Forecasting German Day-Ahead Electricity Prices using multivariate Time Series Models

Author: Stephan Duffner

Thesis Advisor: Professor Øivind Anti Nilsen

Master Thesis

Bergen, Spring 2012

Profile: Energy, Natural Resources and the Environment

Norges Handelshøyskole – Norwegian School of Economics

This thesis was written as a part of the master program at NHH. Neither the institution, the supervisor, nor the censors are - through the approval of this thesis - responsible for neither the theories and methods used, nor results and conclusions drawn in this work.

Abstract

Using a newly available dataset about the unavailability of power plants and the in-feed of renewable energies to forecast day-ahead electricity prices at the German Power Exchange, this work shows that the predictive power increases considerably when including exogenous variables. While a similar univariate approach based on the year 2001 yielded a Mean Absolute Percentage Error of 13.2%, the use of the presented variables improved the forecasting error to 8.3%. Other findings of this work include that a model based on 24 individual time series produces smaller forecasting errors than one time series which includes all consecutive hours, that the selection of the in-sample and out-of-sample periods varies greatly between different works and that the use of OLS seems to be underestimated in the existing forecasting literature for electricity prices.

Table of Contents

Abbreviations

AR	Autoregressive Model
ARCH	Autoregressive Conditional Heteroscedasticity Model
ARIMA	Autoregressive Integrated Moving Average Model
ARIMAX	Multivariate Autoregressive Integrated Moving Average Model
ARMA	Autoregressive Moving Average Model
EEX	European Energy Exchange, Leipzig
GARCH	Generalized Autoregressive Conditional Heteroscedasticity Model
MA	Moving Average Model
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MGARCH	Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model
MWh	Megawatt Hour
OLS	Ordinary Least Squares
RMSE	Root Mean Square Error

Figures

Tables

1 Introduction

As electricity markets become deregulated, the numbers of market participants at the power exchanges is increasing. To place reasonable bids, the participants have to build an own opinion about the future development of electricity prices at the spot market. There are several important factors which contribute to the settlement of electricity prices, for example the forecasted power consumption, the feed-in of renewable energies as requested by law, the price of input commodities like oil or emission certificates – and the unavailability of power plants. This thesis aims to work out the determinants of the electricity price and then use them to forecast electricity prices for the day-ahead electricity market. That way, the thesis also delivers an understanding for the importance of different variables for the electricity price. The data that will be used for this has to be published by the major utility companies due to an order of the regulation authority since about mid-2009. Since the data in question is rather new, this thesis is among the first scientific works making use of it in an econometric context.

The precise forecasting of electricity prices is of high importance for the market participants: First, market participants that own power plants have to adjust their bids to optimize the profit from their power plants. Second, market participants that have to buy electricity capacities need to decide whether on forward markets or at the spot market. Third, market participants are able to schedule the load of their power plants depending on the electricity prices that can be expected.

Following the modelling approach established by Box and Jenkins, the thesis develops a number of Time Series models, including an ARIMA, ARIMAX, MGARCH and an OLS model. The models will be estimated using two sub-samples of the available dataset. The explanatory power of the different models will then be discussed upon their prediction for a respective off-sample subset of the dataset.

Chapter 2 will introduce the specialities of electricity markets in general and show why the German electricity market is very relevant. The important determinants for the electricity price will be worked out. In Chapter 3, the availability of these variables will be checked and basic properties of the data will be explained. Chapter 4 starts with an overview of the available methodologies to model electricity prices and explain the chosen econometric models. Chapter 5 then presents the development of the models and the obtained results of the analysis. Chapter 6 concludes.

2 Determinants of the Electricity Price

2.1 Introduction on Energy Trading, European Power Exchanges and Market Participants

As the aim of this work is to increase the understanding of electricity prices, this chapter will review the theoretic concepts of power markets in general and the German power market in particular and draw conclusions on the variables which are necessary to model electricity prices with econometric methods.

Today, electricity is a traded commodity but is different from other commodities like oil and gas in a number of aspects. The generation and consumption of electricity have to be balanced at all times to have a constant frequency in the grid. A continued imbalance of generation and consumption and the subsequent deviation from the grid's target-frequency of 50Hz would end with the consequence of malfunction of electrical machines and blackouts. In addition, electricity can only be stored by means of converting it to another form of energy which comes at the cost of efficiency losses and in any case, these storage options are very limited. Because only one grid is economically feasible for a society, electricity transmission is a natural monopoly and needs to be controlled by regulatory authorities in order to enable fair market mechanisms.

During the process of liberalisation, power trading activities across Europe have risen considerably. Within Europe, Germany is the largest economy and power market in terms of electricity consumption. Germany's annual power consumption 2010 amounted to about 590 TWh, with France taking the second place using about 510 TWh of electricity (RTE 2011; BMWi 2012). The four largest electricity producers RWE, E.ON, Vattenfall and EnBW hold a generation capacity of about 80% of the German market according to the federal competition authority (Bundeskartellamt 2011). The high voltage grids are also operated by only four transmission system operators (TSOs). The relevant power exchange for the spot market is the "EPEX Spot" which covers the markets Germany, France, Switzerland and Austria and is connected to the Belgian, Dutch and the Nordic market via market coupling mechanisms. The borders to Poland and the Czech Republic have explicit auctions (Tarjei 2011). The EPEX Spot has 211 members (EPEX SPOT 2011b), including the major power utilities of central Europe, transmission system operators, local energy companies and municipalities as well as pure energy trading companies and banks. Small companies which do not have direct access to the EPEX/EEX trading system can trade via separate accounts of other trading members. One main group of market participants are generators and retailers with intrinsic physical long or short positions, i.e. they have a certain customer base and a certain generation capacity and need to trade the difference on the power exchange. A second main group of market participants are pure traders and banks who typically aim to exploit prices differences to gain profit from arbitraging and take speculative positions.

Being the biggest power market in Europe, the large influence of renewable energies along with still large shares of conventional energy production and a diverse structure of market participants makes the German electricity market a very relevant one for an econometric analysis of power exchanges.

2.2 Power Generation and Demand Characteristics in Germany

In this section, important determinants for the power price will be worked out for both the supply and the demand side of the market. Germany has a number of different technologies in use to generate electricity, each with different characteristics and dependencies towards the electricity price. Because of the balancing needs that have been described earlier, production is characterised by a mixture of heterogeneous types of power plants that have varying costs structures reflecting the need for flexibility. Base-load power plants usually operate for most of the time of the year and are characterised by high fixed costs and low marginal costs, while peak-load power plants are only used as needed and have comparatively low fixed costs but typically high marginal costs (Ockenfels et al. 2008).

Before going into the details of production, it is important to understand the concept of marginal costs in the context of electricity production. The marginal costs of electricity production include mainly the fuel costs and other variable costs of production. In addition to this, the marginal costs consist of the opportunity costs that arise if the production resources are not used in the manner with the highest possible value. As Ockenfels explains, occasionally the marginal costs cannot be defined clearly. For example, this can be the case if there are so-called “complementarities” or “non-convexities”, which are caused by start-up costs. Start-up costs are incurred upon every re-start of a power plant. These include costs for heating-up the power plant, network synchronisation and the increased wear and maintenance costs due to the temperature fluctuations (Ockenfels et al.

2008). These cannot always be calculated precisely and thus marginal costs cannot be defined clearly.

In the following, the different types of power plants in the German market will be presented as their properties are in turn an important determinant for the characteristics of the market. The most important power plants are those using rivers, wind and solar power, nuclear power, lignite and hard coal, gas and oil and finally pump storage power plants. When ordering power ascending order in terms of marginal costs, run-of-the-river power plants come first. These power plants are installed in large rivers like the Rhine River and make use of the constant water flow. They have very limited possibilities to store water and usually are operational around the clock as a typical base-load power plant. They have no fuel costs and little maintenance costs as they do not need constant supervision. These power plants can become unavailable in the event of low river levels or maintenance.

Second, there are wind and solar power plants. Due to little marginal costs, these power plants will also run whenever possible but as opposed to run-of-the-river power plants they have a much bigger variability in their power generation due to fluctuating wind and cloud coverage. The electricity companies employ meteorologists to predict wind speeds and insolation and therefore the generated power. The German law requires the power suppliers to feed the electricity generated by solar and wind power plants into the grid and they can be disconnected only in the case of emergencies for grid control. Therefore, wind and solar energy cannot be put into either the base- or peak-load category. Germany’s wind and solar capacities have grown considerably in the last years due to high feed-in tariffs, with jumps in capacity taking place prior to changes in these tariffs (Tarjei 2011).

Third, nuclear power plants generate electricity by the fission of radioactive molecules. Due to the high energy density of uranium they have little marginal costs once they are running, but it is considerably expensive and time-consuming to start and stop a nuclear power plant. However, it is possible to moderate the nuclear reaction by using the control rods and thereby control the power output to some extent. These characteristics make a nuclear power plant a typical base-load power plant. Unavailabilities can occur due to scheduled maintenances, which take about one month every year, low river-levels in summer and unscheduled shut downs out of security or political considerations.

Fourth, lignite and hard coal power plants facilitate coal combustion to generate electricity. Through the oxidation of coal, CO_2 gets produced which by itself is a natural climate gas but contributes to the human-induced global warming due to the quantities humans exhaust of it currently. So called CO_2-certificates are needed when one wants to run a coal power plant in order to make its use less attractive and motivate reduction of CO_2. Coal power plants have less start-up and stop-costs than nuclear power plants but also need a couple of hours to reach full capacity. Modern coal power plants need less time than old ones as they are already designed for greater controllability needs. Also, coal power plants can be put into a standby-mode from which it can produce electricity on shorter notice than from a cold-start. Therefore, coal power plants have both base-load and to some extend peak-load usability. Unavailabilities can occur due to maintenance and also due to low river levels: Given low river levels, cooling might not be possible and ships might not be able to use the rivers for the transport of coal.

Fifth, gas power plants are considered very efficient power plants as they have a high efficiency-factor. They exhaust less CO_2 than coal power plants for a given amount of electricity but are considered more expensive than coal as per electricity produced, though prices have declined in the last years due to new discoveries in North America. They can start very quickly and therefore have good peak-load capabilities. Because gas is transported through pipelines, unavailabilities occur mostly due to maintenance. Transport interruptions of gas due to e.g. political decisions from Russia so far have not been an issue yet in Germany as there are plenty storage capabilities for gas from both tanks and the grid: Contrary to the electricity grid, the gas grid can be a storage in itself as one can increase the pressure of the gas. The gas price is often linked to the oil price in the long-term contracts between gas producers and the retailers, which was originally argued with the possibility of substitution between the two (Stern 2007). Oil power plants are similar in terms of usage to gas power plants but are only used in rare cases for peak-load purposes as oil is an expensive energy carrier and there are considerable CO2-emissions. Unavailabilities can occur for maintenances.

Sixth, there are pump storage power plants which do have a free energy source but that can come with considerable marginal costs: due to their superior possibilities of both producing electricity on very short notice within seconds and the possibility to pump water up into the basin, they can be sold as “Primary Reserve” in a separate market, the market for ancillary services, which is needed from TSOs for very short-term generation capabilities. Pump storage power plants are peak-load power plants and have little unavailabilities for maintenance. Other generation facilities like biomass or geothermal energy do not play a significant role yet. With their unique ability to "store" generated electricity for later use, they have the ability to smoothen prices: Electricity is used to pump up water to the top basin during low prices and then used to generate power again when prices are high. The smoothing price effect of water power can be observed in Nord Pool's electricity prices, which are less volatile than prices on the EPEX due to high capacities of water power plants. The total composition of energy sources for 2011 in Figure 1 shows, that energy production is still dominated by fossil fuels with lignite having a share of 25% of total production and coal and gas having 19% and 14% respectively.

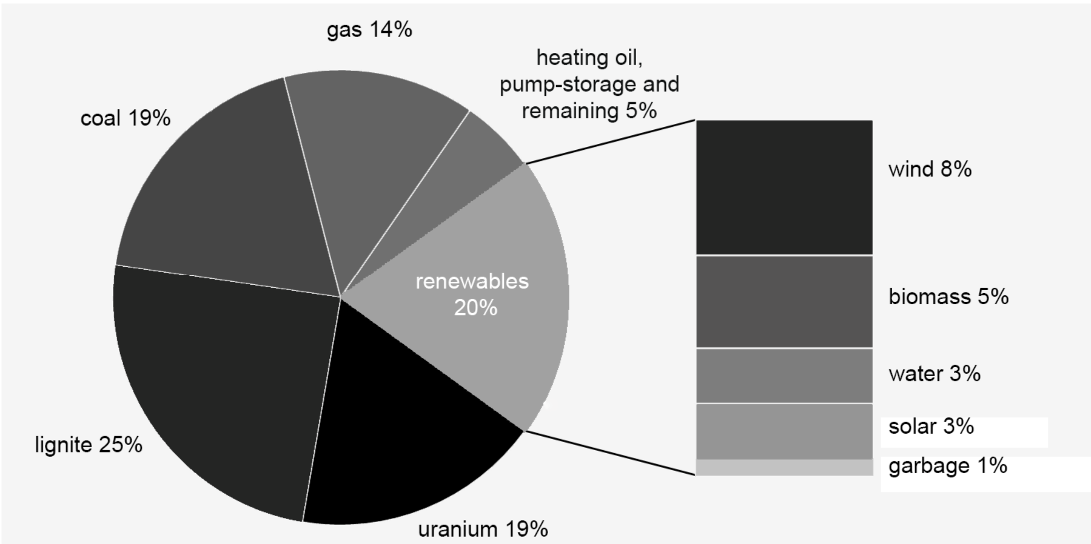

Figure 1: composition of German Power Generation 2011, adapted from: BDEW (2012)

When shifting the focus from the supply side to the demand side of the electricity market, Bourbonnais and Méritet (2008) work out several factors on why electricity demand has characteristics that are different from most other commodities. Electricity demand is highly inelastic as it is a necessary product with very limited substitutes. In Addition, the demand is highly dependent on unforeseeable factors like climate and weather conditions. Also, electricity displays seasonal patterns due to economic activity and weather conditions. Seasonality can occur on various levels, including an hourly, daily, weekly, or monthly seasonality (Bourbonnais & Méritet 2008). Differences in electricity demand between countries can subsequently occur due to variations in climate and weather and also in the composition of electricity buyers, namely how much electricity is needed from households and from different kinds of industry. In Germany, temperature is less important for total demand compared to other countries due to the relatively large dependence on industrial activity of around 45% of total demand, the relatively little dependence on electricity for heating and few necessities for air conditioning (Tarjei 2011).

As a last important determinant of electricity prices, there are to mention congestion issues that can occur at national or regional borders. Congestion means the limited grid capacities between two adjacent networks. The more congestion issues there are, the more important issues about market power within a region will be and the electricity price will be higher. The less important congestion issues become, the lower the electricity price will be in areas that originally had a high price (Lise et al. 2008).

Through this analysis of the generation characteristics of the German electricity market, important variables for the price settlement could be determined: The composition of different base- and peak-load power plants, weather forecasts, the demand forecast, the amount of water power plants, interconnection capacities, planned and unplanned maintenance, the in-feed of renewable energies and the price of input commodities. These variables will be checked for availability and usability in the econometric analysis in Chapter 3.

2.3 Traded Products and Relevant Markets

For the purpose of this work, it is important to select the appropriate market and products for the econometric analysis. The two main marketplaces for day-ahead trading in Germany are represented by the power exchange EPEX Spot and electronic OTC trading.

Due to its liquidity and number of market participants, the EPEX Spot is the central trading point of the German day-ahead power market. Currently, the daily auction for the next day takes place at 12.00 pm, on each day of the weak including statutory holidays (EPEX SPOT 2011a). Liquidity on the Intraday Market, which covers the period after the day-ahead auction and the actual delivery period, is only a small fraction of the day-ahead auction and is only used for minor balancing purposes. Real time imbalances in the power system are balanced using generation units that can provide positive or negative primary, secondary and tertiary reserve energy under supervision of the TSOs. TSOs procure these types of reserve energy on separate markets (Johannes 2011).

Contrary to exchange-based trading, OTC trading takes places directly between the counterparties and is often facilitated by broker companies. The transactions are either executed via electronic broker platforms or bilaterally via telephone. According to

Johannes, most day-ahead trading activities take place between 8 a.m. and 12 p.m. on the day prior to the delivery day. Johannes points out that the continuous OTC market is important for market players to hedge larger volumes prior to the exchanged based auction at 12 p.m. Thus, the OTC-market can be considered to be the last forward market before the final EPEX Spot exchange clears (Johannes 2011). According to Tarjei, most of the trading volume of German power is in the OTC market. Similar to the trading on the exchange, spot contracts require physical delivery while the futures market can be physical or financial (2011). However, even though a large volume of the power trades are made via OTC and thereby independently from the systems of the power exchange, the price settlement through the power exchange will serve as a reference point, from which continued price data is available in a standardized form. This is why an econometric analysis should be based on EEX data and not on OTC data.

There is a range of spot price products with different hourly combinations that can be traded during the daily auction. Out of arbitrage considerations, the price of a product which includes a set of hours, e.g. a Base, Peak or Off-Peak contract, has to be equal to the sum of the individual hours. If this would not hold true, there would be riskless arbitrage opportunities for the market participants by shorting e.g. a high-priced product which has a combination of contracts, and closing the position again by buying the low-priced set of contracts. This “value additivity” not only holds true for the spot market, but also for the Futures market when one also considers the time value of money (Bjerksund et al. 2010). This means that by focusing on the individual hours for forecasting, the same conclusions can be drawn on the price of related products.

Besides the day-ahead and intraday markets which are linked to the physical delivery of electricity, there are derivatives markets which are purely financial and in which contracts on future deliveries are traded. These consist of futures contracts for weekly, monthly or yearly delivery usually up to three years in advance. Besides regular futures contracts, there is a wide range of other derivatives like e.g. European, American and Asian Options which are traded either on the EEX or OTC. Many energy suppliers use the derivatives market for hedging purposes and close open positions so as to limit risks and secure a certain profit margin – giving up possible higher prices in return. The percentage of power that is already hedged in advance is determined through the individual hedging, where conservative strategies involve hedging up to 100%.

As Ockenfels et al. (2008) explain, even though a large part of the energy is traded in long-term contracts and only a comparatively small part is traded day-ahead in the spot market auctions, it is sensible to concentrate the econometric analysis on the spot market. This is due to the fact that the “prices in all upstream electricity markets actually reflect the expected spot market price” and hence it is the spot price that determines “the costs of electricity even in the long run”. This becomes even more compelling when considering the special conditions the spot price is subject to, as it is linked to the physical aspects of electricity while the derivatives are not linked to physical constraints considering their purely financial nature. Because of the lack of storability of electricity, power exchanges require comparatively complex rules and regulations along with a careful consideration of numerous ancillary technical conditions in the generation and transmission of electricity. Ockenfels et al. (2008) point out that the spot market auction complies with these demanding requirements.

2.4 Auction & Price mechanisms

In the auction, both bid prices for an individual hour and block bids comprising several contiguous hours can be submitted. The maximum admissible bid price has to be between -3,000 EUR/MWh and 3,000 EUR/MWh for all contracts. This wide range is used as the power exchange does not want to constraint price formation. Allowing negative prices is due to the possibility of negative marginal costs for some power plants in times of low demand. For instance, in a time of low demand like a Sunday, most power is generated by base-load power plants that run 24/7. For a limited time and in special cases, it might be cheaper for the owner of a nuclear power plant to increase power consumption by offering money to a consumer, rather than to shut down the nuclear power plant and lose all the profits for remaining time until it becomes operational again. In this case, the owner of the nuclear power plant is willing to pay a price to someone who can consume the energy. Negative Prices have been observed on various occasions in the past. In the dataset that will be used later on, 48 of 16,776 hours had prices below 0 EUR/MWh. The market participants making use of this opportunity will most likely be the owners of pump storage power plants, who will use the abundant power to pump up water into their storage basin.

The bids must be sent to EPEX Spot before 12PM on the day before delivery. Then, all bids are aggregated into supply and demand functions and converted into linearly interpolated sell or buy curves. The market price is established on the basis of the intersection of these supply and demand functions and thereby a market clearing price for every hour of the following day is generated. Every market participant who supplies electricity during a given hour receives the respective price for that hour and every market participant who buys electricity during that hour pays that price. Since all participants have the same price, this mechanism is also referred to as the “uniform price auction”. In case the transmission capacity is not sufficient for the execution of the schedules determined in the auction, the market can be divided into price zones. However, this case has never occurred so far as the transmission system capacities within the trading area of EEX are currently sufficient compared to the quantities traded (Ockenfels et al. 2008).

In theory and in practice, the resulting price is the marginal cost of the most expensive power plant from the group of least expensive power plants that are sufficient to cover the power demand. Figure 2 shows this “merit order”: power plants with the least marginal costs will be offered to the market at first because they will yield the highest profit. The rank of different technologies in the merit order can change as fuel prices change, i.e., gas and hard coal power plants may switch their respective ranks in the merit order when fuel prices change (Tarjei 2011).

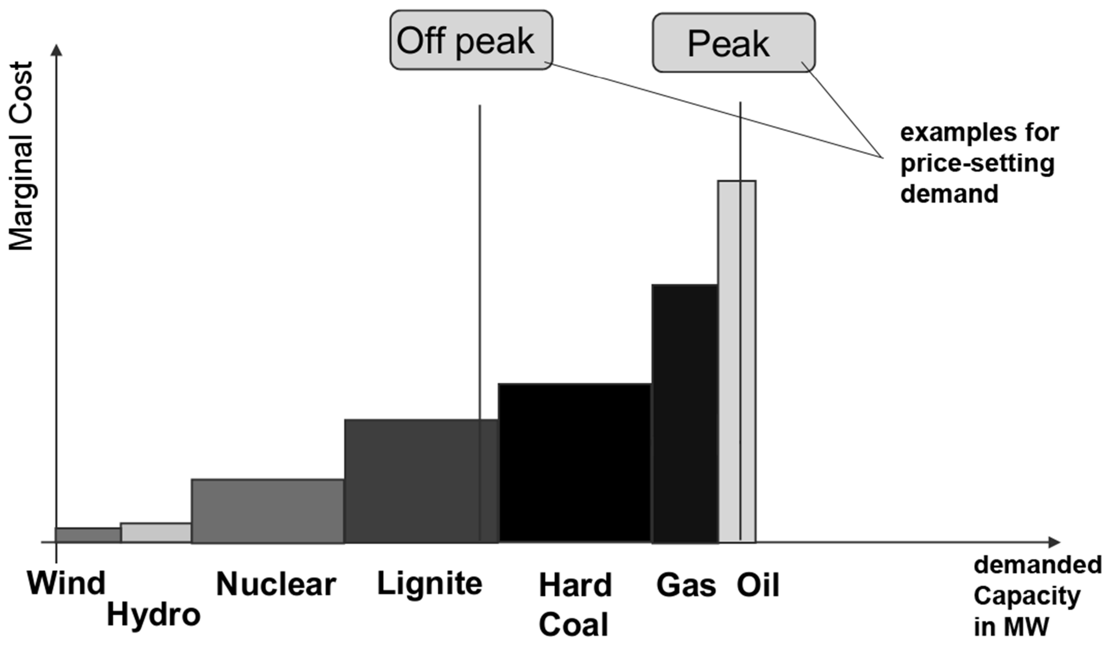

Figure 2: Merit Order of Power Plants, adapted from Skrivarhaug (2010)

The merit order principle in theory allows for the possibility of exercising market power by withholding generation capacities, which is an issue widely discussed in public. By withholding a power plant, a market player with many other power plants can increase the profits of all other power plants on the market as the new settlement price increases

(Ockenfels et al. 2008). That way, another aspect in modelling electricity prices can be the concentration of market power among certain market participants. Market power can be measured in numerical terms, e.g. through the use of the Lerner Index or the Concentration Index (Möst & Genoese 2009). Other works use a comparison of the marginal costs and the electricity price (Müsgens 2004). However, to be able to determine the influence of market power in an econometric setting, the time horizon of the dataset has to be sufficiently large to cover different magnitudes of market power. Such an analysis could e.g. be done by using data from the years of the regulated electricity market in which there were monopolies and ranging to the times of the deregulated energy market which supposedly will exhibit more competition. A cross-sectional analysis within a given point in time over different industries will not yield sufficient results for market power, as different industries exhibit idiosyncrasies so that comparative statistics will not reveal information about market power (Vassilopoulos 2003).

2.5 Role of Forecasts

As Weron and Misiorek elaborate, extreme price volatilities have forced the market participants of the electricity market to not only hedge against volume risks but also against price risks in the electricity market. Thus, opinions about future price movements formed through forecasting have become a crucial input in decision making and strategy development. This accelerated research in modelling and forecasting electricity prices with differences in the used methodologies and the used horizon. Weron & Misiorek distinguish between short-term, medium-term and long-term price forecasting (Rafał Weron & Misiorek 2006). The objectives of the three categories differ. While long-term forecasting is used for investment profitability analysis and planning, like determining future sites or fuel sources of power plants, medium-term or monthly time horizons are used for balance sheet calculations, risk management and derivatives pricing. Short term forecasts are e.g. used by a company that adjusts its production schedule depending on the forecasted hourly pool prices and its production costs and thereby maximizes profits. Accordingly, for spot markets the short term forecasts are of main importance (R. C. Garcia et al. 2005; Rafał Weron & Misiorek 2006). Every major market participants who takes part in the auction of the electricity price in one way or the other will need to form an opinion about the future development of the prices so as to be able to determine a reasonable bidding behaviour. Statkraft for example heavily bases its decisions for "Energy Management", trading and hedging on the findings from the analysis and forecasting unit (Skrivarhaug 2010).

3 Data

In chapter two, the possible range of variables that have an influence on electricity prices have been worked out. Now, it is necessary to check these variables for availability for the public and to discuss their usability for econometric purposes. It will only be tried to obtain publicly available data. That way, it can be simulated what a 3rd party or a possible market entrant is capable of. Apparently, an existing electricity supplier will have superior data concerning its customer base than what is publicly available and thereby will be able to make more precise calculations on the electricity price.

3.1 Available Data

3.1.1 EEX: Spot Prices and CO2-Certificates

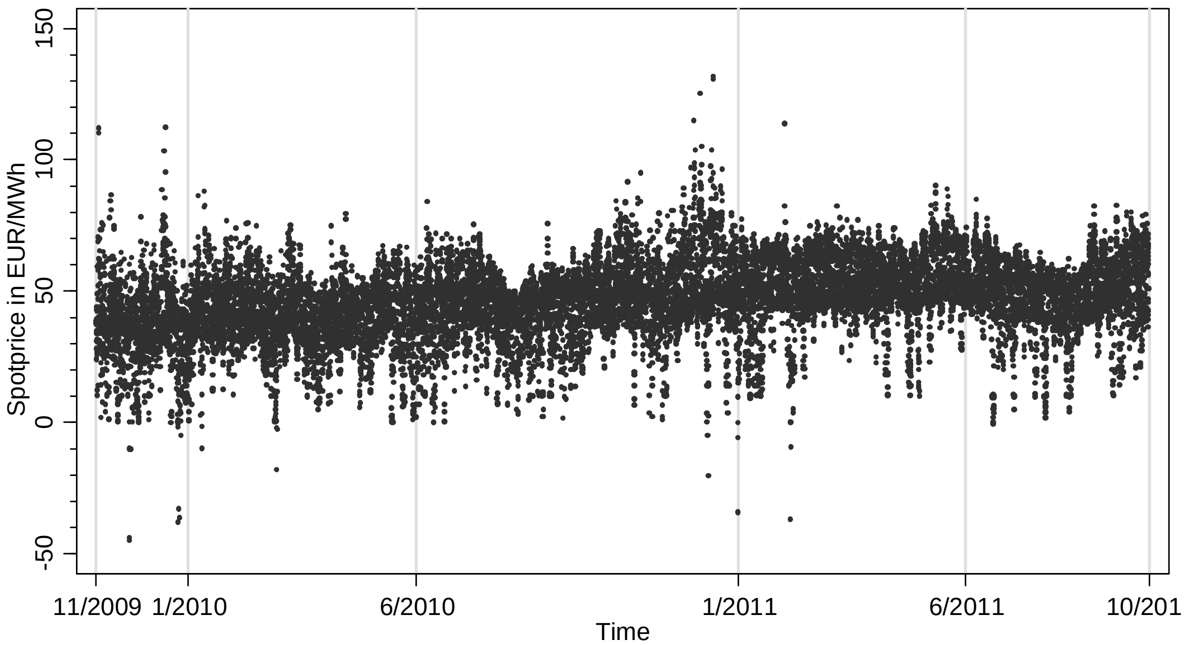

Figure 3: Spot Price EEX in the available dataset, 11 values below -50 EUR/MWh omitted

The spot prices are determined through the auctioning as described earlier and are available from EPEX on a per-hour basis, quoted in EUR/MWh. The data is available since 2002, which was the start of the EEX after the fusion of the power exchange Frankfurt and Leipzig. Since other relevant data is only available since November 2009, this work will start using data on spot prices from this date on. The available data reaches until the 30th September 2011 and accordingly includes nearly two years. This sample size is similar to earlier works as can be seen e.g. in the compilation of Aggarwal (2009). In the sample, the price ranges between -199.99 EUR and 131.79 EUR and has a mean of 46.48 EUR. Peak-prices average at EUR 52.79 and off-peak prices at 40.16 EUR. Table 1 reveals considerable differences between the hours, especially the high values for the kurtosis in the morning hours will lead to interesting results in the forecasting performance.

Hour	Mean	Std. Dev.	Skewness	Kurtosis
1	39.1	10.9	-4.9	67.3
2	35.4	12.0	-3.7	42.0
3	32.1	13.4	-3.1	28.7
4	29.5	15.1	-4.2	44.6
5	30.3	13.0	-2.6	27.4
6	34.2	12.8	-3.1	32.0
7	40.2	17.3	-4.6	56.3
8	48.6	19.1	-3.7	43.6
9	52.1	16.4	-2.1	20.0
10	53.8	13.1	-0.9	6.5
11	54.7	11.7	-0.5	4.1
12	56.1	11.3	-0.3	3.8
13	54.3	10.5	-0.4	3.7
14	51.5	11.5	-0.6	3.9
15	48.2	12.1	-0.6	4.2
16	47.5	12.0	-0.6	4.4
17	47.9	11.8	-0.4	5.0
18	53.4	13.7	1.0	7.8
19	56.9	12.2	0.8	6.2
20	55.9	10.7	0.3	3.2
21	52.5	9.7	0.1	2.9
22	48.6	8.3	0.1	3.2
23	48.6	7.4	-0.1	3.4
24	42.8	8.6	-2.4	21.3
all hours	**46.48**	**15.2**	**-1.5**	**19.5**

Table 1: Summary Statistics for the 24 Spot Price Hours

The EEX also determines the price for CO2-certificates, which is done on each weekday since 2005. In the available dataset, the price has a mean of 14.24 EUR/t. For the use of this work, the price of Fridays has been assumed for Saturday and Sunday as CO2-certificates are not quoted on the weekend.

Electricity prices exhibit a phenomenon which is called spikes or more general: outliers. In some rare events, e.g. when cross-border capacity is remarkably low due to maintenance, there is extraordinarily high wind in-feed or there are very low river levels, exceptionally high or low prices might occur. Some works filter the data by removing outliers, thereby reaching a smaller forecasting error. Since observations cannot be simply removed in a time series dataset, outliers can be filtered by capping the values at a certain threshold or using an average value instead of the outlier. Other works specifically focus on forecasting these exceptional events, like Christensen et al (2011) and Trueck (2007). In order to stay comparable with earlier works, ensure reproducibility and make the results more realistic, outliers have not been removed in the current analysis.

3.1.2 Transparency: Availability of Generation Capacity, Wind- and Solar Feed-in

Due to European and national regulations, large power suppliers have to publish “market-relevant” information considering generation and consumption of electricity since 2009. The EEX publishes this data as a service for these companies on a central website after checking the data for plausibility, anonymizing it to some extent and aggregating it. Not all the owners of power plants have to publish this data, but about 91% of all generation capacity was available on this website at the time the data was downloaded (EEX 2011). In addition to being required by law to publish certain data, some power suppliers publish more data on a voluntary basis. The published data on the transparency website contains: the planned and unplanned unavailabilities of power plants, the planned and actual in-feed of solar/wind energy along with the planned and actual generation of conventional power plants.

The planned unavailability of power plants can be e.g. scheduled maintenances, which are known up to several years in advance. Periods of unplanned unavailability can be due to emergency situations or low river levels that forces thermal power plants to shut down because of environmental concerns and regulations. An unavailability is stated as a time frame, e.g. the data set contains the information that 144 MW of a coal power plant within Germany is being unavailable between 4.11.09 18:00 until 9.11.09 5:00. To use the data in this work, the amount of unavailable power for the various types has been calculated for each individual hour of the dataset. In cases where unavailabilities are stated to start at some point within an hour, e.g. 18:25, the unavailability will be counted in the dataset from the next full hour, i.e. 19:00. As has been discussed, it will take some time for a power plant to start and stop operation in practice so the decline of power being generated will be rather smooth than sudden. As an example, the non-scheduled unavailabilities are shown for nuclear power for the whole data-sample in Figure 4, which includes the decision to shut down nuclear power plants in the aftermath of the Fukushima-catastrophe.

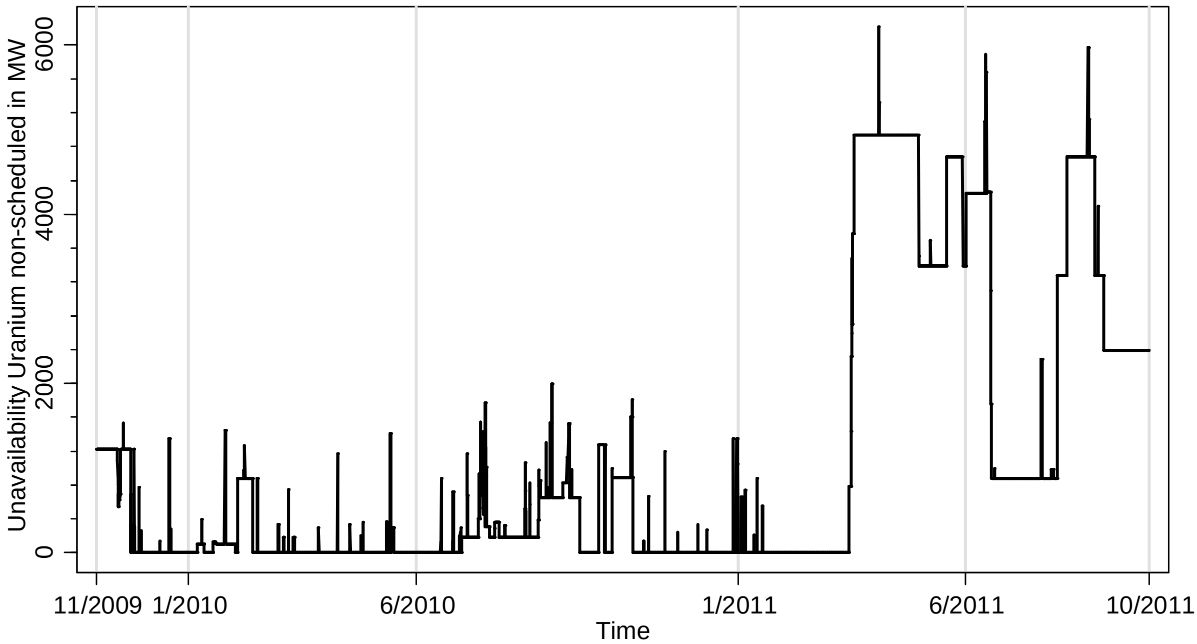

Figure 4: Non-scheduled unavailability for nuclear power plants in MW

The solar and wind in-feed is stated as induced power per each 15 minutes. To use the data in this work, the 4 quarters have been averaged to calculate the amount of MWh that is generated within one hour. The data for solar in-feed is not available until the 19th of July 2010 because the data has not been published on the transparency platform until that date and is noted as zero in the current dataset until that point in time.

As has been worked out in the preceding chapter, the expected demand in a given hour is an important variable in determining the electricity price. The total demand itself is not published on the transparency website. However, generation has to follow demand in an electricity grid and for that all the necessary information is given on the transparency website: There is the information about the total power generated from conventional power plants and the power generated from wind and solar is known as well. Not known are transmission and distribution losses, the power generated by smaller power producers that are not obliged to publish data on the transparency website, and small, decentralised power generation like industrial autogeneration, geothermal or block heat and power plants (Burger et al. 2007). However, this shouldn’t have major consequences for the price formation on the power exchange within the framework of a statistic model as it can be expected that the remaining demanded capacity should follow the same trends which the available data exhibits. Therefore, there are sufficient replacements for total power demand in the transparency data.

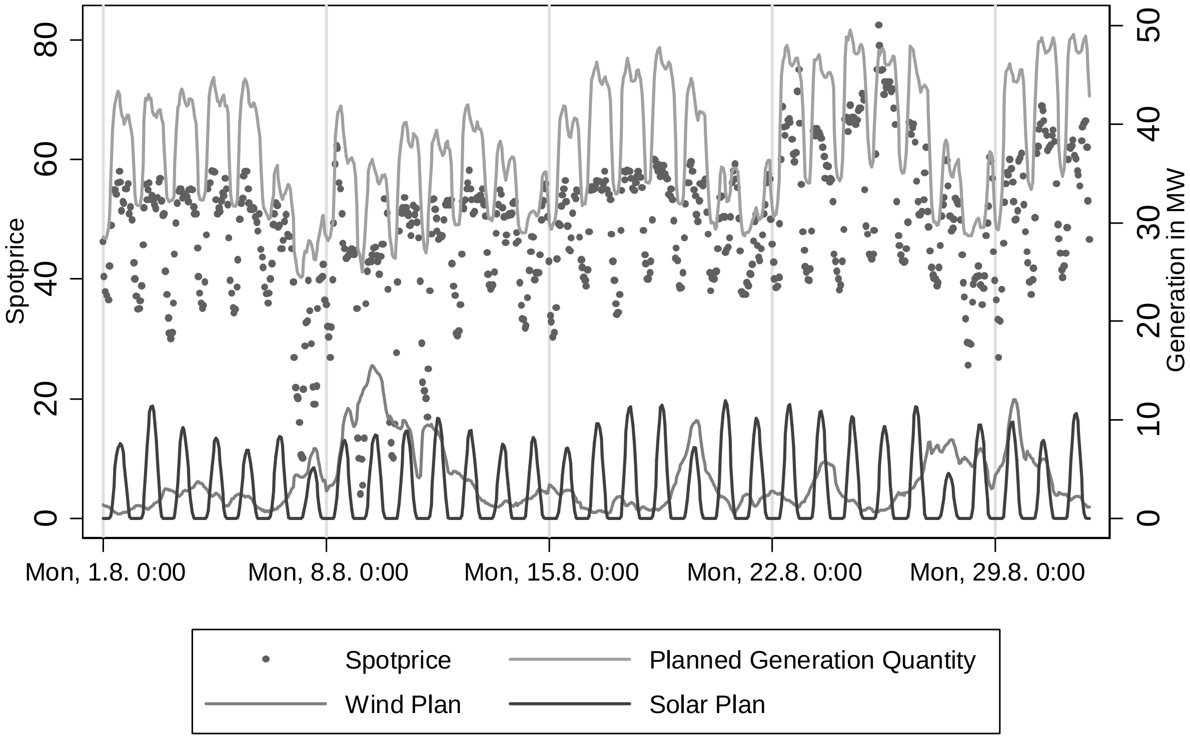

Figure 5: Comparison of Spot Prices, load and Feed-in of renewables

In total, the data published on the transparency website adds significant information that can be used for econometric analyses that has not been available for earlier works. Actually, many earlier works recommend to re-estimate their findings with exactly the data that has now been published on the transparency website, e.g. Swider & Weber (2007) and Garcia et al. (2005). Figure 5 shows an example of the explanatory power of this data: Periods of high wind are accompanied with drops in the generation of both conventional power and the spot price.

3.1.3 Remainder: Weather, Oil-Price, Transmission Capacity and River Levels

As weather plays an important role in energy consumption, temperature is included in the dataset (e.g. Huurman et al. 2010). The data has been obtained from DWD, Deutscher Wetterdienst. Because the purpose of this work is to forecast day-ahead electricity prices, forecasts should have been obtained: However, due to availability reasons, only actual temperature data has been incorporated. That way, the implicit assumption is made that on average, the forecasted temperatures for the next day are exact. Weron & Misiorek (2008) use the same assumption when they calculate an arithmetic average of big cities within the examined market to have a proxy for the average air temperature of the whole region. However, as only daily data for Frankfurt am Main was available for the use of this work, it has become the weather-data that is used as a proxy for the general temperature in Germany. Still, this produces significant estimates which will be shown later because Frankfurt is a rather central city.

In order to keep the number of variables that are used for the econometric analysis in a reasonable size, the oil price is used as a common proxy for the price of the input energy carriers like coal, gas, oil, and uranium. This seems reasonable due to the connection of oil and gas price through long-term contracts. For coal, there are many different prices depending on quality and origin and therefore there is no “one” price that can simply be added. According to Tarjei (2011), “Brent” is the relevant oil price for Germany as it is the crude oil blend from the North Sea. The oil price data is obtained from Thomson Reuters Datastream in USD per Barrel. Since the oil price is only listed for weekdays, the oil-price from Friday has been assumed to stay constant over the following weekend.

Another important issue for the determination of electricity prices can be congestion issues within the market or at its connections to other markets. Considering the auction of the electricity price, the internal transmission capacity has not yet lead to differences in pricing, even though the limited transmission capacity from Northern to Southern Germany has already led to various challenges for the TSOs. Transmission constraints are built into the auctioning system of EPEX Spot and there is the theoretic option of zonal prices but this has not yet occurred in the auction for the German electricity market. Accordingly, transmission constraints are not an issue right now for a statistical approach considering the forecasting of electricity prices but might be an issue in future in case wind production keeps increasing in northern Germany and the grid capacity cannot keep up with this increase. However, collecting the respective data will be difficult because congestions are not part of the transparency-system so far and the various congestion points differ in who manages them and whether they are part of the EPEX auction or auctioned separately.

As McDermott & Nilsen (2011) show, the river levels also have an influence on electricity prices. However, considering the sample size of about two years and the number of variables already included in this analysis, the river levels have not been obtained both due to availability and also to reduce risks of overspecification. As a proxy, temperature is included as it is somewhat correlated with the river levels which are especially low in summer (McDermott & Nilsen 2011).

Variable	Obs	Mean	Std. Dev.	Min	Max
Spotprice	16,776	46.5	15.2	-200.0	131.8
emissionprice	16,776	14.2	1.4	10.4	16.8
CrudeOilBrent	16,752	91.7	17.4	70.7	125.4
temperature	16,776	11.0	7.9	-12.2	28.5
non-usability planned lignite	16,776	1,915.3	1,220.8	0.0	5,916.0
non-usability planned gas	16,776	1,639.5	1,024.1	0.0	5,193.0
non-usability planned oil	16,776	230.9	239.3	0.0	1,658.0
non-usability planned pump-storage	16,776	537.0	442.2	0.0	2,320.4
non-usability planned coal	16,776	2,246.1	1,587.8	0.0	7,963.5
non-usability planned uranium	16,776	2,274.9	2,119.9	0.0	10,978.4
planned non-usability total	16,776	8,926.0	4,878.8	0.0	23,002.5
non-sched. non-usability lignite	16,776	1,148.1	713.0	0.0	4,684.0
non-sched. non-usability gas	16,776	486.2	415.8	0.0	2,334.0
non-sched. non-usability oil	16,776	13.9	74.5	0.0	772.0
non-sched. non-usability pump-storage	16,776	65.0	109.9	0.0	900.0
non-sched. non-usability coal	16,776	1,059.7	628.9	0.0	4,121.3
non-sched. non-usability uranium	16,776	1,077.6	1,647.1	0.0	6,220.1
non-sched. non-usability total	16,776	3,852.7	2,122.4	0.0	11,245.3
Planned Generation Capacity	16,776	44,667.4	8,342.1	20,714.0	67,666.2
Actual Generation Capacity	16,776	41,621.2	7,868.3	21,453.4	63,781.0
solar infeed plan	10,536	2,037.3	3,074.3	0.0	13,982.7
wind infeed plan	16,776	4,530.9	3,769.6	234.8	22,661.0

Table 2: Summary statistics for the dataset

In total, the dataset compromises 16,776 observations from the 1st of November 2009 to 30th September 2011 where each observation represents one hour and includes information about the spot price, the price for CO2-certificates, Crude Oil, temperature, planned wind and solar feed-in, the planned and unplanned non-availabilities for the different kinds of power plants and the total planned power generation for all other power plants that are part of the EEX transparency system. The resulting summary statistics are shown in Table 2.

Three hours in the dataset are affected by the changes to and from Daylight Saving Time. In order to have a complete dataset without any gaps, the mean of the hour before and the hour after has been used for values that are missing during these hours.

3.2 Properties of the Electricity Spot Price

In this section, the most relevant of the available variables, the spot price data, will be examined for a number of statistical properties that are important for econometric time series modelling. Some of the variables like CO_2- or oil-prices might exhibit special properties as well but examining them as well apparently would be outside the scope of this work.

3.2.1 Autocorrelation

Autocorrelation is the correlation of a given variable with itself, most commonly with values earlier in time. This is an important feature of many time series compared to a cross sectional analysis as, at least in an economic context, a value will often depend on its earlier value and will not be randomly distributed.

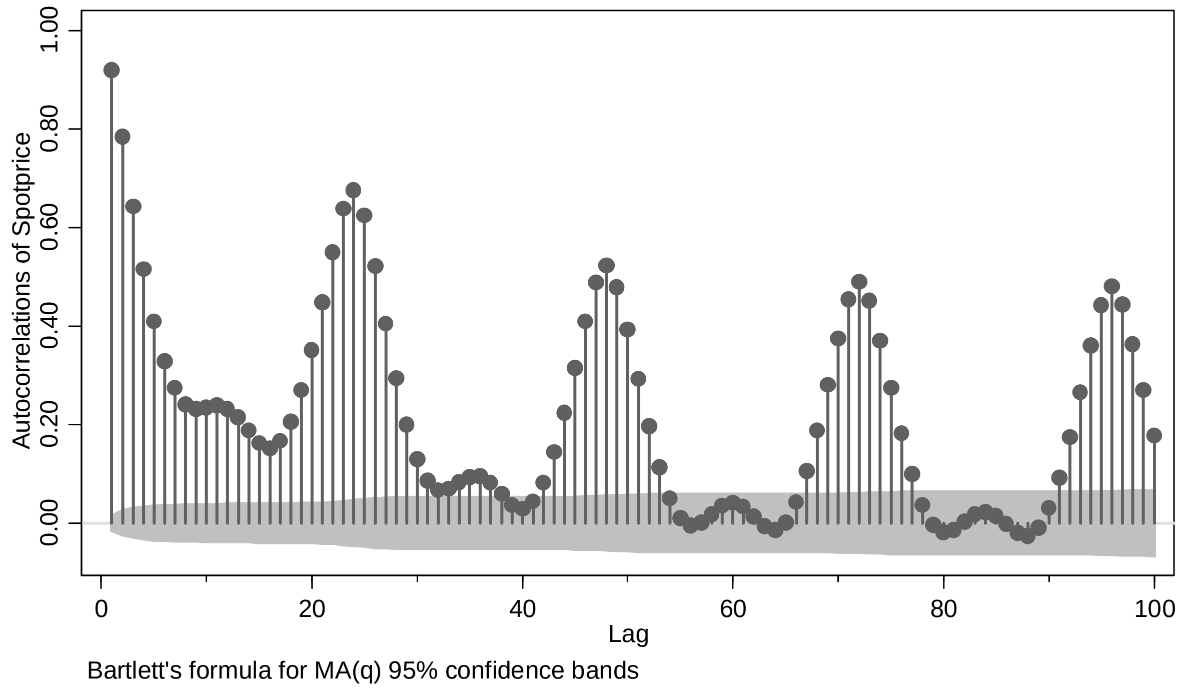

Figure 6: Correlogram of the spot price, full time horizon

Autocorrelation can be visualised by using a correlogram which will plot the correlation of a variable given its lagged values as can be seen in Figure 6, which shows a plot of the full time series of the electricity prices against its own lagged values. Clearly, electricity prices have a strong autocorrelation towards the same hours of the former days, which is why the correlogram shows a peak at the marks at each 24 hours. All peaks lie outside the shaded area that represents the 95% confidence interval. There are also high correlations within the same day, which can be seen for the first lags. However, these correlations cannot be used for the purpose of forecasting electricity prices as they won't be known beforehand: All 24 hours of one day are auctioned simultaneously.

3.2.2 Stationarity

A time series exhibits stationarity, when the joint probability distribution remains stable over time and consequently also mean and variance do not change over time. A time series which does not have these characteristics is called non-stationary. The assumption of stationarity is needed for time series analysis because otherwise the relationship between two variables would change arbitrarily and one could not track correlations between the two in a regression analysis. Using non-stationary time series in a regression analysis can be risky as one might compute significant correlations even though there are none as both variables increase independently from each other. This is called a "Spurious Regression". Non-stationarity is especially problematic in combination with highly persistent time series. A time series is highly persistent when it has a long memory towards even small shocks and therefore does not return to its former mean or variance, thereby becoming a non-stationary process (Verbeek 2008).

To test for stationarity and highly persistent time series, one can use a graphical analysis, a correlogram or the dickey fuller test. The dickey fuller test has the H0 that the time series has a unit root and therefore is non-stationary. This H0 is rejected at the 99% level for all 24 hours of the dataset when tested with the full time horizon, i.e. the time series does not exhibit non-stationarity in general. However, when examining periods of a shorter length of only about 50 days and for some off-peak hours in the dataset used for this thesis, stationarity can be a problem as the H0 of a dickey fuller test cannot always be rejected at high confidence levels. This could be the reason why some other authors explicitly examine issues connected with non stationarity on German electricity prices, as does Liebl (2010) for example

3.2.3 Heteroscedasticity

A sample exhibits heteroskedasticity, when the variance of the error term changes depending on the explanatory variables. When estimating the coefficient by the use of OLS, one has to use heteroskedasticity-robust standard errors so that the standard errors and, consequently, the t- and F-scores remain valid. The estimates of the coefficients however will remain unbiased also in the occurrence of heteroskedasticity (Wooldridge 2008).

To test for heteroskedasticity, there are for example the White test, the Breusch-Pagan test or graphical tests. The White test basically consists of estimating the explanatory variables against the error term u as the explained variable. Should any of the explanatory variables turn out to be significant, one has to reject the H0 that the error term is homoscedastic. Using the OLS-regression that will be presented later in conjunction with a white test for the full time horizon, the H0 of homoscedasticity has to be rejected for every individual hour of the dataset at the 99% significance level besides hour 14, where the H0 is rejected with a 97% significance level according to the chi-square distribution. The Breusch-Pagan / Cook-Weisberg test for heteroskedasticity yields the same conclusion with a 99% significance for all hours. The difference in the significance level for the two tests could be due to the fact that the White test uses a relatively large number of regressors and therefore uses many degrees of freedom (White 1980; Wooldridge 2008).

The findings match with those of earlier works that describe electricity prices to exhibit a “nonconstant mean and variance” (R. C. Garcia et al. 2005) and significant heteroeskedasticity (Swider & Weber 2007).

4 Estimation Models and Methodology

4.1 Overview of available Models

Many different models are used in the literature to forecast electricity prices. Aggarwal et al. (2009) describe Game Theory models, Simulation models and Time Series models which are divided between parsimonious stochastic models, regression models and Artificial Intelligence models.

Game Theory Models try to model the strategies of the market participants and identify a solution of those games. A key point is the analysis of the strategic market equilibrium, which can be based on models like the Nash equilibrium, the Cournot model and others. As Game Theory Models require many assumptions, solutions can vary widely between different models (Haili Song et al. 2002; Cunningham et al. 2002; Aggarwal et al. 2009).

Simulation Models, also described as Fundamental Models, try to build an exact model of the system and the solution is found using algorithms that consider the physical phenomena the process is bound to. This mimics the actual dispatch with system operating requirements and constraints. As Tarjei (2011) explains, the supply and demand side of the electricity market are described and the price at which the two curves intersect is calculated. This price then “equals the marginal cost of the marginal power plant supplying power”. Fundamental models are used by utility companies as they have access to extensive datasets, e.g. Statkraft uses purely fundamental modelling in the spot market and forecasts the hourly dispatch for each of approximately 2500 modelled power plants in Europe (Skrivarhaug 2010). These models can provide detailed insights into the system prices though suffer two major drawbacks: First, they require detailed system operation data and second, the simulation methods are complicated to implement and the computational cost are very high. Furthermore, Simulation models make the assumption that a “fair” value will emerge, which can neglect market trends (Aggarwal et al. 2009).

Time Series analysis focuses on the past behaviour of the observed variable. There are models like multiple regression, autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroeskedasticity (GARCH) models. Normally these are univariate, i.e. focusing only on one variable and its passed values but can also be extended with exogenous variables, then being called multivariate models. Besides these “parsimonious stochastic” time series models, there are also so called Artificial Intelligence models. According to Aggarwal et al. (2009), these nonparametric models “map input-output relationships without exploring the underlying process”. AI models are said to have the ability to learn complex and nonlinear relationships that are difficult to model with conventional methods. However, as Weron & Misiorek (2006) point out, these models are not intuitive and don’t have a simple physical interpretation attached which makes understanding the power market’s behaviour difficult.

An additional method to forecast spot prices is examined by Redl et al. (2009). In theory, when assuming an efficient market hypothesis, the forward price should be a reasonable indicator for the upcoming spot price. However, these approaches do not allow the development of an own opinion and can be influenced by speculations. Redl et al. find questionable results on the predictive power of forward prices as the trading strategies of the market participants actually seem to rely on the spot price: spot prices can be explained well by their own lagged prices whereas lagged forward prices do not significantly influence spot prices. The weak predictive abilities of futures are supported by findings from Hipòlit Torró (2007). Accordingly, one of the other methods described earlier is necessary to make an own forecast of spot prices possible.

For the purpose of this thesis, parsimonious time series models are the most suitable. Contrary to Game Theory Models, they need fewer assumptions because they can rely on more actual data and are therefore easier rooted to the examined circumstances. Yet, they do not need as much data as fundamental models which basically try to simulate the complete market conditions. Compared to Artificial Intelligence Models, the results of time series models will still be intuitive and accessible for interpretation. Drawbacks of Time Series Models are the reliance on past data for forecasting: Per definition it is not possible to forecast completely new market developments. In Addition, it is unlikely to precisely forecast extreme events, as a forecast based on past data will have a tendency towards the mean. These two issues might be tackled by sophisticated fundamental models.

Thus, a number of stochastic time series models will be used for this analysis and discussed in more detail in the following. The most common approaches for time series modelling of electricity prices are a multiple regression using Ordinary Least Squares, and autoregressive moving average models along with conditional heteroskedasticity models using Maximum Likelihood estimation.

4.2 Time Series

4.2.1 Multiple regression and Ordinary Least Squares

The most common method to analyse time series is a multiple regression analysis and using “Ordinary Least Squares” (OLS) to estimate the coefficients. OLS computes the coefficients by minimizing the squared residuals between the observations and a fitted line. In the following, a “multiple regression” is always meant to be estimated with OLS.

The Gauss-Markov Assumptions for time series regressions, which need to be met for a multiple regression OLS analysis, are: First, that the stochastic process follows the linear model

$$y_t = \beta_0 + \beta_0 x_{t1} + \cdots + \beta_k x_{tk} + u_t,$$

where y is the explained variable for time period t, β_k are the coefficients of the explanatory variables x_t and u_t is the error term for t. Furthermore, the assumptions require that there is no perfect collinearity and that the error term u has an expected value of zero for any value of the explanatory variable in any given time period. If these first three assumptions hold, it can be shown that the OLS estimators are unbiased (Wooldridge 2008). Additionally, if the variance of the error term u is the same for all time periods and does not depend on any of the explanatory variables and the errors of two different time periods are uncorrelated for all explanatory variables, the OLS estimators can be shown to are “BLUE", the best linear unbiased estimators depended on the explanatory variables (Wooldridge 2008).

Electricity prices exhibit heteroscedasticity. This means that even though the estimation of the coefficients will still be correct, the standard errors and therefore the t-statistics will be biased. This will be corrected by using “robust” standard errors.

4.2.2 ARMAX and Maximum Likelihood

Many works in the field of forecasting day-ahead electricity prices with econometric methods rely on special time series models like ARMA which will be explained in this chapter and GARCH, which will be explained in the next. ARMAX is a special time series model that includes both an autoregressive term (“AR”), a moving average term (“MA”) and additional exogenous variables (“X”). While the nature of autoregression has already been explained in Chapter 3.2., it is necessary to point out what the notion of a moving average is about: While autoregression captures correlations of the dependent variable with its own former values, the moving average component of the ARMAX model captures past deviations or shocks of the dependent variable and its own lagged values. These could be e.g. due to trends, a seasonality, or variables which are captured by neither the autoregressive term nor the exogenous variables. Thus, the moving average term is useful in describing time series in which events have an immediate effect that only lasts a short period of time (Wei 1990). Another addition that can be done in this context is the differencing of adjacent observations, which is helpful to in order to cope with stationarity issues. Then, the incremental development of the data is observed instead of the absolute values. In this case, the model is called an ARIMA model, the “I” standing for “integrated”, because after the estimation of the models the data needs to be integrated to reverse the initial differencing.

An ARMA approach differs from a “standard” multiple regression in two ways: While an AR term still can be explicitly modelled in a multiple regression analysis, the moving average term cannot. In addition, the estimation of the coefficients is not done using ordinary least squares but rather a maximum likelihood routine, because a non-linear fitting procedure is necessary. The reliance on ARMA and related models by the existing time series literature is believed to be also partly due to historic reasons; when autocorrelation was considered a nuisance when it was not modelled explicitly in the OLS model and that way standard errors were wrong and the estimates no longer efficient (Golder 2007).

When looking at a more formal definition, there are disturbances ε_t which are defined to be the error after fitting and can be calculated through $\varepsilon_t = \hat{y}_t - y_t$ where $\hat{y}_t$ is the predicted and y_t the actual price at time step *t*. The ARMA model is based on the assumption that the error term follows a white noise process, generally assumed to be normally distributed with the form $\varepsilon_t = i.i.d.N(\mu, \sigma^2)$. The time invariant parameters mean μ and standard deviation σ can be estimated by maximizing the log-likelihood function. In the ARMA(p, q) model the relation between the observations y_t and the disturbances ε_t is given by

$$y_t = \hat{y}_t + \varepsilon_t = \sum_{z=1}^{p} \propto_z y_{t-z} + \sum_{z=1}^{q} \beta_z \varepsilon_{t-z} + \varepsilon_t$$

(Swider & Weber 2007). The model is based on considering previous values of the process as a combination of an autoregressive (AR) and a moving-average (MA) part, where the AR part has the order p and the MA part the order q. In this notation, p is referring to the number of previous values of y_t and q is referring to the number of previous values of the disturbances ε_t. The ARMA(p, q) model may then be extended by additional consideration of exogenous variables $x_{z,t}$ with which the model then can be described as

$$y_t = \hat{y}_t + \varepsilon_t = \sum_{z=1}^{p} \propto_z y_{t-z} + \sum_{z=1}^{q} \beta_z \varepsilon_{t-z} + \sum_{z=1}^{r} \gamma_z x_{z,t} + \varepsilon_t$$

where r describes the number of exogenous variables and the model can then be referred to as ARMAX(p, q, r). The parameters $\propto_z$, β_z and γ_z can then be estimated by maximizing the log-likelihood function (Swider & Weber 2007).

As Enders explains, the Maximum Likelihood (ML) estimation uses the following principle: If values of $\{\varepsilon_t\}$ are drawn from a normal distribution with a mean of zero and a constant variance σ^2, from standard distribution theory the likelihood L_t of any realisation of ε_t would be

$$L_t = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) exp\left(\frac{-\varepsilon_t^2}{2\sigma^2}\right).$$

As the realisations ε_t are independent from each other, the joint realisation for all values of t is the product of the individual likelihoods. Hence, given the same variance for all realisations, the likelihood for the joint realisation is

$$L = \prod_{t=1}^{T} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) exp\left(\frac{-\varepsilon_t^2}{2\sigma^2}\right).$$

The method used in maximum-likelihood estimation is to select the distributional parameters so as to maximize the probability of drawing the observed sample (Enders 2010). As a simple example, ε_t could be generated from the model

$$\varepsilon_t = y_t - \beta x_t$$

Accordingly, maximizing the log-likelihood function would involve solving for the values for σ^2 and β.

Literature is not clear if one has to correct for heteroscedasticity to get correct test statistics, so as a precaution, robust standard errors will be used in the following when estimating the models. Earlier works that forecast electricity prices using ARMA techniques usually did not state explicitly that robust standard errors are used. Even though this does not influence the outcome of the forecasts, it can be a factor in the development of the forecasting models, when the significance of the coefficients needs to be evaluated.

4.2.3 GARCH and Maximum Likelihood

Electricity prices have the special property that large price changes are often again followed by large price changes. For example, in the case of little demand and large wind in feed, prices will fall abruptly but when this special situation is gone, prices will change by a similar magnitude to the original level again ("Mean Reversion").

By applying a GARCH approach, this conditional heteroscedasticity can be considered. Conditional heteroscedasticity means a time variant variance $\sigma^2\ t$ in which large changes tend to follow large changes, and small changes tend to follow small changes, which is described as the volatility clustering. GARCH(p, q) models are designed to capture this changing volatility by calculating the variance in the following way:

$$\sigma_t^2 = \omega + \sum_{z=1}^{p} \propto_z \sigma_{t-z}^2 + \sum_{z=1}^{q} \beta_z \varepsilon_{t-z}^2$$

Accordingly, the time variant variance is described with a constant part ω, an AR-part of order p and a generalized MA-part of order q. A necessary condition is that the variance is positive at any time step t, i.e. that $\omega > 0$, $\alpha_z \geq 0z$ and $\beta_z \geq 0$. The GARCH term will be included in a regular ARIMA model to model the white noise ε_t and the parameters can then be estimated by maximizing the log-likelihood function. As the name indicates, GARCH is a more generalized version of the ARCH model for which Robert F. Engle received the Nobel Prize in Economics 2003. In the ARCH model, the volatility is only depended on the realisation of the error term in the previous period(s) and not also on its own realisation in the previous period(s). Accordingly, a GARCH(0,1) model is the same as an ARCH(1) model. A GARCH model that makes use of exogenous variables is called an MGARCH. In the estimation of the GARCH model, no heteroskedasticity-robust standard errors will be used as the model explicitly models the variance (Swider & Weber 2007; Enders 2010).

4.3 Methodology for Model Estimation

The usage of an ARIMA or GARCH model for forecasting can be facilitated by following the Box-Jenkins Methodology. This is a three-stage method which is divided into the identification stage, the estimation stage and the diagnostic checking stage. First, one will visually examine the time series by plotting the variables, the autocorrelation function and the partial correlation function. This will allow to check for trends and missing values and give a first grasp for plausible models, e.g. for the number of lags which will be necessary to develop a model with a good fit. Then the tentative models will be estimated and checked for a good fit, which will be done in this work by the forecasting performance for the two out-of-sample periods (Enders 2010). Forecasting time series has often been described as “more art than science”, since there are many different approaches and the aims for modelling can be very different (Burman & Shumway 2006). In this work, the focus will lie on developing a model that can be used for forecasting.

The statistical hypothesis testing will mainly include the t and z-tests of the individual variables. The t-test basically sets the estimated coefficient in relation to the standard deviation and in connection with an assumed distribution can make inferences on the statistical significance of the coefficient. The z-test is a quite similar test and is used instead of the t-test for the ARIMA and MGARCH mode due to slightly different assumptions about the standard deviation (Gaten 2000).

To measure how well the model forecasts the spot price when it is out-of-sample, the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE) will be used. The MAE is the average of the absolute Forecast Errors:

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

where t and T is referring to the number of hours that are forecasted and t=1 is the first hour of the out-of-sample period (Wooldridge 2008). Accordingly, the MAPE is defined as

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

Following other works, also the Root Mean Square Error (RMSE) will be shown:

$$RMSE = \sqrt{T^{-1}\sum_{t=1}^{T}\left(P_{t,fc} - P_t\right)^2}$$

Due to the quadratic term, the RMSE is usually more sensitive towards outliers and therefore usually has a higher value then the MAE. These three measures are widely used in the corresponding literature and so both the possibility for evaluation of the results and the comparability with other works is ensured (Aggarwal et al. 2009).

4.4 Time Horizon

Since the purpose of this work is to develop a forecasting-model, in-sample and out-of-sample subsets are created: The first subset is used to estimate the model parameters, and the forecasting abilities of the model are then evaluated using the out-of-sample subset. There are different ways to define the estimation window. First, there is the “Rolling Window” technique, in which a specified number of observations is used to make the model estimation. When one day passes, the estimation window is moved one day forward and the same number of observations is used to estimate the model and the new day replaces the oldest day of the subsample. In this model, both the starting date and the ending date of the estimation window are variable. In the “Jackknife” model, the estimation window is extended with each new day that can be included in the dataset, i.e. the starting date is fixed but the ending date is variable. The third way is to have a fixed calibration period and use the estimates from this calibration period to estimate all the days of the out-of-sample period. The existing research uses mostly this last way because the forecasting periods normally are not longer than a month and the estimates of the coefficients should not change drastically by changing the sampling size by a small number of days. The tests that have been done with the dataset showed that changing the sample size was leading to fluctuations in the area of about 0.4%-points regarding the average MAPE. As a optimization of this magnitude doesn’t justify the extensive additional computation time necessary for “Jackknife” or “Rolling Window” estimations, the third way of a fixed sampling will also be used in this work, as it has been used in other works before (Dias 2010).

Another question that arises considering the time horizon of the data is how to model the individual hours of the day. In the spot price auction of the daily prices, the prices for each hour are simultaneously determined on the day before at noon. Accordingly, the participants get to know the prices for all hours simultaneously – That means that it is not possible to forecast e.g. hour 10 by using information from hour 9, since the prices for hour 9 and 10 will be published together. Accordingly, every model that will be used in this work has to be refined in a way that only the realistically available data is considered in order to effectively simulate the forecasting process. To include this consideration into the econometric model, one could either use lags that are larger than the number of the currently estimated hour or it could be done by estimating each hour dependently and using lags of one, thereby referring to the day before. The second option which produces 24 independent time series has been used by Cuaresma et al. (2004) and, as it turns out, produced more precise forecasts than modelling the electricity prices as one single time series. This could be confirmed in preliminary datasets with this sample and additionally it has been found that the computation time decreases when only individual hours and autoregressive lags in the magnitude of 1 are estimated against the estimation of the complete sample and the respective lags in the magnitude of 24, which might also be due to the amount of exogenous variables that are included. While the estimation time on the used system, an Intel i7 processor with 2.7 GHz using STATA, is a few minutes for the individual hours, all hours taken together needed more than 60 Minutes estimation time for a similar model.

4.5 Existing Forecasting Results from Time Series Models

There is a number of other works who employed ARMA/GARCH models on the spot prices of EEX in the past for the purpose of forecasting:

- Keles et al. (2011) use Mean Reversion, ARMA, and GARCH models on the full hourly time series and try to explicitly model negative prices. The time range for the Calibration Period is 2002-2005 and the simulation is run for 2006-2009. They state a MAPE for this time from 16.02% to 21.06% depending on the model and a RMSE of 8.43 to 23.53.
- Swider & Weber 2007: With ARMAX and GARCH Models, a MAPE from 12.92% to 13.49% and a MAE of 4.33 to 4.51 is reached for the different models and for an in-sample subset which ranges from June 2002 to May 2004. As exogenous variables, known price information from other markets like the reserve market for ancillary energy is used. In another, similar article from one of the authors, a MAPE between 17.96% and 19.22% is reported where the hour 12 was modelled individually for an out-of-sample subset (Swider 2006).

- Cuaresma et al. 2004: With an univariate approach, a MAPE of down to about 13.2% and a MAE between 2.60 to 7.13 with a calibration sample ranging from June 2000 to September 2001 and a number of different AR(1) and ARMA models is achieved. The out-of-sample period ranges from September to mid October 2001. These authors also specifically employ the approach of modelling the 24 hours individually. Spikes have been removed in this work.

In Addition to the works based on the EEX, there are many articles that focus on other markets:

- Rafał Weron & Misiorek 2008: The authors examine two markets using a wide range of advanced Time Series models: For the Nordpool market, they report an MAPE of only 3.2%. The other market they examine is the Californian market where a MAPE of 12.96% is the best result for a model in which spikes have been pre-processed in a way that they are dampened.
- Gianfreda & Grossi 2011: In a recent study of the Italian Electricity market, the authors report a MAPE between 10.69% and 12.63% and a RMSE of 9.57 to 12.17 for the different zones of the Italian System. The forecasting was done with a Reg-ARFIMA-GARCH model and included exogenous variables for technologies, market power and network congestion using a rolling window approach.
- Aggarwal et al list a compilation of papers on forecasting for different markets and methods and state that the reported MAPE is usually in the range of about 3% up to about 20% (2009).

These earlier works show that there are considerable differences regarding the precision of the forecasts that can be reached in the different markets. Especially a heavy influence of hydropower plants reduces the volatility of the prices, as the owners of the power plants have few marginal costs besides the opportunity costs and will try to only use water to generate power in times of sufficient prices, thereby limiting the range the electricity prices can have. The EEX spot price is harder to predict than the prices from other markets, as it exhibits more fluctuations than other markets. This can be due to its size on the one hand and on special production characteristics on the other hand that lead to more price changes than e.g. water-dependent Nordpool.

5 Analysis of the Results

5.1 Obtained Models and influence of Transparency-Data

5.1.1 Model Development

5.1.1.1 General Procedure

During the development of the models that will be used for forecasting, several findings have been made. As already described, the models are determined using a Box-Jenkins approach. To generate suitable models, the following steps have been used:

1. The average MAPE/MAE that are achieved by the use of an adjusted model are compared to the model that has been used before
2. second in order to determine how the MAPE/MAE can be increased further, the significance tests are used to determine which variables might have to be dropped

For each model category (i.e. ARIMAX, ARIMA, MGARCH, OLS) the same model is used for all individual 24 hours. Since however the dataset is slightly different for all those 24 hours, the estimates will vary. The 24 estimation results for the complete time horizon for all 4 model categories can be found in the Appendix.

Only variables that are “known” at the point in time where the auction is conducted are included. That way, the forecasting process is simulated just like a market participant would experience it because he will not have access to the actual information ex ante. Accordingly, forecasts are used, like the planned generation capacity for the next day or the forecasted wind in-feed. In case there are no forecasts available, a lagged variable will be used. This is why the oil price and the emission price are used in lagged form.

An important principle of the Box-Jenkins approach is parsimony, which means sparseness or stinginess. While additional coefficients increase the fit, they lead to a reduction of degrees of freedom. A parsimonious model will fit the model well without incorporating needless coefficients – and these needless coefficients will not be projected into the future by using them for forecasting. Consequently, it is recommended to eliminate weak coefficients or coefficients with strong correlation between each other for the purpose of forecasting. This is not only true for ARIMA and GARCH models but also for multiple regression models (Enders 2010).

5.1.1.2 Specific Adjustments

Many of the variables that have been discussed in chapter 3 also are included in the model. Even though economic reasoning can explain that they are important determinants of the electricity price, some of these variables are not statistically significant in the current dataset and therefore will not be used in the current setting to reduce issues associated with overspecification and accordingly stick to the concept of parsimony. The same decisions have been made for all models simultaneously in order to keep the results comparable. The following variables have been dropped in the process of the model development:

- The variables for the unavailability of oil and pump storage power plants have been dropped: As they have very few periods of unavailability, it is difficult to reliably estimate price effects – which is why they turned out to be very insignificant in the estimated models.
- The data for solar power is only available from about halfway through the dataset – it wasn’t made public before – and only reaches a certain magnitude on a few summer days. Both effects taken together also lead to a low significance of solar power given the amount of other exogenous variables that are used in this examination and therefore also the coefficient for solar power has been dropped.

For periods of high wind and little demand, relatively high MAE and MAPE have been discovered and therefore, different variations to include the wind data have been used during the development phase of the models. However, neither the use of dummies for different wind levels, quadratic forms nor combinations of those could significantly improve the forecasting errors. Possibly forecasting can improve when one can facilitate a larger sample which exhibits more of these extreme situations.

Electricity prices exhibit seasonality besides the hourly basis. At first, dummies for each weekday have been included in the model. However, dummies for each day have turned out to be insignificant for the working days, because there are only few differences between them. There are substantial differences between the working days and the weekends though. Accordingly, a weekend-dummy has been used to keep the balance between the significance of additional variables and parsimony. Public holidays are included through an additional dummy and include holidays where the majority of the German federal states have statutory holidays.

Many earlier works that have been conducted for electricity price forecasting use a logarithmic form for all the variables, because that way, a more homogenous variance and smoother volatility is reached (R. C. Garcia et al. 2005; Nogales et al. 2002). However, using the current dataset and the explained models, a level-log form generated the best forecasting performance. “Level-log” form means that the explanatory variables are in log-form, while the explained variable remains in level-form, i.e. not using a logarithmic transformation. It turned out that while MAPE and MAE of course were lower for the log-log-form when comparing the logarithmic price and its logarithmic forecast, MAPE and MAE were inferior when calculating the results back to the level-form. For both selected periods, the forecasting of a level-log model yielded better performance measurements than the respective log-log model. Taking MAPE as an example, in the September Period the MAPE for a log-log model was 9.9% while it was only 8.3% for the level-log model. This is why a level-log form was used in this work. The logarithmic form was used for all explanatory variables. To calculate the logarithmic form of the temperature, a constant has been added to all values in order to avoid negative terms. Other variables do not have negative values and therefore did not need an adjustment.

For the ARIMAX and MGARCH models, there is a wide range of possibilities to include the autoregressive and moving average terms. The AC and PAC suggest the use of several lags. As it turns out, for both the AR and the MA term, significance for lags higher than one is very low. Since the aim of the model development is to identify the smallest and most simplistic model that still adequately describes the data, both the AR and the MA terms have been added for one lag, which, as every hour is modelled individually, then refers back to the same hour of the day before. These findings match the results of earlier works. Including more lags in the ARIMAX and MGARCH models needed considerable more computation time and only marginally improved the MAPE error, by around 0.1%-points.

The ARIMAX and ARIMA model are used in a first differenced form, which seems to avoid possible weak stationarity issues as the MAPE dropped by a few percentage-points. Using a first differenced form was not possible for the MGARCH model, because the model gets to complex and finding an iterative solution for the maximum likelihood operation is not possible anymore.

As for the third step in the Box-Jenkins Methodology, the models are checked that the residuals follow a white noise process. To do that, Enders speaks of the option to use the out-of-sample forecast performance as a measurement which is convenient given the focus of this work (Enders 2010).

To be able to compare the results of the multivariate approach that has been conducted in this work with a univariate approach, a simple univariate ARIMA model with an AR(1) and a MA(1) term has been included. It is differenced one time, making it an ARIMA(1,1,1) model.

5.1.2 Model Results

Since every hour is estimated individually, and four models are estimated for each hour, there are in total 96 estimation results. However, the general picture of the coefficients will be similar – the same model is estimated using the data of the various hours. Therefore, the coefficients will be explained on the example of Hour 13, going from 12:00 to 13:00, for the four estimated models, since it is the hour with the highest traded volume. In order to show the complete picture, the remaining hours are given in the Appendix. While the magnitude of the coefficients might change in other hours, the tendency and the sign of the coefficient is very similar in most cases for all hours. The results will normally be explained by referring to the ARIMAX results as it is the model with the lowest forecasting errors and be supplemented with information from the other models where appropriate.

Variable	explanation
log_PlannedGeneration	the total amount of electricity generation from conventional power plants that is planned in Germany for the next day for the particular hour, in
log_windplangermany	the planned amount of electricity generated from wind energy, based on wind forecasts, in MW; log-form
L1.log_emissionprice	the current price for CO2-certificates, in EUR/t; log-form
L1.log_CrudeOilBrent	the current price for the relevant oil sort, in $/BBL; log-form
log_temperature	the actual temperature for the next day for a particular hour in °C; log-
log_unav_lignite_planned	planned unavailability for lignite power plants for each hour in MW, log-
log_unav_Gas_planned	planned unavailability for gas power plants for each hour in MW, log-
log_unav_coal_planned	planned unavailability for coal power plants for each hour in MW, log-
log_unav_uranium_planned	planned unavailability for nuclear power plants for each hour in MW, log-
log_unav_lignite_nonsched	unplanned unavailability for lignite power plants for each hour in MW, log-
log_unav_Gas_nonsched	unplanned unavailability for gas power plants for each hour in MW, log-
log_unav_coal_nonsched	unplanned unavailability for coal power plants for each hour in MW, log-
log_unav_uranium_nonsched	unplanned unavailability for nuclear power plants for each hour in MW,
d_weekend	weekend-dummy variable, 1 if the hour is within a weekend
d_holiday	holidy-dummy variable, 1 if the hour is within a statutory holidy

Table 3: Explanation of used variable names

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	34.18101***		33.803052***	19.633506***
log_windplangermany	-2.5541961***		-2.5246853***	-2.7752924***
L1.log_emissionprice	12.39411000		14.43999900	9.5815852***
L1.log_CrudeOilBrent	13.73995600		16.923396**	12.7667***
log_temperature	-4.23377800		-3.787267*	-3.7443177*
log_unav_lignite_planned	0.24994634		0.23895162	.5076965***
log_unav_Gas_planned	-0.29971647		-0.20663191	-0.28694623
log_unav_coal_planned	.49334207**		.50889382**	.55880774***
log_unav_uranium_planned	0.11255363		0.08470135	.27669921**
log_unav_lignite_nonsched	0.09749453		0.05122216	-0.05375778
log_unav_Gas_nonsched	-0.04550184		-0.03018660	0.03438182
log_unav_coal_nonsched	.52797819**		.36739753*	.50797372**
log_unav_uranium_nonsched	.4153203*		.39921933**	.39678567***
d_weekend	-4.3652562**		-4.6244609***	-7.5259589***
d_holiday	-5.7572398***		-5.2969489***	-6.0687815***
_cons	-380.48493***	0.02381044	-398.78192***	-225.68903***
ARMA				
L1. ar	.96447413***	.31991917***	.96002614***	.21239769***
L1.ma	-.75071009***	-1.0488318***	-.71037717***	
sigma				
_cons	5.2337186***	8.3331067***		
ARCH				
L1.arch			.22118258***	
_cons			21.857504***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 4: Estimation Results ARIMAX, ARIMA, MGARCH and OLS for hour 13, whole data sample, level-log form

As heteroeskedasticity has been described as an important characteristic of electricity prices, one has to use heteroeskedasticity-robust standard errors for the t- and z-statistics to remain valid. This is done for all models besides the MGARCH model where the volatility is modelled explicitly. Many other works on the forecasting of electricity price do not explicitly describe taking this step. Hence, the test statistics for the following results could be considered as being rather conservative compared to other works. The significance tests are only important for interpreting the model results and to make decisions about whether or not include a variable; a low significance has in itself no effect on the forecasting.

As the Estimation Results in Table 4 show, wind in-feed lowers the price at the power exchange. On average, a 10% increase in wind in-feed changes the electricity price by -0.255 EUR/MWh following the ARIMAX model (MGARCH: -0.252 EUR/MWh, OLS: -2.78 EUR/MWh). The interpretation of the coefficients is done using the template by

Wooldridge, which states the interpretation of a coefficient in level-log form to be $\Delta y = \left(\frac{\beta}{100}\right) \%\Delta x$ (Wooldridge 2008). While wind is significant at the 0.01%-level, this was not the case for solar in-feed. For Solar, even positive estimates were within the 95% confidence interval of all the model results – which of course is against the economic logic. To avoid overspecification issues, the solar in-feed has therefore been dropped. The low significance can be due to the limited time for which the data on solar energy is available, which does not start until mid-July 2010 and also the amount of solar in-feed during this period was lower than wind in-feed, with solar energy amounting only to about 43% of the electricity generation from wind. This makes it more difficult for the estimation routines to calculate significant estimates. The finding of negative coefficients is consistent with earlier findings that discussed the price effect of renewables. Recent studies about the impact of wind energy on market prices in Denmark (Morthorst 2007) and Germany (Neubarth et al. 2006) observed price reductions of about 12–15% in the long run. Sensfuß et al state that the so called "merit-order effect" caused by renewables for the year 2006 reached a volume in the magnitude of 3–5 billion EUR (2008).

Both the oil-price and the price for CO2-certificates are positive, which intuitively makes sense as they are input-factors for electricity generation. A 10% increase in the oil-price increases the electricity price by about 1.37 EUR/MWh using the ARIMAX estimates (MGARCH: 1.69 EUR/MWh, OLS: 1.28 EUR/MWh). The results are significant at high confidence levels for all models, though the significance is somewhat lower for the ARIMA and MGARCH models. In the ARIMAX model, the estimate for the influence of the oil price is only significant at the 72%-level. The difference in the significance between the MGARCH model and the ARIMAX model is probably due to the robust standard-errors for the ARIMAX model while the volatility is modelled explicitly for the MGARCH model and therefore the standard errors can be calculated more precisely. The picture is similar for the estimate of the emission-price coefficient, where a 10% increase is estimated to increase the electricity price by 1.24 EUR/MWh for the ARIMAX model (MGARCH: 1.44 EUR/MWh, OLS: 0.96 EUR/MWh). The significance is above the 90% significance level for ARIMA and MGARCH and above 99% for OLS.

Even though temperature will not be as important in Germany as it is for other electricity markets, it still remains statistically significant in the MGARCH and OLS model. On average, a 10%-temperature increase will decrease the electricity price by 0.42 EUR/MWh following the ARIMAX model (MGARCH: -0.38 EUR/MWh, OLS: -3.7 EUR/MWh). The reason for this coefficient is probably that heating with electricity is still done in some instances and apparently outweighs the use of Air Conditioning in the warmer months. Because one could expect electricity prices to rise, when a lot of energy is needed for air conditioning when river levels are low, a quadratic form of temperature has also been estimated. However, the calculated turning point was far above normal temperature levels (Wooldridge 2008), so a quadratic form of the temperature does not improve the model. This is probably due to the case that in the sample period of the two years, the summers haven't been particularly hot and there hasn't been any significant outages due to low river levels, so when extending the dataset to times with higher temperatures and the resulting impacts on the electricity price a quadratic form of temperature could be considered – which however won't be easily possible yet as the transparency data is only available after a certain point in time.

Concerning the unavailabilities, only the estimates for coal power plants, both for planned and non-scheduled outages and uranium power plants for non-scheduled outages can be considered statistically significant for the ARIMAX and MGARCH model. All their estimates are positive, which makes sense since a cut in supply should increase prices. The other estimates cannot be considered statistically significant at high significance levels for ARIMAX and MGARCH which may be due to the relatively limited and the complexity of the model. However, the other estimates are also positive within their respective 95% confidence intervals.

A finding that has been expected is that the coefficients of the different types of power plants line up in the way they also align in the merit order that has been explained in Chapter 2, i.e. that the price effect of an unavailable nuclear power plant is the highest, while the coefficient then gradually decreases for the power plants which are higher up in the merit order and therefore have a lesser price effect. This expectation however, cannot be proven with this data as the significance levels are too weak: right now actually the price effect of a coal power plant seems to be higher than that of a nuclear power plant. However, the theory that the coefficients line up in the merit order might still be true, as the confidence intervals of the estimates still allow for this possibility. Further research and a more extensive dataset will be necessary to estimate the coefficients with more precision.

The weak significance of the coefficients can also be a sign of competition: a high significance would indicate a lot of power for the power plant owner as he could be sure about the price effect when turning off a power plant – which, as the current analysis shows, he can only be for the large base-load power plants which are further down in the merit order and that way definitely will cut supply. On the other side, he cannot easily influence prices by not offering a power plant to the market which has higher marginal costs and is easy to control. When one market participant does not offer a gas power plant on the market, the others might just balance out that move by offering more of their own available capacities or even going above the own capacity for some time by e.g. running a power plant at over 100% capacity if profitable.

For the gas-based peak power plant, also negative coefficients are in the range of the confidence interval – which is understandable as they are on top of the merit order. Power suppliers might even schedule maintenance periods in times with expectedly low prices. The iteration routing that estimates the model will connect the low price with the unavailability and that way generates the possibility of a negative coefficient, even though a negative coefficient should not occur when supply is cut considering basic economic theory.

When researching the use of dummies for the weekdays, using individual dummies for each day yielded a weak significance for the weekdays and only Saturday and Sunday were considered statistically significant at a high significance level. Probably the variations within the weekdays are not large enough as they could be assigned to a certain day. To reduce the impact of overspecification, only a weekend dummy has been utilised. The weekday dummy is statistically significant at the 99%-level and the ARIMAX model predicts that prices for the weekend are about 4.4 EUR/MWh lower when controlling for all other factors (MGARCH: -4.62 EUR/MWh, OLS: -7.52 EUR/MWh). However, demand is already explicitly controlled for through the use of the total planned generation, so it remains unclear why there is an additional discount on the prices for the weekend that goes above the limits that are set by total demand.

All the models seem to be consistent with economic reasoning in both magnitude and significance. In total, the OLS model seems to find a higher significance in the exogenous variables but does not have the moving average term which is estimated for the ARIMAX and MGARCH model. This moving average term possibly comes on cost of degrees of freedom, thereby reducing the significance of some of the other estimated coefficients. A comparison of the out-of-sample forecasting performance will yield answers on which model is most suited for forecasting.

5.2 Forecasting Performance

The forecast that is being undertaken is a conditional forecast: some information from the future like the planned generation capacity is known and can be used for forecasting (Wooldridge 2008). The forecasts for the out-of-sample period are calculated based on the estimates made in the in-sample period. This is done for two different periods to check how the approach works in different seasons, for both September 2011 and for February 2011. In both cases, the sample size has been left equal so that differences can only be due to the differences in season and not due to the differences in the time horizon. Again, each hour is modelled individually.

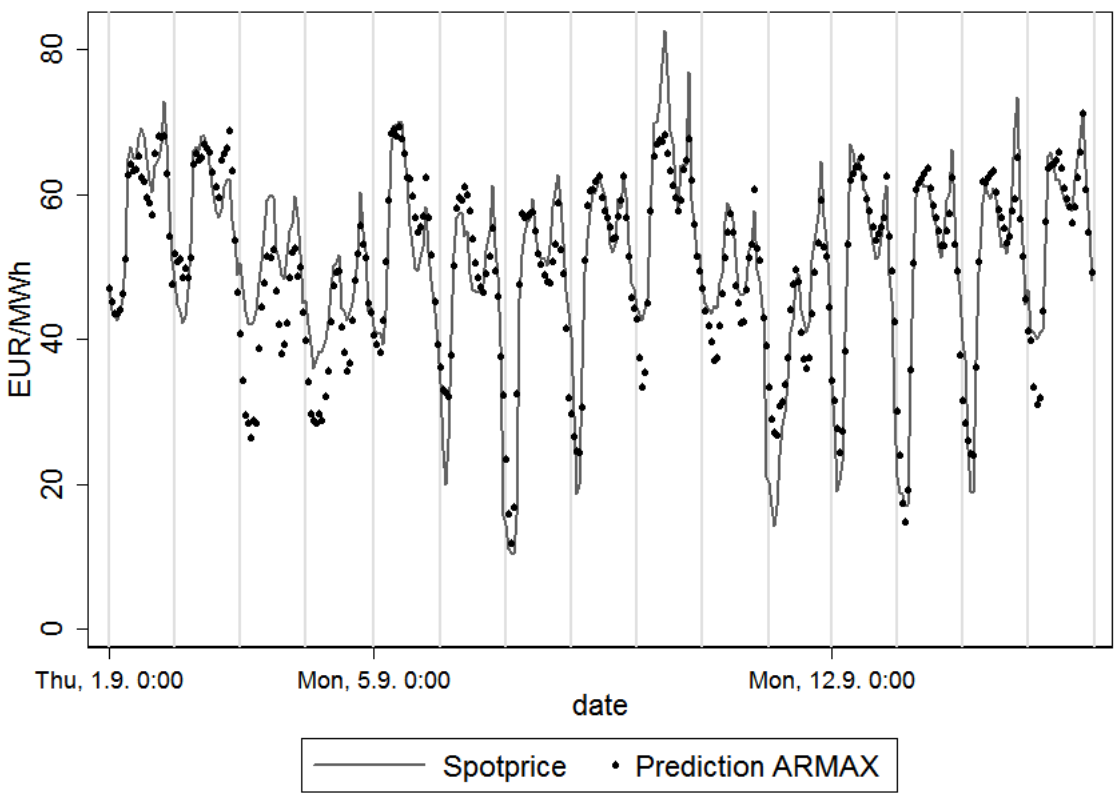

Figure 7: Forecasted and actual Spot Price for first two weeks September 2011, based on estimation period from June 2010 to August 2011

Table 5 shows the forecasting Errors for September 2011 for the different models that have been developed. All models share a relatively low forecasting error for the time between 7am until midnight, but are significantly higher for the early morning hours. This is another indication why hourly forecasts are better than modelling all hours together: Hours with a high unpredictability will not influence hours with a better predictability. A common characteristic of those hours with high deviations are little demand and lots of wind-in-feed. As already pointed out, a number of variations for the modelling have been tried but did not yield significant improvements in the forecasting performance, so possibly also other factors play a role which are not captured in this dataset, like congestion or different bidding strategies by the market participants.

	ARIMAX			**ARIMA**			**OLS**			**MGARCH**		
Hour	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE
1	14.1%	4.4	6.5	20.6%	6.1	9.1	15.4%	4.9	6.8	15.4%	4.9	7.0
2	13.5%	4.3	5.7	21.3%	6.0	8.4	15.1%	4.6	6.2	14.5%	4.5	5.7
3	16.4%	4.6	6.1	26.7%	6.6	8.9	19.4%	5.1	6.4	18.6%	4.9	6.1
4	19.5%	5.0	6.3	30.9%	7.2	9.9	24.3%	5.8	7.2	20.5%	5.1	6.2
5	17.8%	4.8	6.0	29.1%	7.0	9.5	20.2%	5.0	6.4	20.9%	6.7	9.6
6	14.2%	4.4	6.1	19.5%	5.8	8.0	14.3%	4.6	5.7	14.5%	4.8	5.9
7	8.1%	3.9	5.7	18.2%	8.4	10.7	11.3%	5.7	7.4	8.8%	4.4	6.1
8	7.9%	4.3	5.0	20.9%	11.1	13.6	9.1%	5.3	7.0	22.0%	12.1	16.0
9	4.8%	2.8	3.6	17.5%	9.8	11.7	6.5%	3.9	5.2	12.6%	7.3	9.8
10	5.0%	2.9	3.9	13.0%	7.6	9.6	6.2%	3.9	5.4	14.5%	8.7	11.5
11	5.7%	3.4	4.3	12.0%	7.1	9.0	6.0%	3.9	5.5	12.4%	7.7	10.8
12	5.4%	3.5	4.8	10.8%	6.7	8.8	5.5%	3.6	5.5	13.7%	8.6	11.9
13	5.5%	3.4	4.4	9.9%	6.0	7.7	5.7%	3.6	4.9	11.5%	7.1	10.2
14	5.7%	3.3	3.9	12.1%	6.9	8.5	5.4%	3.1	4.2	11.9%	6.8	9.7
15	6.7%	3.6	4.2	12.4%	6.6	8.1	5.9%	3.2	3.9	10.9%	6.0	9.1
16	7.0%	3.6	4.4	10.9%	5.7	7.1	5.9%	3.1	3.8	14.3%	7.4	10.7
17	6.6%	3.3	3.9	9.2%	4.8	6.0	5.9%	3.0	3.7	15.7%	8.0	11.3
18	5.6%	3.0	3.8	8.5%	4.8	6.0	5.1%	2.8	3.4	10.1%	5.6	7.9
19	4.9%	2.9	3.4	6.9%	4.3	5.1	4.8%	2.9	3.5	5.6%	3.5	4.1
20	4.8%	3.2	4.0	6.1%	4.1	5.1	6.6%	4.5	5.6	5.7%	3.8	4.8
21	5.3%	3.5	4.2	5.8%	3.9	5.3	10.5%	7.3	8.1	5.3%	3.6	4.6
22	5.8%	3.3	3.8	6.6%	3.7	4.4	7.1%	4.2	4.7			
23	4.3%	2.2	2.6	5.0%	2.6	3.2	4.0%	2.1	2.6	3.7%	1.9	2.4
24	5.4%	2.3	3.5	7.8%	3.4	4.5	5.4%	2.3	3.4	5.0%	2.1	3.5
ø	8.3%	3.6	4.6	14.2%	6.1	7.8	9.4%	4.1	5.3	12.5%	5.9	8.0

Table 5: Forecasting Errors for September 2011

The results for the February period are considerably higher for the MAPE figure, for which a number of reasons has been identified. First, the February-period exhibits a lot of extreme prices with very small spot prices at just about a few cents, while the forecasting models predict prices of about 20 EUR/MWh, which of course leads to large deviations. In the case of MAPE, the deviations for those hours are in the magnitude of several hundred percent. This shows the importance of using not only the MAPE, but rather several measurements and that the MAPE has weaknesses for exceptionally low prices as the percentage deviations can get very large. Second, the two estimation periods differ in terms of their volatility and also the "representativeness" of the in-sample period for the out-of-sample period differs. Third, the stated "planned generation" seems to be flawed in some instances. There are cases in which the stated figures are not consistent: This can be seen in periods of high wind in-feed and the amount of planned generation is stated, as an example, it is unrealistic that a generation capacity of about 40.000 MW is necessary from conventional power plants when the forecasted wind is about 22.000 MW on a Sunday. The actual generation for this example then resolved to be at around 30.000 MW. Examples like that can be found in dozens of cases for the February period, while there seem to be no cases in the September period in which the planned generation capacity was so far off compared to the actual generation the day later. It seems that the data quality of the planned generation capacity improved considerably in between the two periods: While the mean difference between planned and actual generated capacity was 5,387 MW for the February period, it was only -124 MW for the September period.

	ARIMAX			**ARIMA**			**OLS**			**MGARCH**		
Hour	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE
1	40.0%	6.3	8.4	46.1%	7.7	10.1	38.5%	5.3	7.7	37.6%	8.8	10.0
2	32.1%	6.2	7.7	41.0%	7.9	9.4	36.8%	5.7	7.5	39.9%	6.1	7.7
3	22.6%	6.0	7.3	30.6%	8.2	10.0	25.8%	5.7	7.1	30.2%	6.8	8.7
4	138.8%	5.3	6.2	638.5%	7.7	9.1	145.6%	5.4	6.4	343.9%	5.9	6.6
5	574.9%	5.4	6.9	1721.1%	7.6	9.3	675.4%	5.2	6.2	895.5%	5.2	6.8
6	458.8%	6.5	7.9	1137.4%	7.9	10.3	802.2%	5.4	6.9	709.2%	6.1	7.7
7	20.7%	6.6	8.8	39.5%	11.0	15.1	21.3%	6.0	8.4	17.4%	6.2	8.6
8	755.5%	5.3	7.4	2457.2%	14.4	18.0	865.7%	4.8	6.6	794.4%	4.9	6.7
9	8.4%	4.4	5.8	27.5%	12.2	14.2	9.4%	4.0	5.2	8.8%	3.9	5.1
10	6.1%	3.2	4.4	18.0%	9.0	10.6	6.8%	3.3	4.4	6.2%	3.0	4.1
11	7.0%	3.6	4.9	15.9%	7.9	9.7	7.9%	3.8	5.0	6.9%	3.2	4.6
12	7.2%	3.6	5.1	14.9%	7.3	9.1	8.2%	3.9	5.2	6.9%	3.2	4.7
13	7.5%	3.5	5.1	14.5%	6.9	8.9	8.0%	3.6	5.1	7.3%	3.2	4.9
14	6.9%	3.0	4.2	16.7%	7.5	9.5	7.2%	3.1	4.1	6.7%	2.8	4.1
15	8.3%	3.3	4.3	19.0%	7.8	10.0	9.5%	3.5	4.5	7.9%	3.0	4.2
16	8.4%	3.2	4.5	18.2%	7.3	9.3	9.4%	3.6	4.6	8.5%	3.5	4.5
17	7.8%	3.3	4.5	15.3%	6.4	8.4	12.8%	5.8	6.7	8.6%	3.8	4.9
18	7.4%	4.3	5.9	8.1%	4.5	5.8	9.6%	5.8	6.9	6.3%	3.7	5.0
19	7.5%	4.9	8.7	9.0%	6.1	10.3	6.9%	4.6	5.9	5.8%	3.8	5.0
20	5.9%	3.6	5.1	8.0%	5.0	6.6	6.8%	4.1	5.4	5.1%	3.1	4.5
21	9.8%	4.7	6.6	13.3%	6.5	8.6	9.9%	4.4	6.6	9.4%	4.5	6.1
22	9.7%	3.7	5.7	13.6%	5.3	7.7	9.5%	3.4	5.7	8.6%	3.1	5.2
23	9.6%	3.9	6.0	12.6%	5.1	7.6	9.6%	3.7	5.7	9.8%	4.3	5.7
24	21.4%	8.9	16.2	28.2%	11.4	18.9	18.0%	6.0	14.8	16.0%	8.7	15.0
ø	90.9%	4.7	6.6	265.2%	7.9	10.3	115.0%	4.6	6.4	124.9%	4.6	6.3

Table 6: Forecasting Errors for February 2011

The large difference between February and September might also give an indication about the temptation some scientists must have for selecting a “good” out-of-sample window. Because the existing literature is about different market, time horizons and sample sizes, among the wide range of models it seems to be difficult to judge which model actually is the superior one.

The comparison of the univariate ARIMA Time Series model and the multivariate models shows a substantial improvement of forecasting performance when more relevant variables are accessible. The MAE could be reduced on average by more than 2 EUR/MWh from the univariate to the multivariate ARIMAX model. Given the value of the transparency data, their inclusion should clearly be considered in future works about forecasting as well.

When regarding the forecasting performance for both periods together, the ARIMAX model performs best, as it performs best in the September period and quite good in the February period. This can be seen in all chosen performance measurements. However, the fact of a changing volatility of Electricity prices is a good fundamental reason to explicitly model it – which is done in the conditional heteroskedasticity model. Given a larger sample size, this model might perform better as it will be easier to find iterative solutions for the maximum likelihood function. Knowing the results of the OLS approach, it is surprising to what extent former works rely only on ARMA(X) and (M)GARCH models to forecast electricity prices, considering the forecasting errors of OLS are quite similar to the other models. In addition, the iterations for OLS are usually done so quickly by a statistical program that it requires much less computation time to estimate a multiple regression.

5.3 Discussion

When considering the results of this work, it is clear that they are within the range of other researchers who were conducting similar kinds of research. Moreover, one needs to keep in mind that the German electricity market has become much more volatile in recent years: While Cuaresma et al (2004) report a standard deviation of 9.50 EUR/MWh for their whole sample ranging between June 2000 to September 2001 before the spikes have been removed, the dataset used in this work exhibits a standard deviation of 15.22 EUR/MWh. This is mainly due to the addition of large quantities of renewable power plants and in the reduction of nuclear power after political decisions. Consequently, given that Cuaresma et al. have chosen the same month for the out-of-sample period, exhibited less volatility and achieved a MAPE of about 13% with a univariate approach, a significant improvement could be reached by using the newly published exogenous variables that have been described in this work since for the September period, the MAPE achieved with the ARIMAX model is 8.6%.

While there is a considerable improvement in the forecasting quality by the use of exogenous variables, there are still a number of aspects that could improve the forecasting performance. In the current work, outliers or spikes have not been considered in a special way. Spikes can be particular interesting when negative prices are observed, which are unlikely to be forecasted with the model employed in this work. Keles et al. (2011) specifically model the possibility of negative prices and thereby reach an improvement in the forecasting precision. A possible solution for outlier treatment could be to have one model for prediction during regular time periods, as the models presented here work fairly precise as long as there are no extreme situations. A second model would be used to detect the possibility of extreme events, for example by including data about grid capacities, congestion issues and river levels.

Timing-Issues have not been examined exhaustively within this data, i.e. the question about what length of calibration period yields the best results as the focus of this work also was to compute reliable estimates for the coefficients of the transparency data. Also, there is the question of the comparability of the in-sample and out-of-sample period: Prediction will work best if the model has been estimated on a comparable time period – but the question remains how this comparable time period can be selected. This topic also is not exhaustively covered in other works. A common standard for the selection of time periods would mean that different works can be compared more easily and the presentation of “good” out-of-sample time windows can be avoided.

This work is focused on using the transparency data and the newly available exogenous variables only on Time Series models, but there is a wide range of different models that have been presented in Chapter 3.1 that could be used with the same data. However, a more complex model doesn’t necessarily improve the forecasting results: As the use of the MGARCH model in the current work has shown, more complex models also need a certain amount of data to produce reliable results.

6 Conclusion

To the author's knowledge, this is the first available work which makes use of the data on the EEX transparency platform for forecasting. Accordingly, it is also the first work which includes exogenous variables like the in-feed of renewables, the unavailabilities of power plants and total generation for the forecasting of the German electricity price.

First, based on energy market characteristics and existing literature, important determinants for the electricity price have been worked out. Then, statistical methods have been discussed that can be used to forecast the day-ahead electricity prices. After the discussion of those models, a selection of those has been used to actually make this prediction. This has been done using the Box-Jenkins Approach with a number of important econometric models.

This work shows that the predictive power increases considerably when including the transparency data that is published by the EEX compared to former works that did not have access to this data. While a similar univariate approach based on the year 2001 yielded a MAPE of 13.2%, the use of the presented variables improved the forecasting error to 8.3%. Other findings of this work are that also for the usage with exogenous variables, a model based on 24 individual time series works better than one time series which includes all consecutive hours because computation time is far less for the former and because hours with high volatilities like the early morning hours do not interfere with other hours. Also, it has been shown that using MAPE as a measurement for the forecasting performance has weaknesses especially in the occurrence of price spikes. As a fourth major point, it can be concluded that OLS has not been used exhaustively in former scientific works while this work shows that OLS does actually not perform worse in forecasting and uses less computation time.

Further research should make use of the transparency data to make judgements about market power and the potential to control prices by withholding capacity, include other important determinants of the electricity price like river levels, transmission capacities and prices or load from neighbouring markets. The remainder of the wide range of available models should also be used on this dataset and some consensus should be found for standardized time horizons. Considering the increase of renewables and the coupling of the different European markets, the forecasting of electricity price will remain an important topic in future and will experience a lot of exciting development

7 References

Aggarwal, S., Saini, L. & Kumar, A., 2009. Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power & Energy Systems*, 31, pp.13–22.

BDEW, 2012. *Erneuerbare Energien und das EEG: Zahlen, Fakten, Grafiken (2011)*, Berlin. Available at: http://www.bdew.de/internet.nsf/id/3564E959A01B9E66C125796B003CFCCE/$file/BDEW%20Energie-Info_EE%20und%20das%20EEG%20%282011%29_23012012.pdf [Accessed June 8, 2012].

Bjerksund, P., Rasmussen, H. & Stensland, G., 2010. Valuation and Risk Management in the Norwegian Electricity Market. In E. Bjørndal et al., eds. *Energy, Natural Resources and Environmental Economics*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 167–185. Available at: http://www.springerlink.com/content/r89025jgr5411t2u/ [Accessed February 1, 2012].

BMWi, 2012. BMWi - Stromversorgung. Available at: http://www.bmwi.de/BMWi/Navigation/Energie/stromversorgung,did=292510.html [Accessed April 5, 2012].

Bourbonnais, R. & Méritet, S., 2008. Electricity Spot Price Modelling: Univariate Time Series Approach. In *The Econometrics of Energy Systems*.

Bundeskartellamt, 2011. *Sektorenuntersuchung Stromerzeugung Stromgroßhandel*, Bonn. Available at: http://www.bundeskartellamt.de/wDeutsch/download/pdf/Stellungnahmen/110113_Bericht_SU_Strom__2_.pdf [Accessed March 15, 2012].

Burger, M., Graeber, B. & Schindlmayr, G., 2007. *Managing Energy Risk*,

Burman, P. & Shumway, R.H., 2006. Generalized Exponential Predictors for Time Series Forecasting. *Journal of the American Statistical Association*, 101(476), pp.1598–1606.

Cuaresma, J.C. et al., 2004. Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy*, 77, pp.87–106.

Cunningham, L.B., Baldick, R. & Baughman, M.L., 2002. An empirical study of applied game theory: transmission constrained Cournot behavior. *IEEE Transactions on Power Systems*, 17(1), pp.166–172.

Dias, A.V., 2010. *Repositório do ISCTE-IUL: Forecasting hourly prices in the portuguese power market with ARIMA models*. Available at: http://repositorio-iul.iscte.pt/handle/10071/2040 [Accessed January 13, 2012].

EEX, 2011. Transparency in Energy Markets - Market Information. Available at: http://www.transparency.eex.com/de/Information/marktinformationen/newsdetails/news/Stilllegung_von_Kernkraftwerken [Accessed March 15, 2012].

Enders, W., 2010. *Applied econometric time series*, Hoboken, N.J.: Wiley.

EPEX SPOT, 2011a. EPEX SPOT SE: Day-ahead auction. Available at: http://www.epexspot.com/en/product-info/auction/germany-austria [Accessed February 6, 2012].

EPEX SPOT, 2011b. EPEX SPOT SE: List of Members. Available at: http://www.epexspot.com/en/membership/list_of_members/participants/trading [Accessed February 6, 2012].

Garcia, R.C. et al., 2005. A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*, 20(2), pp.867– 874.

Gaten, T., 2000. Z-tests and t tests. Available at: http://www.le.ac.uk/bl/gat/virtualfc/Stats/ttest.html [Accessed April 11, 2012].

Gianfreda, A. & Grossi, L., 2011. *Forecasting Italian Electricity Zonal Prices with Exogenous Variables*, Università di Verona, Dipartimento di Scienze economiche. Available at: http://econpapers.repec.org/paper/verwpaper/01_2f2011.htm [Accessed January 13, 2012].

Golder, M., 2007. Time Series. Available at: https://files.nyu.edu/mrg217/public/timeseries.pdf [Accessed April 2, 2012].

Haili Song, Chen-Ching Liu & Lawarree, J., 2002. Nash equilibrium bidding strategies in a bilateral electricity market. *IEEE Transactions on Power Systems*, 17(1), pp.73–79.

Hipòlit Torró, 2007. *Forecasting Weekly Electricity Prices at Nord Pool*, Fondazione Eni Enrico Mattei. Available at: http://ideas.repec.org/p/fem/femwpa/2007.88.html [Accessed January 13, 2012].

Huurman, C., Ravazzolo, F. & Zhou, C., 2010. The power of weather. *Computational Statistics & Data Analysis*, (0). Available at: http://www.sciencedirect.com/science/article/pii/S0167947310002665 [Accessed January 13, 2012].

Johannes, V., 2011. Risk premiums in the German day-ahead Electricity Market. *Energy Policy*, 39(1), pp.386–394.

Keles, D. et al., 2011. Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. *Energy Economics*. Available at: http://www.sciencedirect.com/science/article/pii/S0140988311001721 [Accessed January 13, 2012].

Liebl, D., 2010. Modeling hourly Electricity Spot Market Prices as non stationary functional times series. Available at: http://mpra.ub.uni-muenchen.de/25017/ [Accessed October 25, 2011].

Lise, W., Hobbs, B.F. & Hers, S., 2008. Market power in the European electricity market—The impacts of dry weather and additional transmission capacity. *Energy Policy*, 36(4), pp.1331–1343.

McDermott, G.R. & Nilsen, Ø.A., 2011. Electricity Prices, River Temperatures and Cooling Water Scarcity. Available at: http://swopec.hhs.se/nhheco/abs/nhheco2011_018.htm [Accessed January 12, 2012].

Morthorst, P.E., 2007. *Impacts of Wind Power on Power Spot Prices*, Available at: http://www.optres.fhg.de/events/workshop-2006-10-12/Copenhagen/Morthorst%20Cph%281206%29.pdf [Accessed March 16, 2012].

Möst, D. & Genoese, M., 2009. Market power in the German wholesale electricity market. *The Journal of Energy Markets*, Volume 2 / Number 2, pp.47–74.

Müsgens, F., 2004. Market Power in the German Wholesale Electricity Market by Market Power in the German Wholesale Electricity Market. *Fuel*, 2(May), pp.47–74.

Neubarth, J. et al., 2006. Beeinflussung der Spotmarktpreise durch Windstrom-erzeugung. *ET. Energiewirtschaftliche Tagesfragen*, 56(7), pp.42–45.

Nogales, F.J. et al., 2002. Forecasting next-day electricity prices by time series models. *IEEE Transactions on Power Systems*, 17(2), pp.342–348.

Ockenfels, A., Grimm, V. & Zoettl, G., 2008. *The Pricing Mechanism of the Day Ahead Electricity Spot Market Auction on the EEX*, Available at: http://cdn.eex.com/document/38615/gutachten_eex_ockenfels_e.pdf [Accessed February 27, 2012].

Redl, C. et al., 2009. Price formation in electricity forward markets and the relevance of systematic forecast errors. *Energy Economics*, 31, pp.356–364.

RTE, 2011. *Electrical Energy Statistics for France 2010*, Available at: http://www.rte-france.com/uploads/Mediatheque_docs/vie_systeme/annuelles/Statistiques_energie_electrique/an/statistiques_annuelles_2010_an.pdf [Accessed April 5, 2012].

Sensfuß, F., Ragwitz, M. & Genoese, M., 2008. The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. *Energy Policy*, 36(8), pp.3086–3094.

Skrivarhaug, A.V., 2010. Energy Markets & Price Analysis in Practice, Presentation during Lecture of the Course ENE421 at 07.10.2010 at NHH, Bergen, Norway.

Stern, J.P., 2007. Is there a rationale for the continuing link to oil product prices in continental European long-term gas contracts? *International Journal of Energy Sector Management*, 1(3), pp.221–239.

Swider, D.J., 2006. Erweiterte ARMA-Ansätze zur Prognose von Spotmarktpreisen in Europa. *Zeitschrift für Energiewirtschaft*, 30(1), pp.31–42.

Swider, D.J. & Weber, C., 2007. Extended ARMA models for estimating price developments on day-ahead electricity markets. *Electric Power Systems Research*, 77(5–6), pp.583–593.

Tarjei, K., 2011. Power Trading Analytics and Forecasting in Germany. *The Electricity Journal*, 24(8), pp.41–55.

Tim Christensen, Stan Hurn & Ken Lindsay, 2011. *Forecasting Spikes in Electricity Prices*, National Centre for Econometric Research. Available at: http://ideas.repec.org/p/qut/auncer/2011_1.html [Accessed October 25, 2011].

Trueck, S., Weron, Rafal & Wolff, R., 2007. *Outlier Treatment and Robust Approaches for Modeling Electricity Spot Prices*, University Library of Munich, Germany. Available at: http://ideas.repec.org/p/pra/mprapa/4711.html [Accessed January 21, 2012].

Vassilopoulos, P., 2003. *Models for the Identification of Market Power in Wholesale Electricity Markets*,

Verbeek, M., 2008. *A Guide To Modern Econometrics*, John Wiley & Sons.

Wei, W.W.S., 1990. *Time series analysis: univariate and multivariate methods*, Addison-Wesley Pub.

Weron, Rafał & Misiorek, A., 2006. Short-term Electricity Price Forecasting with Time Series Models: A Review and Evaluation. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.129.6648.

Weron, Rafał & Misiorek, A., 2008. Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4), pp.744–763.

White, H., 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4), pp.817–838.

Wooldridge, J.M., 2008. *Introductory Econometrics* International ed of 4th revised ed., Cengage Learning Emea.

8 Appendix

Appendix 1: Detailed regression results for Hour 13 for the various models

Detailed ARIMAX regression results for Hour 13:

```
ARIMA regression

Sample:  3 - 699                               Number of obs      =       697
                                               Wald chi2(17)      =  11923.43
Log pseudolikelihood = -2157.617               Prob > chi2        =    0.0000
```

Spotprice	Coef.	Semirobust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Spotprice						
log_PlannedGenerationQuantity	34.18101	5.415336	6.31	0.000	23.56715	44.79487
log_windplangermany	-2.554196	.285294	-8.95	0.000	-3.113362	-1.99503
log_emissionprice						
L1.	12.39411	7.356297	1.68	0.092	-2.023968	26.81219
log_CrudeOilBrent						
L1.	13.73996	14.01328	0.98	0.327	-13.72557	41.20548
log_temperature	-4.233778	2.807919	-1.51	0.132	-9.737199	1.269643
log_unav_lignite_planned	.2499463	.1757632	1.42	0.155	-.0945432	.5944359
log_unav_Gas_planned	-.2997165	.236123	-1.27	0.204	-.762509	.163076
log_unav_coal_planned	.4933421	.183229	2.69	0.007	.1342198	.8524643
log_unav_uranium_planned	.1125536	.2087093	0.54	0.590	-.2965092	.5216164
log_unav_lignite_nonsched	.0974945	.1763535	0.55	0.580	-.248152	.4431411
log_unav_Gas_nonsched	-.0455018	.1076619	-0.42	0.673	-.2565154	.1655117
log_unav_coal_nonsched	.5279782	.1811315	2.91	0.004	.172967	.8829894
log_unav_uranium_nonsched	.4153203	.1673355	2.48	0.013	.0873488	.7432918
d_weekend	-4.365256	1.366925	-3.19	0.001	-7.044379	-1.686133
d_holiday	-5.75724	1.487348	-3.87	0.000	-8.672388	-2.842092
_cons	-380.4849	71.77356	-5.30	0.000	-521.1585	-239.8113
ARMA						
ar						
L1.	.9644741	.0409611	23.55	0.000	.8841918	1.044756
ma						
L1.	-.7507101	.0714538	-10.51	0.000	-.890757	-.6106632
/sigma	5.233719	.2186925	23.93	0.000	4.805089	5.662348

Detailed ARIMA regression results for Hour 13:

```
ARIMA regression

Sample:  2 - 699                                 Number of obs      =       698
                                                 Wald chi2(2)       =   2058.10
Log pseudolikelihood = -2519.935                 Prob > chi2        =    0.0000
```

D.Spotprice	Coef.	Semirobust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Spotprice						
_cons	.0238104	.0239817	0.99	0.321	-.0231929	.0708138
ARMA						
ar						
L1.	.3199192	.064986	4.92	0.000	.192549	.4472893
ma						
L1.	-1.048832	.0343039	-30.57	0.000	-1.116066	-.9815973
/sigma	8.333107	.45808	18.19	0.000	7.435286	9.230927

Detailed MGARCH regression results for Hour 13

ARCH family regression -- ARMA disturbances

Sample: 3 - 699	Number of obs	=	697
Distribution: Gaussian	Wald chi2(17)	=	6146.84
Log likelihood = -2132.537	Prob > chi2	=	0.0000

Spotprice	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
Spotprice						
log_PlannedGenerationQuantity	33.80305	2.603696	12.98	0.000	28.6999	38.9062
log_windplangermany	-2.524685	.2355351	-10.72	0.000	-2.986326	-2.063045
log_emissionprice						
L1.	14.44	7.978441	1.81	0.070	-1.197458	30.07746
log_CrudeOilBrent						
L1.	16.9234	5.612953	3.02	0.003	5.92221	27.92458
log_temperature	-3.787267	1.71237	-2.21	0.027	-7.14345	-.4310836
log_unav_lignite_planned	.2389516	.1683776	1.42	0.156	-.0910625	.5689657
log_unav_Gas_planned	-.2066319	.1742228	-1.19	0.236	-.5481023	.1348384
log_unav_coal_planned	.5088938	.1738502	2.93	0.003	.1681538	.8496339
log_unav_uranium_planned	.0847013	.1264888	0.67	0.503	-.1632122	.3326149
log_unav_lignite_nonsched	.0512222	.1534323	0.33	0.738	-.2494996	.3519439
log_unav_Gas_nonsched	-.0301866	.1027026	-0.29	0.769	-.23148	.1711068
log_unav_coal_nonsched	.3673975	.1793762	2.05	0.041	.0158267	.7189684
log_unav_uranium_nonsched	.3992193	.1224033	3.26	0.001	.1593133	.6391254
d_weekend	-4.624461	.7208901	-6.41	0.000	-6.037379	-3.211542
d_holiday	-5.296949	.9955526	-5.32	0.000	-7.248196	-3.345702
_cons	-398.7819	49.05024	-8.13	0.000	-494.9186	-302.6452
ARMA						
ar						
L1.	.9600261	.0158963	60.39	0.000	.92887	.9911822
ma						
L1.	-.7103772	.0387339	-18.34	0.000	-.7862941	-.6344602
ARCH						
arch						
L1.	.2211826	.0473925	4.67	0.000	.128295	.3140701
_cons	21.8575	1.328156	16.46	0.000	19.25437	24.46064

Detailed OLS regression results for Hour 13

```
Linear regression                                      Number of obs =     697
                                                       F( 16,   680) =   85.58
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6859
                                                       Root MSE      =  5.9727
```

Spotprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Spotprice						
L1.	.2123977	.0308277	6.89	0.000	.1518689	.2729265
log_PlannedGenerationQuantity	19.63351	3.586481	5.47	0.000	12.5916	26.67541
log_windplangermany	-2.775292	.2621152	-10.59	0.000	-3.289945	-2.26064
log_emissionprice						
L1.	9.581585	2.400833	3.99	0.000	4.867649	14.29552
log_CrudeOilBrent						
L1.	12.7667	1.772105	7.20	0.000	9.287245	16.24615
log_temperature	-3.744318	1.468842	-2.55	0.011	-6.628329	-.8603062
log_unav_lignite_planned	.5076965	.143979	3.53	0.000	.2249997	.7903933
log_unav_Gas_planned	-.2869462	.1677846	-1.71	0.088	-.6163844	.0424919
log_unav_coal_planned	.5588077	.1570458	3.56	0.000	.2504549	.8671606
log_unav_uranium_planned	.2766992	.0939863	2.94	0.003	.092161	.4612374
log_unav_lignite_nonsched	-.0537578	.160686	-0.33	0.738	-.3692581	.2617425
log_unav_Gas_nonsched	.0343818	.0955414	0.36	0.719	-.1532098	.2219735
log_unav_coal_nonsched	.5079737	.1537535	3.30	0.001	.2060851	.8098623
log_unav_uranium_nonsched	.3967857	.0896811	4.42	0.000	.2207006	.5728707
d_weekend	-7.525959	.9251198	-8.14	0.000	-9.342393	-5.709524
d_holiday	-6.068781	1.809998	-3.35	0.001	-9.622639	-2.514924
_cons	-225.689	40.14157	-5.62	0.000	-304.5053	-146.8727

Appendix 2: combined regression results of all hours for the various models

Hour 1:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	37.80228***		24.60817***	20.916045***
log_windplangermany	-3.6566649***		-3.0279332***	-4.1276485***
L1.log_emissionprice	17.75903900		12.661732***	12.321046***
L1.log_CrudeOilBrent	0.11464695		22.445524***	21.115508***
log_temperature	-6.3437269***		-5.108219**	-4.4434372**
log_unav_lignite_planned	0.16834868		.72492924***	.52885935**
log_unav_Gas_planned	-0.20518605		-0.26056873	-0.22210832
log_unav_coal_planned	-0.22786173		.48305174***	-0.21456920
log_unav_uranium_planned	0.22725293		.35978447***	.54277278***
log_unav_lignite_nonsched	0.82739322		-.32907417*	0.42810613
log_unav_Gas_nonsched	0.00582665		0.04533547	0.00745567
log_unav_coal_nonsched	0.27371378		-0.05327172	0.28957681
log_unav_uranium_nonsched	.41787369**		.39189076***	.54411888***
d_weekend	2.2950937**		.90145721**	0.75365787
d_holiday	-4.59595190		2.2160996***	-5.02879310
_cons	-356.90043***	0.02112581	-322.41129***	-278.3315***
ARMA				
L1. ar	.99654808***	0.28446055	.77775005***	0.15434763
L1.ma	-.90645213***	-.95431379***	-.36924355***	
sigma				
_cons	7.4538292***	9.015391***		
ARCH				
L1.arch			1.2691595***	
_cons			14.457272***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 2:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	51.268699***		38.624666***	22.977551***
log_windplangermany	-4.3530186***		-3.5321575***	-4.993511***
L1.log_emissionprice	15.93829900		3.61206150	10.931225***
L1.log_CrudeOilBrent	0.53715505		28.554133***	22.444739***
log_temperature	-5.3154141**		-1.99926910	-4.0162398*
log_unav_lignite_planned	0.11387944		.47747303**	.57618282**
log_unav_Gas_planned	-0.54276517		0.09009744	-0.26500555
log_unav_coal_planned	-0.12116890		.59944476***	-0.16610405
log_unav_uranium_planned	0.20057586		.31874499*	.56903945***
log_unav_lignite_nonsched	0.79296257		-.58856378**	0.34925190
log_unav_Gas_nonsched	0.00418757		-0.02645743	0.04970529
log_unav_coal_nonsched	0.28917394		-0.00311807	0.30748571
log_unav_uranium_nonsched	.35089265*		-0.02595310	.49365531***
d_weekend	2.167644**		0.02096275	-0.58144236
d_holiday	-5.47942760		2.8626649***	-6.40672130
_cons	-494.89983***	0.02409850	-480.66562***	-301.71745***
ARMA				
L1. ar	.99553433***	.36761154**	.920006***	.22310482**
L1.ma	-.86671126***	-.96365859***	-.63962275***	
sigma				
_cons	7.7701201***	9.7861842***		
ARCH				
L1.arch			.91978579***	
_cons			20.128158***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 3:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	36.927526***		45.568891***	22.387963***
log_windplangermany	-6.1356262***		-3.9078869***	-5.7550826***
L1.log_emissionprice	11.57011*		11.786076*	9.3120262**
L1.log_CrudeOilBrent	33.793153***		36.524938***	23.399087***
log_temperature	-5.71325020		-3.15291300	-6.9702452***
log_unav_lignite_planned	0.40772919		.32246663*	.65360481*
log_unav_Gas_planned	-0.30314538		0.06799320	-0.18182105
log_unav_coal_planned	-0.15597248		0.25185147	-0.30149646
log_unav_uranium_planned	1.0747426***		.53428318***	.71548734***
log_unav_lignite_nonsched	0.06328409		.49969187***	0.34661356
log_unav_Gas_nonsched	0.14038775		-0.11897479	0.06384351
log_unav_coal_nonsched	0.33230188		0.05319753	0.24121525
log_unav_uranium_nonsched	.60111705***		0.04337756	.42469594**
d_weekend	0.52561005		1.0335905**	-1.44529000
d_holiday	-4.97532620		1.21701100	-6.16794300
_cons	-483.05997***	0.02509617	-611.90496***	-283.83727***
ARMA				
L1. ar	.38250675***	.4521076***	.92376494***	.28664573***
L1.ma	-0.00000146	-.96975869***	-.70497316***	
sigma				
_cons	8.4847182***	10.670406***		
ARCH				
L1.arch			1.1630967***	
_cons			18.742712***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 4:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	63.674478***		50.774207***	22.197633***
log_windplangermany	-5.4987738***		-4.3764272***	-6.1408358***
L1.log_emissionprice	10.36984800		-6.86870290	8.2791905*
L1.log_CrudeOilBrent	14.83263600		39.427998***	25.873117***
log_temperature	-11.237169***		-2.96436570	-9.8476668***
log_unav_lignite_planned	0.04952267		-.65129285**	.69355104*
log_unav_Gas_planned	-0.21293191		0.07868574	0.07316870
log_unav_coal_planned	-0.07065987		.7251534***	-0.28344164
log_unav_uranium_planned	.84121729*		-.70581962***	.96150825***
log_unav_lignite_nonsched	0.72766808		-0.14493857	0.41872205
log_unav_Gas_nonsched	0.02523161		0.09826230	0.07178556
log_unav_coal_nonsched	0.46905336		0.19820639	0.40267699
log_unav_uranium_nonsched	0.04954287		.27003663*	0.36511810
d_weekend	2.7791196**		2.1627335***	-1.49495820
d_holiday	-3.37227290		3.139345**	-5.29553180
_cons	-663.02891***	0.02973405	-615.23589***	-284.14341***
ARMA				
L1. ar	.99046159***	.43798596**	.93026803***	.28300698*
L1.ma	-.83092656***	-.96837735***	-.50940756***	
sigma				
_cons	9.971369***	12.080276***		
ARCH				
L1.arch			.92161441***	
_cons			25.102053***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 5:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	63.758967***		64.306959***	22.60092***
log_windplangermany	-4.5757852***		-3.713386***	-5.1001707***
L1.log_emissionprice	9.17774730		-15.02923900	11.680514***
L1.log_CrudeOilBrent	13.92112700		36.790448***	23.018528***
log_temperature	-7.8567158**		-3.80147810	-6.9110599***
log_unav_lignite_planned	0.25271863		-0.10200086	.59754605**
log_unav_Gas_planned	-0.22165543		0.00453733	0.04227868
log_unav_coal_planned	-0.12209879		.46902251**	-0.20286231
log_unav_uranium_planned	.51572046*		-0.10038672	.72155776***
log_unav_lignite_nonsched	0.84378690		-0.20186775	0.59890036
log_unav_Gas_nonsched	0.01715034		0.08456934	0.12734362
log_unav_coal_nonsched	0.38771838		0.21913979	0.28257456
log_unav_uranium_nonsched	0.13215262		0.06171226	.46320927***
d_weekend	0.97591389		1.4083245*	-3.505816***
d_holiday	-2.36607150		4.223811***	-4.75181000
_cons	-674.54611***	0.02513085	-736.02659***	-301.61527***
ARMA				
L1. ar	.99089033***	.45468125***	.96447902***	.29310181***
L1.ma	-.79142144***	-.96154129***	-.63872163***	
sigma				
_cons	7.5936844***	9.9838979***		
ARCH				
L1.arch			.7508098***	
_cons			22.903099***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 6:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	31.890981***		25.758191***	19.13872***
log_windplangermany	-4.0039117***		-3.5282224***	-3.7123873***
L1.log_emissionprice	14.972124***		13.237325***	11.359272***
L1.log_CrudeOilBrent	31.457239***		30.23428***	21.933093***
log_temperature	-3.65495550		-5.293696**	-5.3262337***
log_unav_lignite_planned	.46358935**		0.43380091	.50049971***
log_unav_Gas_planned	-0.22188133		0.02967436	-0.06278090
log_unav_coal_planned	-0.09193868		0.24270836	-0.15631688
log_unav_uranium_planned	.79316077***		.48603893***	.56862702***
log_unav_lignite_nonsched	0.07959116		-.93072354***	0.29547389
log_unav_Gas_nonsched	0.00929201		0.06231792	-0.06244881
log_unav_coal_nonsched	.3725523*		0.35461165	0.30927164
log_unav_uranium_nonsched	.60770705***		.39834759***	.43597048***
d_weekend	-6.3273826***		-6.7400165***	-8.3256056***
d_holiday	-7.0830099*		-1.08401700	-8.24396710
_cons	-449.17731***	0.02407188	-366.7947***	-267.19914***
ARMA				
L1. ar	.3347955***	.41153775***	.44531337***	.28818207***
L1.ma	-0.00000556	-1.0327318***	-0.13246604	
sigma				
_cons	7.6613091***	9.9025245***		
ARCH				
L1.arch			.61299041***	
_cons			25.352831***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 7

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	30.981299***		25.843225***	22.728432***
log_windplangermany	-3.637164***		-2.6725533***	-3.4087475***
L1.log_emissionprice	16.598153***		8.0224427*	14.06104***
L1.log_CrudeOilBrent	26.775754***		26.603745***	22.334494***
log_temperature	-0.06519959		-5.1769029***	-2.62972580
log_unav_lignite_planned	.5534098*		0.27822873	.46006485*
log_unav_Gas_planned	0.04999313		.77484682***	0.28634938
log_unav_coal_planned	-0.26759541		0.16184572	-0.19955485
log_unav_uranium_planned	.8545118***		.69836513***	.79444764***
log_unav_lignite_nonsched	0.72263232		0.00698214	0.74231240
log_unav_Gas_nonsched	-0.20380398		-0.07815967	-0.30217785
log_unav_coal_nonsched	0.24502345		.65290839***	0.21562244
log_unav_uranium_nonsched	.65220581***		.27471308**	.50521532***
d_weekend	-15.849297***		-14.084379***	-16.912671***
d_holiday	-15.644461**		-4.9500317***	-17.659031*
_cons	-434.91486***	.03092117*	-349.20147***	-320.12544***
ARMA				
L1. ar	.22179916*	.33282536**	.91434271***	.17394995*
L1.ma	-0.00002098	-.98413693***	-.84449164***	
sigma				
_cons	10.222899***	15.171687***		
ARCH				
L1.arch			.77739932***	
_cons			28.887622***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 8:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	27.437818**		26.819609***	26.327475***
log_windplangermany	-4.0968684***		-4.2105717***	-3.7487765***
L1.log_emissionprice	18.32708500		17.61682900	14.951266***
L1.log_CrudeOilBrent	21.26459200		27.586021**	23.103796***
log_temperature	-3.99073210		-3.90410190	-3.35162570
log_unav_lignite_planned	-0.23648321		-0.02218590	0.38122321
log_unav_Gas_planned	-1.1684211*		-1.2107645**	-.53466005*
log_unav_coal_planned	-0.69038883		-0.72534209	-0.34671691
log_unav_uranium_planned	0.20952151		0.33336862	.60865465**
log_unav_lignite_nonsched	0.95199959		.89797834**	0.85406308
log_unav_Gas_nonsched	-0.19518237		-0.23478274	-0.17009279
log_unav_coal_nonsched	0.72567639		0.70686218	0.33729520
log_unav_uranium_nonsched	.67808041**		.63002404*	.64062792***
d_weekend	-17.980773***		-18.300898***	-19.157447***
d_holiday	-17.836673*		-17.659959***	-18.464179*
_cons	-335.67291*	.02956952***	-356.82328***	-344.62383***
ARMA				
L1. ar	.97451633***	.29361451***	.96404924***	.11632081***
L1.ma	-.88426866***	-.99985374***	-.86832256***	
sigma				
_cons	10.611107***	17.350794***		
ARCH				
L1.arch			0.00153706	
_cons			112.5788***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 9:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	33.325601***		32.925618***	25.809265***
log_windplangermany	-3.480623***		-2.7286204***	-3.4932607***
L1.log_emissionprice	17.609448*		6.69889300	15.715741***
L1.log_CrudeOilBrent	22.86275500		20.842713**	19.794732***
log_temperature	-5.91883270		-9.916298***	-5.6712351**
log_unav_lignite_planned	0.43157606		.7220324***	.85242637***
log_unav_Gas_planned	-1.1151446*		-.92980309***	-0.42035603
log_unav_coal_planned	-0.45165769		-0.21762200	-0.24951508
log_unav_uranium_planned	0.23802905		-0.04602932	.58023897***
log_unav_lignite_nonsched	0.52010453		-0.30867947	0.41675378
log_unav_Gas_nonsched	-0.05230753		0.08701385	0.00458863
log_unav_coal_nonsched	0.49289164		-0.06015702	0.18448234
log_unav_uranium_nonsched	.62144292**		.55740902***	.60790365***
d_weekend	-12.682528***		-12.663772***	-15.365571***
d_holiday	-14.03927**		-5.0207615***	-14.973846**
_cons	-404.84011***	.0270713*	-349.86478***	-322.28816***
ARMA				
L1. ar	.96608567***	.32058857***	.9800337***	.14801179***
L1.ma	-.84381875***	-.98689119***	-.83810203***	
sigma				
_cons	8.2591872***	14.506515***		
ARCH				
L1.arch			.71909549***	
_cons			30.100324***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 10:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	34.697186***		36.603719***	24.702462***
log_windplangermany	-2.8896977***		-2.7288763***	-2.9277702***
L1.log_emissionprice	14.198691*		10.71038500	12.309271***
L1.log_CrudeOilBrent	20.536988***		19.585462***	16.087858***
log_temperature	-4.04309450		-1.57285560	-5.7430137***
log_unav_lignite_planned	.52624091*		.82842401***	.82069672***
log_unav_Gas_planned	-.67779612*		0.00141164	-0.31432167
log_unav_coal_planned	-0.07992700		0.12350853	0.03544408
log_unav_uranium_planned	0.26213942		0.16813236	.49916876***
log_unav_lignite_nonsched	0.37047942		0.13958606	0.31540242
log_unav_Gas_nonsched	-0.03947824		-0.05533374	0.03502365
log_unav_coal_nonsched	.64770906*		.7457873***	0.41007868
log_unav_uranium_nonsched	.56754123***		.27156826*	.51339833***
d_weekend	-8.1664451***		-8.5766926***	-10.97828***
d_holiday	-8.9102773***		-5.9301091***	-9.6915023***
_cons	-418.16812***	0.02532470	-441.27301***	-293.50371***
ARMA				
L1. ar	.9409465***	.32016592***	.94435148***	.17516049***
L1.ma	-.76299542***	-1.0246643***	-.72774551***	
sigma				
_cons	6.1858224***	10.964992***		
ARCH				
L1.arch			.44475248***	
_cons			25.757948***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 11:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	35.177559***		36.262636***	21.989716***
log_windplangermany	-2.5167581***		-2.4814507***	-2.7269989***
L1.log_emissionprice	14.384388*		14.79252300	10.775511***
L1.log_CrudeOilBrent	18.401183**		16.387599**	13.115677***
log_temperature	-4.88143430		-5.529277**	-5.7583534***
log_unav_lignite_planned	0.46565688		.54593863**	.66805253***
log_unav_Gas_planned	-0.36645697		-0.24247199	-0.21735131
log_unav_coal_planned	0.17027138		0.08605519	0.30125174
log_unav_uranium_planned	0.30815815		0.21829180	.42611543***
log_unav_lignite_nonsched	0.30109992		0.34417986	0.19108581
log_unav_Gas_nonsched	-0.00975044		-0.01506676	0.09738297
log_unav_coal_nonsched	.60052802*		.7008965***	.50499653*
log_unav_uranium_nonsched	.4423993*		.3524047*	.43509905***
d_weekend	-6.1688315***		-6.0792513***	-9.4464456***
d_holiday	-7.4464026***		-7.4112522***	-8.1870527***
_cons	-417.04394***	0.02360769	-420.37944***	-250.8986***
ARMA				
L1. ar	.94263265***	.33077111***	.95787339***	.20560212***
L1.ma	-.73425562***	-1.0255604***	-.72316567***	
sigma				
_cons	5.6739135***	9.7205983***		
ARCH				
L1.arch			.18745459***	
_cons			27.09925***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 12:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	33.599894***		34.749486***	18.125378***
log_windplangermany	-2.596942***		-2.6172395***	-2.8226178***
L1.log_emissionprice	15.34800600		13.66038000	11.194814***
L1.log_CrudeOilBrent	13.90772800		11.57915500	10.448787***
log_temperature	-5.76333870		-6.7993139***	-5.1450751**
log_unav_lignite_planned	0.32827067		.49211397**	.58565519***
log_unav_Gas_planned	-0.29841231		-0.16107622	-0.26998913
log_unav_coal_planned	.45819136*		.48832315*	.55264215**
log_unav_uranium_planned	0.06754711		0.13041624	.3415591***
log_unav_lignite_nonsched	0.29777397		0.26043201	0.09092553
log_unav_Gas_nonsched	0.03533158		0.05737000	0.13932431
log_unav_coal_nonsched	.63081801**		.59137513***	.58855116**
log_unav_uranium_nonsched	.41095864*		.31860946*	.38344973***
d_weekend	-6.0124916***		-6.0957779***	-9.4665231***
d_holiday	-5.3516249***		-5.3868685***	-6.1490245**
_cons	-377.85158***	0.02363891	-374.87096***	-199.91891***
ARMA				
L1. ar	.95937327***	.32138863***	.96408188***	.22210209***
L1.ma	-.73440745***	-1.0487175***	-.71060183***	
sigma				
_cons	5.7175239***	9.1149445***		
ARCH				
L1.arch			.21481738***	
_cons			26.726738***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 13:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	34.18101***		33.803052***	19.633506***
log_windplangermany	-2.5541961***		-2.5246853***	-2.7752924***
L1.log_emissionprice	12.39411000		14.43999900	9.5815852***
L1.log_CrudeOilBrent	13.73995600		16.923396**	12.7667***
log_temperature	-4.23377800		-3.787267*	-3.7443177*
log_unav_lignite_planned	0.24994634		0.23895162	.5076965***
log_unav_Gas_planned	-0.29971647		-0.20663191	-0.28694623
log_unav_coal_planned	.49334207**		.50889382**	.55880774***
log_unav_uranium_planned	0.11255363		0.08470135	.27669921**
log_unav_lignite_nonsched	0.09749453		0.05122216	-0.05375778
log_unav_Gas_nonsched	-0.04550184		-0.03018660	0.03438182
log_unav_coal_nonsched	.52797819**		.36739753*	.50797372**
log_unav_uranium_nonsched	.4153203*		.39921933**	.39678567***
d_weekend	-4.3652562**		-4.6244609***	-7.5259589***
d_holiday	-5.7572398***		-5.2969489***	-6.0687815***
_cons	-380.48493***	0.02381044	-398.78192***	-225.68903***
ARMA				
L1. ar	.96447413***	.31991917***	.96002614***	.21239769***
L1.ma	-.75071009***	-1.0488318***	-.71037717***	
sigma				
_cons	5.2337186***	8.3331067***		
ARCH				
L1.arch			.22118258***	
_cons			21.857504***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 14:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	41.271177***		42.917016***	23.064352***
log_windplangermany	-2.3229732***		-2.3382799***	-2.8489351***
L1.log_emissionprice	16.34489*		15.493621*	9.0338762***
L1.log_CrudeOilBrent	-3.90753400		4.62742440	14.71026***
log_temperature	-4.5010305*		-4.3806625**	-2.9320166*
log_unav_lignite_planned	0.18370317		0.18353848	.452397**
log_unav_Gas_planned	-0.38452104		-0.27219880	-.3927713*
log_unav_coal_planned	.34241262*		.40748741*	.64250644***
log_unav_uranium_planned	0.23853290		0.21371025	.35103002***
log_unav_lignite_nonsched	0.08659657		0.07459799	-0.09893779
log_unav_Gas_nonsched	-0.04983668		-0.06058410	0.04677774
log_unav_coal_nonsched	.46518167*		.43365839**	.41363993*
log_unav_uranium_nonsched	0.23709350		0.15565349	.40937585***
d_weekend	-5.3179597***		-5.505127***	-9.6124266***
d_holiday	-5.9596348***		-5.1458196***	-7.0663897***
_cons	-384.30593***	0.02184959	-448.88625***	-269.5503***
ARMA				
L1. ar	.99319155***	.34820818***	.9900392***	.15851905***
L1.ma	-.82392967***	-.98002474***	-.79834258***	
sigma				
_cons	5.2828821***	9.7998416***		
ARCH				
L1.arch			.24900033***	
_cons			21.989072***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 15:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	44.284846***		45.514854***	23.454875***
log_windplangermany	-2.3633724***		-2.3893342***	-3.033221***
L1.log_emissionprice	13.92668200		12.19222400	9.1469797***
L1.log_CrudeOilBrent	-3.79139210		5.80810470	15.045163***
log_temperature	-5.0312937*		-4.6514866**	-1.54989260
log_unav_lignite_planned	0.13235115		0.18533498	.40536085**
log_unav_Gas_planned	-.52608194*		-.46551518**	-.54595947**
log_unav_coal_planned	0.21074055		0.23169781	.60473567***
log_unav_uranium_planned	0.08635531		0.10337648	.36083118***
log_unav_lignite_nonsched	0.17484839		0.16263812	-0.01316782
log_unav_Gas_nonsched	-0.00153649		0.00183183	0.05328312
log_unav_coal_nonsched	.60316101**		.4171801*	.54038999***
log_unav_uranium_nonsched	0.20404042		0.11391070	.43734335***
d_weekend	-5.1366576***		-5.515561***	-10.078288***
d_holiday	-6.5200476***		-5.1128798***	-8.0185987***
_cons	-407.32381***	.02278912*	-470.53853***	-280.63057***
ARMA				
L1. ar	.99466561***	.37654034***	.98920023***	.16629552***
L1.ma	-.82545149***	-.98488751***	-.79275879***	
sigma				
_cons	5.3750809***	10.185259***		
ARCH				
L1.arch			.2439576***	
_cons			23.065092***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 16:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	47.68508***		48.162981***	25.170207***
log_windplangermany	-2.1314987***		-2.1938913***	-2.8912969***
L1.log_emissionprice	13.31817800		12.31161500	6.9881852**
L1.log_CrudeOilBrent	-8.07006900		-2.05565640	15.081602***
log_temperature	-5.2870359**		-4.7081323**	-1.26191740
log_unav_lignite_planned	0.05572936		0.16435615	.3162176*
log_unav_Gas_planned	-.55721187*		-.4030958*	-.72261355***
log_unav_coal_planned	0.11134154		0.09657613	.56906403***
log_unav_uranium_planned	0.06006284		0.03093016	.37394823***
log_unav_lignite_nonsched	0.30067901		0.21786109	0.02897763
log_unav_Gas_nonsched	-0.06889492		-0.08779958	0.01355622
log_unav_coal_nonsched	.52857787*		.43824146**	.47683974**
log_unav_uranium_nonsched	0.22769303		0.14231329	.5171144***
d_weekend	-3.671807**		-4.3357836***	-8.9267637***
d_holiday	-5.1266488**		-4.1523223***	-6.5030635**
_cons	-420.15931***	0.02151146	-469.04309***	-295.97812***
ARMA				
L1. ar	.99639641***	.37942466***	.99454385***	.1833057***
L1.ma	-.83506007***	-.97688883***	-.81179358***	
sigma				
_cons	5.4163213***	10.02995***		
ARCH				
L1.arch			.25893165***	
_cons			23.477231***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 17:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	40.638738***		41.421586***	23.434908***
log_windplangermany	-2.4522639***		-2.4882927***	-2.8664892***
L1.log_emissionprice	13.42008400		12.31276900	6.1671559*
L1.log_CrudeOilBrent	-5.64370880		4.60219560	15.507098***
log_temperature	-5.1524258**		-4.592941*	-1.72040910
log_unav_lignite_planned	0.05775033		0.17649329	0.17011721
log_unav_Gas_planned	-.7192883**		-.59492771***	-1.2195777***
log_unav_coal_planned	-0.00676847		0.01372686	.55022923***
log_unav_uranium_planned	0.15103815		0.15574275	0.14315786
log_unav_lignite_nonsched	.50251888*		.46459724**	0.11837968
log_unav_Gas_nonsched	0.13670987		0.10534177	0.11595972
log_unav_coal_nonsched	0.17312209		0.14921940	0.21974886
log_unav_uranium_nonsched	0.07917150		0.05550424	.41459611***
d_weekend	-4.1724104***		-4.403471***	-7.602702***
d_holiday	-4.1891316*		-3.8408972***	-4.923577*
_cons	-355.68101***	0.01543064	-418.91713***	-270.53558***
ARMA				
L1. ar	.99520915***	.35830721***	.99069522***	.21811192***
L1.ma	-.83582277***	-.94340281***	-.80559593***	
sigma				
_cons	5.4625426***	9.5154166***		
ARCH				
L1.arch			.11660317**	
_cons			26.818613***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 18:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	30.06911***		36.742532***	21.267973***
log_windplangermany	-3.2806397***		-2.3653539***	-2.8921851***
L1.log_emissionprice	12.46726700		3.75163740	5.4780641*
L1.log_CrudeOilBrent	9.82196380		-11.52795200	14.444378***
log_temperature	-7.24878**		0.36557777	-5.0178927**
log_unav_lignite_planned	0.03394862		.44131711***	-0.11860062
log_unav_Gas_planned	-1.1732223**		-.75046624***	-1.8737651***
log_unav_coal_planned	-0.09736947		0.04804238	.60222448**
log_unav_uranium_planned	0.07768616		0.27657607	-0.23050829
log_unav_lignite_nonsched	.52648229*		0.18693973	0.11535176
log_unav_Gas_nonsched	0.15764834		0.16217672	.24255861*
log_unav_coal_nonsched	0.16294469		0.13816838	0.35636828
log_unav_uranium_nonsched	0.01725223		0.04019839	.40507243***
d_weekend	-4.640001***		-3.2502941***	-5.5069347***
d_holiday	-3.7499774*		-2.8227736***	-3.40094670
_cons	-291.82991000	0.00032849	-284.80757***	-223.37445***
ARMA				
L1. ar	.98003303***	.36648483***	.98797071***	.29805638***
L1.ma	-.7937967***	-.91215372***	-.78652419***	
sigma				
_cons	6.4775887***	9.4751676***		
ARCH				
L1.arch			.79039885***	
_cons			18.188574***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 19:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	28.193226***		28.178825***	29.017619***
log_windplangermany	-3.5973472***		-3.5095759***	-2.7202578***
L1.log_emissionprice	7.22426010		6.81751160	5.4592005*
L1.log_CrudeOilBrent	22.905954***		6.36720420	15.540737***
log_temperature	-9.8518305***		-10.058907***	-8.4705332***
log_unav_lignite_planned	0.12846274		.37203245*	0.14791285
log_unav_Gas_planned	-.84955378**		-.45001962**	-.94835241***
log_unav_coal_planned	0.12007899		0.06019070	0.41856752
log_unav_uranium_planned	0.13992850		0.22662520	0.05298892
log_unav_lignite_nonsched	0.32102231		0.17216836	-0.12838305
log_unav_Gas_nonsched	0.19414148		.20802739*	.31923308**
log_unav_coal_nonsched	0.19308875		-0.05934701	.47620985*
log_unav_uranium_nonsched	0.12069721		0.21576612	.47185465***
d_weekend	-2.7324007**		-2.8381191***	-1.9180469*
d_holiday	-2.06213950		-1.14733450	-1.29803110
_cons	-307.09253***	0.00637866	-235.83055***	-305.55835***
ARMA				
L1. ar	.95367983***	.31100961***	.98344353***	.24473227***
L1.ma	-.74816528***	-.90106759***	-.82961523***	
sigma				
_cons	5.7548101***	8.2358031***		
ARCH				
L1.arch			.63145715***	
_cons			18.189706***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 20:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	25.715961***		25.299293***	28.128639***
log_windplangermany	-3.0607893***		-3.0529566***	-2.1151667***
L1.log_emissionprice	9.18120330		7.70374740	5.6968784**
L1.log_CrudeOilBrent	22.766466*		26.29774***	14.517465***
log_temperature	-6.1966927**		-5.9511253***	-5.7816526***
log_unav_lignite_planned	0.19129693		0.21092476	.57893611***
log_unav_Gas_planned	-0.08918054		0.00873314	0.09497032
log_unav_coal_planned	0.30052445		0.14830518	0.12043278
log_unav_uranium_planned	0.11523631		-0.03635797	.35511848***
log_unav_lignite_nonsched	0.31080500		0.16153774	-0.01873075
log_unav_Gas_nonsched	.21996355*		.23703039*	.2716052**
log_unav_coal_nonsched	0.11068868		0.07893555	0.26745935
log_unav_uranium_nonsched	0.09551349		0.16570254	.35399053***
d_weekend	-1.48600730		-1.6037619*	0.13755218
d_holiday	-1.45454420		-1.73294380	1.08138640
_cons	-308.45472***	0.03174329	-316.91514***	-322.42504***
ARMA				
L1. ar	.97340542***	.31455477***	.95317371***	.35994521***
L1.ma	-.74610488***	-.87208057***	-.67423994***	
sigma				
_cons	5.0143335***	6.7090862***		
ARCH				
L1.arch			.38826014***	
_cons			16.880663***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 21:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	22.362283***		23.606601***	23.920223***
log_windplangermany	-2.7941831***		-2.7277516***	-2.1361971***
L1.log_emissionprice	5.90447860		4.56865630	5.541163*
L1.log_CrudeOilBrent	23.96739700		27.603183***	19.896204***
log_temperature	-4.2447985*		-4.6014834***	-2.355185*
log_unav_lignite_planned	.35908002*		.39245992**	.68714896***
log_unav_Gas_planned	-0.07274793		-0.03370575	0.12211702
log_unav_coal_planned	0.09571129		0.09000431	0.09563243
log_unav_uranium_planned	0.16439223		0.19641233	.34742636***
log_unav_lignite_nonsched	.44158985**		.3568285*	0.19353306
log_unav_Gas_nonsched	0.12177323		0.11462531	0.03969395
log_unav_coal_nonsched	-0.00528526		-0.02613711	-0.09211456
log_unav_uranium_nonsched	0.12096985		0.07948632	.38948819***
d_weekend	-1.37120950		-1.292373*	-0.29710317
d_holiday	-0.65841586		-0.82363163	1.12641630
_cons	-280.13063***	0.04231088	-308.39874***	-310.10247***
ARMA				
L1. ar	.98236088***	.31743725***	.98190404***	.305114***
L1.ma	-.7876479***	-.89775074***	-.76239495***	
sigma				
_cons	4.3106161***	5.7299813***		
ARCH				
L1.arch			.14415241*	
_cons			15.987987***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 22:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	19.968679***		19.75627***	12.091723***
log_windplangermany	-2.3720006***		-2.378019***	-2.3254129***
L1.log_emissionprice	9.86545760		8.93927900	10.053267***
L1.log_CrudeOilBrent	19.07745800		26.274299***	19.447815***
log_temperature	-3.19902040		-3.0176058**	-1.85577150
log_unav_lignite_planned	.29238464*		.32824174**	.53100082***
log_unav_Gas_planned	-0.09131124		-0.06826825	0.03245842
log_unav_coal_planned	0.12832493		0.11493520	0.15158070
log_unav_uranium_planned	0.11444437		0.18719973	.29600312***
log_unav_lignite_nonsched	.28698897*		.26613183*	0.10131502
log_unav_Gas_nonsched	0.07768405		0.06784163	-0.02744808
log_unav_coal_nonsched	0.01069450		0.00235299	-0.06127085
log_unav_uranium_nonsched	0.11114793		0.09034521	.32843425***
d_weekend	-0.29516407		-0.35383576	-.95735448*
d_holiday	0.96040159		1.01538640	1.47688590
_cons	-251.75125***	0.03226369	-282.57532***	-190.51708***
ARMA				
L1. ar	.98391742***	.33076319***	.97567022***	.25959668***
L1.ma	-.77559602***	-.90026561***	-.75781638***	
sigma				
_cons	3.6009939***	4.6342181***		
ARCH				
L1.arch			0.03492667	
_cons			12.600915***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 23:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	14.263912***		13.891256***	8.8588427***
log_windplangermany	-2.3886777***		-2.3987591***	-2.3503405***
L1.log_emissionprice	4.41656620		4.67286560	11.016525***
L1.log_CrudeOilBrent	7.43353240		4.26962920	17.017938***
log_temperature	-3.7004528**		-3.9214285***	-1.53617600
log_unav_lignite_planned	0.22955066		0.23806056	.44859507***
log_unav_Gas_planned	-0.03500992		-0.03947944	-0.09226035
log_unav_coal_planned	0.05785177		0.02091638	0.06793111
log_unav_uranium_planned	0.06186226		0.07891398	.18961352*
log_unav_lignite_nonsched	.30893659*		.31036834**	0.06680460
log_unav_Gas_nonsched	.14928417*		0.15005041	0.07866142
log_unav_coal_nonsched	0.07172104		0.07230006	0.10772475
log_unav_uranium_nonsched	0.10546854		0.08387247	.16860921**
d_weekend	1.04207650		1.0221909*	0.62141964
d_holiday	2.2565445*		2.0617488**	3.0467432**
_cons	-121.54707000	0.02346748	-114.03937**	-147.75668***
ARMA				
L1. ar	.98989467***	0.15504584	.9991723***	.27854674***
L1.ma	-.82338492***	-.84604587***	-.82648903***	
sigma				
_cons	3.5354617***	4.2851841***		
ARCH				
L1.arch			0.02881849	
_cons			12.191944***	

legend: * p<0.05; ** p<0.01; *** p<0.001

Hour 24:

Variable	ARIMAX	ARIMA	MGARCH	OLS
log_PlannedGeneration	14.396106***		11.203437***	13.274795***
log_windplangermany	-3.6947543***		-3.4995838***	-3.4306552***
L1.log_emissionprice	15.968828***		12.205486***	12.677937***
L1.log_CrudeOilBrent	23.809716***		20.522193***	18.993467***
log_temperature	-2.20560000		-3.6701909***	-2.0287939*
log_unav_lignite_planned	.43420307***		0.31650979	.46966943***
log_unav_Gas_planned	-0.13072817		-.58593874***	-0.05764743
log_unav_coal_planned	-0.08095092		-0.13714903	-0.12063854
log_unav_uranium_planned	0.36709827		.81886499***	.40158692*
log_unav_lignite_nonsched	0.07154900		.78517307***	0.00368402
log_unav_Gas_nonsched	0.02504886		-0.03923612	0.01128442
log_unav_coal_nonsched	0.08568627		0.09751121	0.10127377
log_unav_uranium_nonsched	0.25685276		.38179914***	.27391146**
d_weekend	0.04524810		-0.24428238	0.04218470
d_holiday	-1.41249510		-0.36549002	-0.18893381
_cons	-228.43738***	0.02283126	-170.32832***	-197.17905***
ARMA				
L1. ar	.77427567***	.29044574***	.76005563***	.18575654***
L1.ma	-.55740662***	-.93533213***	-.52081989***	
sigma				
_cons	5.2351134***	6.2720355***		
ARCH				
L1.arch			1.2267741***	
_cons			10.05351***	

legend: * p<0.05; ** p<0.01; *** p<0.001