2 Certificates','Availability']]

#This line is to create the title row for the output excel file
np_finalsolution=np.array(['intercept','Demand','Wind','Photovoltaics','Coal Price', 'CO2 Certificates','Availability','Price forecast'])

#This 'for' loop is to filter every hour of a week separately. The loop runs separately for each hour of a week and predicts the intercepts and coefficients
for x in range(1,169):
   useddataset_x=useddataset[useddataset.Hours==x] #This line is to take out only the price    data at hour x, hour 1 means price at 1st hour of monday

#This block finds the intercept and regression coefficients
   from sklearn import linear_model
   import statsmodels.api as sm

#Following two lines declares the independent X variables and dependent Y variable to python
```

```
    X = useddataset_x[['Demand','Wind','Photovoltaics','Coal Price','CO2 Certificates','Availability']]
    Y = useddataset_x['Filtered Price']

#Following four lines fits a regression line and prints the intercept and coefficient for every hour
    regr = linear_model.LinearRegression()
    regr.fit(X, Y)
    print('Intercept: \n', regr.intercept_)
    print('Coefficients: \n', regr.coef_)

#This block is to predict the price using the predicted values of fundamental and exogenous variables taken from the csv file 'futurevalues_targetweek_1.csv'
    exovariables=pd.read_csv(r'C:/Users/ASUS/.spyder-py3/futurevalues_targetweek_1.csv')
exovariables=DataFrame(exovariables,columns=['Hour','Demand','Wind','Photovoltaics',
    'Coal Price', 'CO2 Certificates','Availability'])
    print (exovariables)
    exovariables_x=exovariables[exovariables.Hour==x]
    predicteddata=exovariables_x[['Demand','Wind','Photovoltaics','Coal Price','CO2 Certificates','Availability']]
    print ('Predicted  Price: \n', regr.predict(predicteddata))

 #This block is to stack the coefficients,intercepts and predicted prices obtained for each hour into one single array of 168 rows which can later be exported to an excel file
    coefficient_x=regr.coef_
    intercept_x=regr.intercept_
    predictedprice_x=regr.predict(predicteddata)
    np_stackeddata=np.hstack((intercept_x,coefficient_x,predictedprice_x))
    print(np_stackeddata)
    np_finalsolution=np.vstack((np_finalsolution,np_stackeddata))

 # This block use statsmodels to predict the price directly which also prints certain important statistical parameters like p-value, standard error etc.
    X = sm.add_constant(X) # adding a constant
    model = sm.OLS(Y, X).fit()
    predictions = model.predict(X)
    print_model = model.summary()
    print (x)
    print(print_model)
    #the 'for' loop ends here

#This block finally exports the solution set back to a csv file with name 'finalfiletargetweek_1' which is essentially the output
finalsolutioninexcel=pd.DataFrame(np_finalsolution)
finalsolutioninexcel.to_csv('finalfiletargetweek_1.csv',index=True)
```

10.2 Intercepts, Coefficients and Price forecast for Jan 21-27, 2019

Day	Time	Hour	intercept	Demand C_f	Wind Feed-in C_f	Solar Feed-in C_f	Coal Price C_f	Emission Price C_f	Plant Availability C_f	Price forecast (Euro/MWh)
21-Jan	12:00 AM	1	-31.041	0.000985968	-0.001460342	-2.10E-14	0.213943937	1.018122107	0.000122324	60.72186926
21-Jan	1:00 AM	2	-43.2899	0.001075057	-0.001648278	5.66E-14	0.140548885	1.13912781	0.000450325	63.45541193
21-Jan	2:00 AM	3	-48.6991	0.00095189	-0.001662417	-2.20E-15	0.192485185	1.157143014	0.000591205	60.18310426
21-Jan	3:00 AM	4	-43.1967	0.000945726	-0.001542555	-4.92E-14	0.22346524	1.094431865	0.00039946	57.26684633
21-Jan	4:00 AM	5	-32.2124	0.001083409	-0.001510786	-0.239399364	0.24403859	1.013898508	2.33E-05	58.96773436
21-Jan	5:00 AM	6	-41.4262	0.001802084	-0.001356018	-0.006398724	0.324540446	0.953543743	-0.000599419	74.73146113
21-Jan	6:00 AM	7	-94.5361	0.001578367	-0.00125087	-0.002668923	0.377048703	1.327931692	0.000602859	94.65058068
21-Jan	7:00 AM	8	-84.1675	0.001343446	-0.001390819	-0.002392792	0.405501495	1.311838222	0.000651249	101.4960346
21-Jan	8:00 AM	9	-59.3986	0.001378082	-0.001521301	-0.001986825	0.384003059	1.158263534	0.000111516	97.74301948
21-Jan	9:00 AM	10	-51.9578	0.001432832	-0.001604265	-0.001616843	0.389135087	0.999244215	-0.000131546	89.3632707
21-Jan	10:00 AM	11	-42.5388	0.001403079	-0.001534799	-0.001435513	0.372527164	0.885370158	-0.0002756	83.45776033
21-Jan	11:00 AM	12	-40.7657	0.001406794	-0.001497153	-0.00143683	0.354539824	0.902942986	-0.000310538	80.56545175
21-Jan	12:00 PM	13	-34.0817	0.001579584	-0.001556194	-0.001427371	0.288262906	0.910082233	-0.000600255	79.4750857
21-Jan	1:00 PM	14	-47.5227	0.001530247	-0.001482084	-0.001312352	0.342004716	0.983129858	-0.000361466	80.24692743
21-Jan	2:00 PM	15	-65.9359	0.001262115	-0.001298842	-0.001050119	0.44953704	1.035328468	0.000159422	79.48739868
21-Jan	3:00 PM	16	-71.4177	0.00110332	-0.001257372	-0.000919849	0.538491632	1.091118095	0.000323523	80.09952393
21-Jan	4:00 PM	17	-75.6106	0.001008452	-0.001192914	-0.000825564	0.673025053	1.118313224	0.000333124	82.55428565
21-Jan	5:00 PM	18	-83.0804	0.000965016	-0.001265217	-0.001206552	0.764803736	1.00060627	0.000556693	89.67105365
21-Jan	6:00 PM	19	-50.2852	0.00108513	-0.00139291	-0.002061362	0.511475902	1.051907854	0.000130163	93.19469005
21-Jan	7:00 PM	20	-42.8375	0.001202714	-0.001334447	-0.00325796	0.289931552	1.101096612	2.15E-05	87.87595325
21-Jan	8:00 PM	21	-42.6319	0.000914847	-0.00110068	-0.005448048	0.316186045	1.169892936	0.000212116	76.13493499
21-Jan	9:00 PM	22	-29.262	0.000748018	-0.00103614	0.049025498	0.243551134	1.142954096	0.000201218	68.29749985
21-Jan	10:00 PM	23	-33.579	0.000762643	-0.001012553	-2.60E-14	0.262809672	1.1250959	0.000195121	63.95862383
21-Jan	11:00 PM	24	-13.3169	0.000690239	-0.001142182	3.13E-14	0.316982719	0.992464787	-0.000223113	55.59906798
22-Jan	12:00 AM	25	-36.6883	0.0008338	-0.00115661	4.89E-14	0.295305595	1.004342926	0.000193893	57.94410349
22-Jan	1:00 AM	26	-44.3486	0.00108839	-0.001180272	-2.94E-14	0.316369395	0.864248668	0.000100027	57.49005297
22-Jan	2:00 AM	27	-53.4379	0.001254076	-0.001264128	1.87E-14	0.343884523	0.781412484	0.000119312	56.29010005
22-Jan	3:00 AM	28	-56.2991	0.001342972	-0.00128811	5.35E-14	0.339813784	0.808154701	9.13E-05	56.82053138
22-Jan	4:00 AM	29	-50.687	0.001439929	-0.001200479	-0.116424858	0.328458527	0.787071401	-0.000114645	59.46370586
22-Jan	5:00 AM	30	-54.6179	0.002138415	-0.001176216	-0.007492211	0.390616337	0.577902187	-0.000827626	71.79220092
22-Jan	6:00 AM	31	-113.43	0.001890034	-0.001107233	-0.002463089	0.575278092	0.812383232	0.000312306	88.14918951
22-Jan	7:00 AM	32	-111.737	0.001645738	-0.001417796	-0.002384008	0.6639791	0.847937973	0.000494331	95.97418777
22-Jan	8:00 AM	33	-75.5729	0.001387997	-0.001577907	-0.002081209	0.55323276	1.05423335	0.00021312	92.34488148
22-Jan	9:00 AM	34	-48.0985	0.001309598	-0.001611776	-0.001735474	0.55616858	0.902556176	-0.000268985	80.69604582
22-Jan	10:00 AM	35	-66.3197	0.001468996	-0.001608296	-0.001439938	0.515421199	0.884465384	-6.18E-05	79.56682971
22-Jan	11:00 AM	36	-68.4595	0.001537728	-0.001577819	-0.001502426	0.469444342	0.965343304	-5.99E-05	78.5790429
22-Jan	12:00 PM	37	-75.6017	0.001628438	-0.00153899	-0.001298539	0.430615553	0.927816292	-3.23E-05	75.79436832
22-Jan	1:00 PM	38	-103.379	0.001791416	-0.001476714	-0.001136529	0.531512923	0.86428907	0.000129356	76.07101168
22-Jan	2:00 PM	39	-105.907	0.001664239	-0.001419632	-0.001026285	0.584446212	0.887854559	0.000260576	75.77148597
22-Jan	3:00 PM	40	-88.21	0.001434093	-0.00145033	-0.001135773	0.617712941	0.890525856	0.000214372	75.70995911
22-Jan	4:00 PM	41	-90.9388	0.001287369	-0.001443511	-0.001040678	0.838112082	0.852028849	0.000149513	75.82135956
22-Jan	5:00 PM	42	-89.2025	0.00149014	-0.001495079	-0.001596718	0.779519249	0.92631414	9.45E-06	89.29466305
22-Jan	6:00 PM	43	-69.0181	0.001517095	-0.001698199	-0.002830285	0.789457333	0.843078241	-0.000409924	86.07412721
22-Jan	7:00 PM	44	-80.8439	0.001557962	-0.001544823	-0.003685604	0.749886098	0.807154838	-0.000310341	77.10558965
22-Jan	8:00 PM	45	-70.3049	0.001428419	-0.001350397	-0.00969718	0.747177129	0.810144918	-0.000464627	66.94024218
22-Jan	9:00 PM	46	-51.0731	0.001276709	-0.001250076	-0.110594817	0.541173525	0.863158313	-0.000407314	60.81871019
22-Jan	10:00 PM	47	-42.538	0.000989654	-0.001176109	-4.36E-14	0.336241341	0.955562286	5.69E-05	56.29727564
22-Jan	11:00 PM	48	-22.4091	0.000919229	-0.001306741	-2.73E-14	0.268051942	0.884815532	-0.000146026	49.15586585

Day	Time	Hour	intercept	Demand C_f	Wind Feed-in C_f	Solar Feed-in C_f	Coal Price C_f	Emission Price C_f	Plant Availability C_f	Price forecast (Euro/MWh)
23-Jan	12:00 AM	49	-21.7408	0.000992559	-0.001277406	-1.54E-14	0.275349545	0.779464421	-0.000204107	47.37924548
23-Jan	1:00 AM	50	-22.1214	0.000924293	-0.001366343	3.07E-14	0.228681666	0.851387457	-3.33E-05	45.55139604
23-Jan	2:00 AM	51	-21.0225	0.000953318	-0.001372307	2.88E-14	0.191124119	0.854293987	-3.11E-05	43.89439677
23-Jan	3:00 AM	52	-21.2258	0.001007954	-0.001296722	-4.07E-14	0.182215011	0.839118276	-8.22E-05	44.44620571
23-Jan	4:00 AM	53	-22.3928	0.001212931	-0.001233824	-0.108831918	0.226178504	0.746931694	-0.000321413	47.63939054
23-Jan	5:00 AM	54	-36.7765	0.002153829	-0.001212396	-0.011393983	0.371811147	0.510335317	-0.001200705	62.44222873
23-Jan	6:00 AM	55	-81.5174	0.001745553	-0.001131004	-0.003358776	0.487981645	0.843272878	-6.58E-05	79.06484734
23-Jan	7:00 AM	56	-66.9221	0.001202592	-0.0013105	-0.002368737	0.502166095	1.021885078	0.000331015	83.63420773
23-Jan	8:00 AM	57	-55.3848	0.001142932	-0.001427554	-0.001898266	0.438075397	1.117044944	0.000247265	82.04055156
23-Jan	9:00 AM	58	-41.3052	0.001153779	-0.001511598	-0.001657	0.434527233	0.973846173	-3.09E-05	74.01946641
23-Jan	10:00 AM	59	-39.0898	0.001193718	-0.001506149	-0.001481187	0.468105224	0.812546153	-0.000158296	69.75980339
23-Jan	11:00 AM	60	-34.2675	0.001297293	-0.001465475	-0.001582701	0.448210182	0.705273631	-0.000351033	67.53001363
23-Jan	12:00 PM	61	-35.5269	0.001457495	-0.001409422	-0.001433362	0.442275616	0.622554442	-0.000590826	65.66290931
23-Jan	1:00 PM	62	-39.9414	0.001380162	-0.001374252	-0.001375134	0.487068374	0.661777919	-0.000467962	65.89809488
23-Jan	2:00 PM	63	-54.7347	0.00131129	-0.001385432	-0.001321917	0.588350217	0.785269661	-0.00022543	68.96323388
23-Jan	3:00 PM	64	-60.5008	0.001171627	-0.001377839	-0.00124611	0.65347878	0.853349919	-3.05E-05	70.55014734
23-Jan	4:00 PM	65	-62.6982	0.000967711	-0.001363277	-0.001275113	0.806263756	0.875289348	9.10E-05	72.36087094
23-Jan	5:00 PM	66	-78.1379	0.001041906	-0.001524619	-0.001803341	0.946294788	0.809027667	0.000227216	79.30439472
23-Jan	6:00 PM	67	-67.9551	0.001146099	-0.001625385	-0.002715495	0.884243759	0.771298056	-2.97E-05	78.37636418
23-Jan	7:00 PM	68	-63.8415	0.000997561	-0.001389168	-0.003886391	0.627899032	1.012077357	0.000271889	73.38375071
23-Jan	8:00 PM	69	-60.6056	0.00086958	-0.001181051	-0.006145051	0.549222958	1.082765281	0.000331336	65.45560188
23-Jan	9:00 PM	70	-27.08	0.000443863	-0.000940801	-0.033561401	0.384185412	1.194082633	0.000293458	57.91658686
23-Jan	10:00 PM	71	-29.3588	0.000738012	-0.000997634	-3.302640175	0.289387395	1.165004387	7.65E-05	56.74553737
23-Jan	11:00 PM	72	-6.34864	0.000684005	-0.001043673	6.11E-14	0.090403321	1.224909959	-6.50E-05	52.76063059
24-Jan	12:00 AM	73	-10.4286	0.000863292	-0.001139739	1.36E-14	0.090870451	1.10829159	-0.000112608	51.43499705
24-Jan	1:00 AM	74	-17.1673	0.000889604	-0.001142505	9.81E-14	0.077210637	1.153135339	3.64E-05	51.95181332
24-Jan	2:00 AM	75	-17.4686	0.000640861	-0.001110492	6.45E-14	0.078731466	1.198407034	0.000278401	48.49653178
24-Jan	3:00 AM	76	-14.8009	0.000542977	-0.001046349	2.61E-15	0.088789997	1.198735433	0.000291117	47.48743824
24-Jan	4:00 AM	77	-3.25383	0.000590596	-0.000986592	-0.244596877	0.100141228	1.154019012	-3.43E-06	49.23754897
24-Jan	5:00 AM	78	-5.93568	0.001298616	-0.000965666	-0.011810535	0.266437935	0.938191373	-0.00086923	60.91622708
24-Jan	6:00 AM	79	-37.032	0.001218902	-0.000919502	-0.002776579	0.35894359	1.098732069	-0.000276831	75.47465781
24-Jan	7:00 AM	80	-71.607	0.000972609	-0.001039084	-0.001566418	0.491421775	1.184220044	0.000610925	82.42936488
24-Jan	8:00 AM	81	-47.9714	0.000841951	-0.001219851	-0.001581929	0.430412157	1.194775097	0.000417072	78.65172041
24-Jan	9:00 AM	82	-27.739	0.000949863	-0.001282132	-0.001364994	0.32511553	1.128077746	-8.02E-06	73.24848783
24-Jan	10:00 AM	83	-23.9545	0.000949344	-0.001283729	-0.001219491	0.34142687	1.013400419	-0.000102526	68.84719842
24-Jan	11:00 AM	84	-29.9748	0.000974179	-0.001338477	-0.00122317	0.290316178	1.075430432	7.57E-05	68.32079389
24-Jan	12:00 PM	85	-23.9684	0.000933863	-0.001288529	-0.001119762	0.327843311	1.047506738	-8.96E-05	65.20494288
24-Jan	1:00 PM	86	-30.8742	0.000897766	-0.001264474	-0.001006022	0.32507296	1.138835222	6.72E-05	65.9671035
24-Jan	2:00 PM	87	-35.7584	0.000927254	-0.001216812	-0.000940266	0.341934082	1.159348729	7.78E-05	68.0567659
24-Jan	3:00 PM	88	-40.9214	0.00086877	-0.001155506	-0.000953614	0.424272383	1.138129697	0.000145277	69.75062561
24-Jan	4:00 PM	89	-34.4587	0.000931618	-0.001111652	-0.000974121	0.459480884	1.144630361	-0.000121121	73.97166346
24-Jan	5:00 PM	90	-37.312	0.001148263	-0.001180177	-0.001525826	0.458570836	1.114632993	-0.000253842	84.79345307
24-Jan	6:00 PM	91	-24.1216	0.001261445	-0.001200327	-0.002213932	0.402087374	0.9888564	-0.000596424	83.76283066
24-Jan	7:00 PM	92	-19.4983	0.00119163	-0.001132452	-0.003563247	0.269136218	1.122700875	-0.000517048	77.95694874
24-Jan	8:00 PM	93	-6.16657	0.000954678	-0.00095205	-0.007935957	0.322407077	1.033405628	-0.00068896	64.977434
24-Jan	9:00 PM	94	-7.86248	0.000680976	-0.000821786	-0.023755502	0.199494459	1.161917303	-0.000179929	60.93268418
24-Jan	10:00 PM	95	-19.9228	0.000741915	-0.000807678	1.39E-14	0.215837632	1.118999282	-4.55E-05	58.72197913
24-Jan	11:00 PM	96	-3.58951	0.000512701	-0.000835531	1.81E-14	0.161710037	1.139833378	-5.40E-05	52.65826328

Day	Time	Hour	intercept	Demand C_f	Wind Feed-in C_f	Solar Feed-in C_f	Coal Price C_f	Emission Price C_f	Plant Availability C_f	Price forecast (Euro/MWh)
25-Jan	12:00 AM	97	-12.7255	0.00050927	-0.000911658	2.69E-14	0.206805412	1.129088659	9.32E-05	50.87206456
25-Jan	1:00 AM	98	-13.9168	0.00055499	-0.001023196	-1.70E-14	0.122062218	1.207240953	0.000213893	51.34387603
25-Jan	2:00 AM	99	-15.9379	0.00051555	-0.001121423	-3.90E-14	0.088197091	1.274502522	0.000354964	50.15936887
25-Jan	3:00 AM	100	-21.1458	0.000446983	-0.000964271	5.22E-14	0.131801663	1.211756593	0.000447213	49.40743116
25-Jan	4:00 AM	101	-6.86685	0.000649755	-0.00089574	-0.329632459	0.145007674	1.073210926	-5.08E-05	51.66629721
25-Jan	5:00 AM	102	-4.08554	0.001022954	-0.000807893	-0.011603437	0.322742863	1.006385427	-0.000743512	60.82813714
25-Jan	6:00 AM	103	-28.1842	0.001046337	-0.000814764	-0.003131743	0.291244358	1.192053543	-0.000172158	76.95689232
25-Jan	7:00 AM	104	-23.6439	0.000893506	-0.000962893	-0.002212212	0.201428343	1.430422968	9.59E-05	86.59175922
25-Jan	8:00 AM	105	-11.9717	0.000658283	-0.00103867	-0.001698724	0.228222668	1.435793357	0.000140368	82.29718881
25-Jan	9:00 AM	106	-9.67195	0.000681423	-0.00107767	-0.001255486	0.272064882	1.272183282	-9.60E-06	75.42842611
25-Jan	10:00 AM	107	-5.03321	0.000649665	-0.001089984	-0.001096109	0.291343678	1.218864465	-8.69E-05	72.23967956
25-Jan	11:00 AM	108	-22.7207	0.000692883	-0.001051101	-0.000994585	0.30450711	1.225986509	0.000165658	71.17300022
25-Jan	12:00 PM	109	-21.7628	0.000535132	-0.001014475	-0.000923897	0.298757729	1.279472919	0.00032547	67.68852173
25-Jan	1:00 PM	110	-23.1509	0.000385301	-0.000969227	-0.000822631	0.249359329	1.389462385	0.000580837	66.71282132
25-Jan	2:00 PM	111	-23.9646	0.000358026	-0.000950187	-0.000762428	0.2696172	1.41567553	0.00056982	67.03657655
25-Jan	3:00 PM	112	-39.1758	0.000396573	-0.000971715	-0.000731869	0.382214184	1.360192943	0.000681061	68.56801871
25-Jan	4:00 PM	113	-35.3017	0.000446532	-0.000977483	-0.000853423	0.517046814	1.318986985	0.000378034	72.23343985
25-Jan	5:00 PM	114	-40.8738	0.000653051	-0.000981021	-0.001168292	0.503949968	1.311597061	0.000309805	81.57376776
25-Jan	6:00 PM	115	-23.3493	0.000646661	-0.000906058	-0.001437995	0.402624354	1.311095319	5.53E-05	80.11808237
25-Jan	7:00 PM	116	-23.7599	0.000669721	-0.000927518	-0.002173365	0.326867155	1.295462615	9.89E-05	75.63246533
25-Jan	8:00 PM	117	-13.2833	0.000526542	-0.000902991	-0.003361245	0.247127339	1.254076596	0.000110601	66.41879714
25-Jan	9:00 PM	118	6.164446	0.000402908	-0.000864342	-0.003509393	0.221955587	1.099930738	-0.000135851	58.44137233
25-Jan	10:00 PM	119	-5.07001	0.000580938	-0.00090744	1.18E-14	0.214515249	1.06645753	-0.000122375	57.72916364
25-Jan	11:00 PM	120	-8.06227	0.00052743	-0.001117694	1.70E-13	0.323304579	1.060723363	-0.000162911	52.49501908
26-Jan	12:00 AM	121	-14.6155	0.000844784	-0.001257749	-3.72E-14	0.346872582	0.88648646	-0.000338583	51.70104667
26-Jan	1:00 AM	122	-8.09018	0.00077784	-0.001227133	4.43E-14	0.263227328	0.891435639	-0.000283344	49.44241115
26-Jan	2:00 AM	123	-12.6798	0.000747931	-0.00125924	-1.38E-14	0.249355961	0.921204471	-0.000119547	48.6566868
26-Jan	3:00 AM	124	-26.0552	0.000966229	-0.001251099	-4.31E-14	0.334477642	0.874418499	-0.000179381	49.59080496
26-Jan	4:00 AM	125	-20.5799	0.00079647	-0.001210695	-0.165963246	0.319082402	0.893653932	-0.000115422	47.93832516
26-Jan	5:00 AM	126	-15.504	0.000672799	-0.001216474	-0.007259991	0.382005174	0.941114375	-0.000200337	47.49277736
26-Jan	6:00 AM	127	-7.53558	0.000856652	-0.001178859	-0.002698746	0.405965213	0.908693625	-0.000569866	50.40972697
26-Jan	7:00 AM	128	12.17203	0.000656084	-0.001116605	-0.001830765	0.382223992	0.957511891	-0.000719632	53.69113993
26-Jan	8:00 AM	129	17.26269	0.000494851	-0.001196617	-0.00155963	0.391133581	1.027869323	-0.000590238	56.75208854
26-Jan	9:00 AM	130	28.46226	0.000301587	-0.001230924	-0.001538349	0.324314945	1.069288916	-0.000432032	55.77062246
26-Jan	10:00 AM	131	37.06962	0.000256106	-0.001258252	-0.001445315	0.338485005	0.988066946	-0.000530562	52.65220026
26-Jan	11:00 AM	132	43.68469	0.000428457	-0.001350488	-0.001444854	0.195653613	1.040382362	-0.000628777	53.98278787
26-Jan	12:00 PM	133	33.96706	0.000830522	-0.001386203	-0.001363698	0.070014507	0.99424734	-0.000742562	55.08614855
26-Jan	1:00 PM	134	43.38528	0.000644705	-0.001441289	-0.001331594	0.054588406	1.02439916	-0.000710333	51.30618834
26-Jan	2:00 PM	135	30.715	0.000231539	-0.001223579	-0.001082854	0.195480119	1.124097315	-0.000291898	49.97912776
26-Jan	3:00 PM	136	-0.92387	0.000308948	-0.001073906	-0.000860927	0.33023283	1.147903811	1.19E-05	53.41156671
26-Jan	4:00 PM	137	-20.9154	0.00033332	-0.001045794	-0.000817679	0.513285048	1.154104999	0.000144292	57.2482453
26-Jan	5:00 PM	138	-28.9916	0.000356648	-0.001015523	-0.00095033	0.592931128	1.238268783	0.000227324	64.90376691
26-Jan	6:00 PM	139	-1.49367	0.000467874	-0.001093673	-0.001853476	0.492228842	1.171557955	-0.000300043	64.88378758
26-Jan	7:00 PM	140	1.545178	0.000908042	-0.001229555	-0.002971069	0.326577957	1.090227242	-0.000669807	62.87058345
26-Jan	8:00 PM	141	-5.85022	0.000741886	-0.001304634	-0.003100358	0.226361215	1.254631912	-0.00022162	57.79128122
26-Jan	9:00 PM	142	-4.72318	0.000504459	-0.001336144	0.059495274	0.220602553	1.189818477	4.52E-05	52.66329641
26-Jan	10:00 PM	143	-22.4012	0.001126517	-0.001350796	-8.77E-15	0.161706908	1.046039476	-0.000138513	57.16532824
26-Jan	11:00 PM	144	-21.4933	0.001010023	-0.001530174	3.51E-14	0.282572206	1.005968052	-0.000209747	50.41245258

10.3 Python code for single equation approach

```
#This block is to import important libraries to be used later for data manipulation
import pandas as pd
from pandas import DataFrame
import numpy as np

#This block is to define the dataset to python
combineddataset = pd.read_csv(r'C:/Users/ASUS/.spyder-py3/Historic Data.csv')

combineddataset=DataFrame(combineddataset,columns=['SerialNumber','Year','Hours','Price','Demand','Hydropower','Wind','Photovoltaics','Import','Export','Coal Price','Neutral Gas Price Index','NetConnect Germany Gas Market', 'CO2 Certificates','Currency US-EU','Availability'])

#This line is to make python read the values only upto a certain serial number in the excel file
combineddatsetyear=combineddataset[combineddataset.SerialNumber<16825]
useddataset=combineddatsetyear[['SerialNumber','Year','Hours','Price','Demand','Wind','Photovoltaics', 'Coal Price', 'CO2 Certificates','Availability']]

#This line is to create the title row for the output excel file
np_finalsolution=np.array(['intercept','Demand','Wind','Photovoltaics','Coal Price', 'CO2 Certificates','Availability','Price forecast'])

#This block finds the intercept and regression coefficients
from sklearn import linear_model
import statsmodels.api as sm
X = useddataset[['Demand','Wind','Photovoltaics','Coal Price', 'CO2 Certificates','Availability']]
Y = useddataset['Price']
regr = linear_model.LinearRegression()
regr.fit(X, Y)
print('Intercept: \n', regr.intercept_)
print('Coefficients: \n', regr.coef_)

#This block is to predict the price using the predicted exogeneous variables taken from the csv file 'futurevalues_singlemodel'
exovariables=pd.read_csv(r'C:/Users/ASUS/.spyder-py3/futurevalues_singlemodel.csv')
exovariables=DataFrame(exovariables,columns=['Hour','Price','Demand','Wind','Photovoltaics','Coal Price', 'CO2 Certificates','Availability'])
predicteddata=exovariables[['Demand','Wind','Photovoltaics','Coal Price' 'CO2 Certificates','Availability']]
print ('Predicted  Price: \n', regr.predict(predicteddata))

 #This block is to stack the coefficients,intercepts and predicted prices obtained
coefficient=regr.coef_
intercept=regr.intercept_
predictedprice=regr.predict(predicteddata)
print (predictedprice)
np_stackeddata=np.hstack((intercept,coefficient,predictedprice))
print(np_stackeddata)
X = sm.add_constant(X) # adding a constant
model = sm.OLS(Y, X).fit()
predictions = model.predict(X)
```

```
print_model = model.summary()

#This prints the important statistical parameters
print(print_model)
#This block finally exports the solution set back to a csv file with name 'finalfiletargetweek' which is essentially the output
finalsolutioninexcel=pd.DataFrame(np_stackeddata)
finalsolutioninexcel.to_csv('finalfile_targetweek.csv',index=True)
```

10.4 Autoregressive model programming code

Based on (Jason Brownlee, 2017).

```
"The code below will load the dataset as a Pandas Series"
from pandas import Series
from matplotlib import pyplot
series = Series.from_csv('Demand Forescast JANUARY 2019.csv', header=0)" 'Demand Forescast JANUARY 2019.csv' is the excel file which content the historical hourly price values "
"The code below will Plot the data in the excel file 'Demand Forescast JANUARY 2019.csv' "
print(series.head())
series.plot()
pyplot.show()

" The code below will evaluate the correlation of Lagged Past Values"
from pandas import Series
from matplotlib import pyplot
from statsmodels.graphics.tsaplots import plot_acf
series = Series.from_csv('Demand Forescast JANUARY 2019.csv', header=0)
plot_acf(series, lags=31) " 31 is the number is lagged past values to be evaluated; this value can be change as desired "
pyplot.show()

"This section of the code will import the Autoregression Model"
from pandas import Series
from matplotlib import pyplot
from statsmodels.tsa.ar_model import AR
from sklearn.metrics import mean_squared_error
series = Series.from_csv('Demand Forescast JANUARY 2019.csv', header=0)
"This section of the code divide the historical data in a training set and test set "
X = series.values
train, test = X[1:len(X)-168], X[len(X)-168:]
"This section of the will train the Autoregression Model by using the  training set and test "
model = AR(train)
model_fit = model.fit()
print('Lag: %s' % model_fit.k_ar)
print('Coefficients: %s' % model_fit.params)
window = model_fit.k_ar
coef = model_fit.params
"This section of the will compute the prediction"
# walk forward over time steps in test
history = train[len(train)-window:]
history = [history[i] for i in range(len(history))]
predictions = list()
```

```
for t in range(len(test)):
	length = len(history)
	lag = [history[i] for i in range(length-window,length)]
	yhat = coef[0]
	for d in range(window):
		yhat += coef[d+1] * lag[window-d-1]
	obs = test[t]
	predictions.append(yhat)
	history.append(obs)
	print('predicted=%f, expected=%f' % (yhat, obs))
error = mean_squared_error(test, predictions)
print('Test MSE: %.3f' % error)
"This section of the will plot the predicted values"
# plot
pyplot.plot(predictions, color='red')
pyplot.show()
```

10.5 Linear regression. Wind forecast model programming code

```
#This block is to import important libraries to be used later for data manipulation
import pandas as pd
from pandas import DataFrame
import numpy as np

#This line is to import data from the excel file
Winddataset = pd.read_csv(r'C:/Users/ASUS/.spyder-py3/Wind Forecast.csv')

#This line tells python about the columns available in the csv file
Winddataset=DataFrame(Winddataset,columns=['Year','Hours','Wind speed','Wind offshore','Wind onshore'])
Winddataset2018=Winddataset[Winddataset.Year<=2018]

from sklearn import linear_model
np_finalsolution=np.array([‘Intercept’,’coefficient’,’Wind onshore’])
for x in range(1,169):
Winddataset_x=Winddataset2018[Winddataset2018.Hours==x]
Print(Winddataset)
X=Winddataset_x[[‘Wind speed’]]
Y=Winddataset_x[‘Wind onshore’]
Regr=linear_model.LinearRegression()
Regr.fit(X,Y)
print(‘Intercept: \n’, regr.intercept_)
print(‘Coefficients: \n’, regr.coef_)

Windspeedfuture=pd.read_csv(r'C:/Users/ASUS/.spyder-py3/Windspeedfuture.csv’)
Windspeedfuture=DataFrame(Windspeedfuture,colums=[‘Hours’,’Wind speed’])

#This block is to predict the price using the predicted exogenous variables taken form the csv file’futurevalues’
Windspeedfuture_x=Windspeedfuture[Windspeedfuture.Hours==x]
Predictedonshoreproduction=Windspeedfuture_x[[‘Wind speed’]]
Print (‘Predicted onshoreproduction: \n’, regr.predict(predictedonshoreproduction))

#This block is to stack the coefficients, intercepts and predicted prices obtained for each hour into one single array of 168 rows
```

```
coefficient=regr.coef_
intercept=regr.intercept_
predictedonshoreproduction=regr.predict(predictedonshoreproduction)
print(predictedonshoreproduction)
```

11. References

CLEW, 2017. *Clean Energy Wire CLEW,* Berlin: https://www.cleanenergywire.org/factsheets/germanys-energy-consumption-and-power-mix-charts.

CLEW, 2018. *Clean Energy Wire CLEW,* Berlin: https://www.cleanenergywire.org/factsheets/germanys-energy-consumption-and-power-mix-charts.

Duffner, 2012. *Forecasting German Day-Ahead Electricity Prices using multivariate Time Series Models.* Available at: https://brage.bibsys.no/xmlui/bitstream/handle/11250/169717/Duffner 2012.pdf?se ed. s.l.:Norges Handelshøyskole – Norwegian School of Economics.

Example of Multiple Linear Regression in Python (2019) *Datatofish.com*. Available at: https://datatofish.com/multiple-linear-regression-python/ (Accessed: 17 February 2019).

EY, 2013. *Mapping power and utilities regulation in Europe,* s.l.: https://www.ey.com/Publication/vwLUAssets/Mapping_power_and_utilities_regulation_in_Europe/$FILE/Mapping_power_and_utilities_regulation_in_Europe_DX0181.pdf.

Jason Brownlee, 2017. *Autoregression Models for Time Series Forecasting With Python.* Retrieved January 16, 2019, from https://machinelearningmastery.com/autoregression-models-time-series-forecasting-python/ ed. s.l.:s.n.

Weron, 2006. *Modeling and Forecasting Electricity Loads and Prices Rafal⁄Weron. John Wiley & Sons, Inc..* s.l.:s.n.