Electricity Price Forecasting using Sale and Purchase Curves: The X-Model

Florian Ziel, Rick Steinert

Europa-Universität Viadrina Frankfurt (Oder), Germany

September 2, 2015

arXiv:1509.00372v1 [q-fin.TR] 1 Sep 2015

Abstract

Our paper aims to model and forecast the electricity price in a completely new and promising style. Instead of directly modeling the electricity price as it is usually done in time series or data mining approaches, we model and utilize its true source: the sale and purchase curves of the electricity exchange. We will refer to this new model as X-Model, as almost every deregulated electricity price is simply the result of the intersection of the electricity supply and demand curve at a certain auction. Therefore we show an approach to deal with a tremendous amount of auction data, using a subtle data processing technique as well as dimension reduction and lasso based estimation methods. We incorporate not only several known features, such as seasonal behavior or the impact of other processes like renewable energy, but also completely new elaborated stylized facts of the bidding structure. Our model is able to capture the non-linear behavior of the electricity price, which is especially useful for predicting huge price spikes. Using simulation methods we show how to derive prediction intervals. We describe and show the proposed methods for the day-ahead EPEX spot price of Germany and Austria.

1 Introduction

In the recent decades modeling electricity prices have become a complex and broad field of research. Due to the liberalization of markets and increasing disclosure of data, new insights concerning the structure and behavior of the prices were gained. Researcher pointed out that there are typical characteristics of electricity prices regardless where it has been traded. These are summarized as the stylized facts of electricity prices, see e.g. [23]. One of these stylized facts concerns tremendous deviations of the price pattern from its mean, named as price spikes. This specific feature of electricity prices has huge impacts for research as well as politics and companies. Many electricity companies, e.g. in Germany, are obliged to market some of their electricity at an exchange, which makes their earnings prone to heavy price spikes and creates a complex task for their risk management department. Moreover, many financial contracts such as futures or options are dependent on the variance of the price process and therefore demand eligible estimation techniques. Also long-term cost calculation for investment projects or political programs like the development of renewable energy are dependent on stable and reliable methods for calculation of electricity prices, which can account for the likelihood of price spikes.

Therefore, a great variety of models for estimating the electricity price occurred during the past decades. Those models are often related to well-known models of the finance literature but can originate from many other fields of research. [25] for instance divides electricity price models into five different groups, multi-agent, fundamental, reduced-form, statistical and computational intelligence models. Besides the multi-agent approaches all models have in common

that they focus on the price itself or related time series like renewable energy or electricity demand. Multi-agent models usually focus on the supply and demand of electricity to obtain prices by equilibrium, optimization or simulation ([22], [16]), but hence often do not incorporate the time-series of electricity bids and asks of a real exchange into their approaches.

The distinction between different model approaches can be refined when the explicit or implicit incorporation of price spikes is considered. In the area of time series models the usage of specific heteroscedastic models for the variance of the process are typical (e.g. [1], [15]). But standard GARCH-type models cannot account for all of the extreme price events within the data ([20]). Hence, many researcher developed extended models which can account for severe price movements. These models commonly fall into two main categories. First, there are regime-switching models, which introduce different regimes, usually a base and a spike regime, with different probabilities for a price spike to occur (see, for instance [12], [11], [7]). Second, there are diffusion models, which add a jump component, e.g. a Poisson process, to allow for price spikes (see, for instance [24], [8]). Rarely there are approaches which focus solely on the price spike itself and try to forecast the event without modeling the whole price time series, e.g. in [4].

However, all of these approaches for modeling price spikes have in common that they are focused mainly on the price time series and not of the underlying mechanic which causes the price process. The electricity price can also be seen as the intersection between the part of the electricity supply and demand which was traded at an exchange. The resulting sale and purchase curves, which are also referred to as ask and bid curves or supply and demand curves, contain all the information which is needed to determine the market price but provide even further information on all the other prices for other market volumes. This information can be necessary especially for the estimation of the likelihood of extreme price events, as the elasticity of the price, which can be obtained from the shape of the sale and purchase curves, vastly accounts for price movements.

Our idea therefore tries to connect the different categories of electricity price estimation under considering extreme price movements. We will use the true data generating process, e.g. the sale and purchase curves of the electricity price, to provide better forecasts for extreme price movements while still modeling the time series of electricity prices by an autoregressive approach. We will use the hourly day-ahead electricity price auction data of Germany and Austria provided by the EPEX Spot. It will be shown that incorporating the sale and purchase data yields promising results for forecasting the likelihood of extreme price events. Within our approach we will be able to estimate the full prediction density of electricity prices.

Our paper is organized as follows. The next section focuses on our idea and will describe the data and our observations for the EPEX Spot day-ahead auctions. We will follow up with a detailed description of our model and its specific setup for the auction data. Afterwards we show the empirical results of our approach. Our last section discusses our findings and will provide insights for possible improvements and future research. During the paper we will use the phrase "price curves" for both, the sale and purchase curve. Every price will be provided as price in EUR/MWh and every volume in MW, if not specified otherwise.

2 Price formating process and price curves structure

The electricity price of exchanges is the result of competitive bidding and asking. Focusing merely on the time series of prices therefore neglects their true source. If the true sale and purchase curves were known, the price could be solely determined by the intersection of both curves - regardless of any time dependencies between different prices. Many authors point out that the price is driven by external factors, e.g. wind and solar or electricity demand, see for instance [25]. However, taking a closer look on the underlying price process, it can be stated that it is the buyers and sellers on an electricity exchange who are influenced by those factors and therefore adjust their bids. Reasons for that can be e.g. that these market participants are

electricity companies who are facing heavy overproduction of electricity due to an unexpected change in wind speed or temperature or an underproduction due to outages of power plants.

But those market participants are not equal, they can be investment companies, electricity producers or transmission service operators, among others. Also not all electricity producers are equal - they have distinct production portfolios and are therefore more or less likely prone to e.g. heavy weather conditions. An unexpected shift in wind production levels for instance can therefore lead to a little or vast change in prices, dependent on if the equilibrium price of the market was already mainly driven by producers who have high investments in wind turbines. This diversified information is summarized in the sale and purchase curve of electricity prices. Hence, especially for estimating heavy price movements it is essential to know, if the market is capable of adjusting for external shocks easily or if a tremendous price spike will occur. This sensitivity of the intersection price can therefore be obtained by analyzing the original price curves instead of only their outcome as price time series. To motivate our idea even further, we decided on showcasing the day-ahead price of the 12.04.2015 of the EPEX Spot for Germany and Austria. We will use this day throughout our whole paper, as it provides easily traceable insights for the typical price movement process when an extreme price spike occurs.

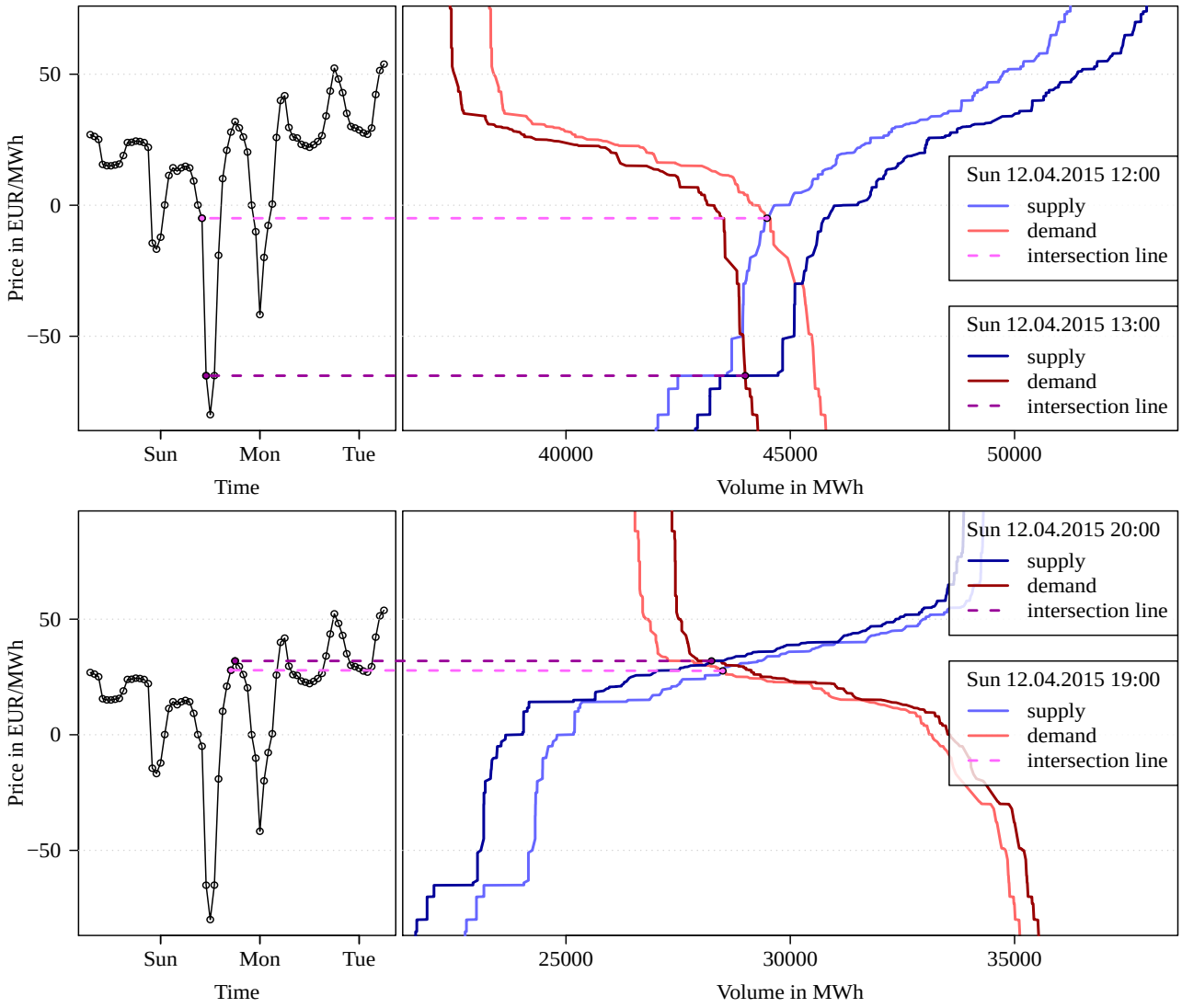


Figure 1: Market clearing price time series with supply and demand curves for 12:00 and 13:00, and 19:00 and 20:00 hours resp. on 12.04.2015 auctions.

The left-hand side of Figure 1 shows the day-ahead electricity price of the 12.04.2015. In the upper and lower right-hand side the price curves for 12:00 and 13:00 and 19:00 and 20:00 respectively are provided. The horizontal axis of this area represents the trading volume and the

vertical axis the price. It is shown that during the afternoon hours the electricity price heavily declined reaching even negative values. Examining the price curves for 12:00 that day on the upper right-hand side two typical phenomena for such an observation can be seen. First, the traded volume is, in comparison to other hours that day, relatively high. Second, the slope of the demand curve for any price and volume combination with a lower price than the market price is extremely negative. Simultaneously, the slope of the supply curve is extremely positive for price and volume combinations with a lower price than the market price - at least for price combinations close to the actual market price. Monitoring the left- hand side of the figure shows that the price exhibited a tremendous price decline from 12:00 to 13:00. Taking into account the phenomena mentioned beforehand gives insights on why such a heavy price spike was even possible. The high amount of supplied electricity shifted the price to a level, where usually only a relatively small proportion of bids can be found, e.g. the supply and demand curves exhibited large "steps". Those "steps" result in the second observation of curves having extreme negative or positive slopes close to the market price. This in turn indicates, that the equilibrium price is very sensitive to external shocks. Any sudden decrease in demand which would lead to a left-shift of the demand curve or any sudden increase in production which would lead to a right-shift of the supply curve has a great impact on the price - especially in comparison to other, higher price levels. But we can also see that the supply curve for 13:00 exhibits a slope of almost zero around the intersection price, indicating that any further decrease in demand or increase in supply will not have the same vast effects than before. And indeed, the price movement from 13:00 to 14:00 was much smaller than the one from 12:00 to 13:00. In contrast, the price curves for 19:00 and 20:00 on the lower right-hand side of the figure show the typical behavior of price curves when the market is not very prone to extreme events. The slopes of the price curves right from the intersection price seem not to have extremely positive and negative slopes respectively. Only the demand curve left of the intersection price seem to have a very negative slope, but matching it with the supply curve it can be seen that any shifts to the right or left will be captured by the supply curve easily and can therefore not result in a heavy price spike.

Under the assumption that not only the price but also the price curves are dependent on time, we will derive a model which is tailor-made for the day-ahead auction market of the EPEX Spot in Germany and Austria for the purpose of estimating the likelihood of heavy price movements. Therefore and in order to introduce our model we need to take a closer look at the EPEX Spot market and the observed bidding structure of their participants.

For our summary statistics we use the data from 01.10.2012 to 19.04.2015. If not mentioned otherwise all prices where given in EUR/MWh and volumes in MW. However, all techniques can be applied to other electricity markets in exactly the same way but under consideration of their corresponding market features.

The day-ahead electricity spot price of the EPEX will be traded in daily auctions at 12:00 CET for the hours of the next day. So there are in general 24 prices everyday. Due to the daylight saving time we have once a year 23 values in March and 25 values in October. For the 24 auctions on a common day we use the labels 0:00, 1:00, ..., 23:00 within this paper.

Since 2008 the electricity spot price is set to be between $P_{\min} = -500$ and $P_{\max} = 3000$. Before that there were no negative prices allowed. The traders at the EPEX can make bids for either selling or buying a certain amount of electricity power. By the EPEX regulations, the minimal order size for Germany and Austria is 0.1 MW for a one hour block and the minimal price difference between different orders is 0.1 EUR/MWh. Hence, there are in total 35001 different possible prices on the full price grid $\mathbb{P} = \{-500, -499.9, \ldots, 2999.9, 3000\}$. But in practice not every of those possible prices is utilized. Also, a single trader can only submit up to 256 distinct price and volume combinations as offers. Usually there are about 700 different featured prices which construct each curve, which we depicted in Figure 2. The illustration shows the histogram for the amount of different prices for both, the demand and the supply curve of electricity at the EPEX Spot. It can be seen that the range of different prices covers approximately 200 to 1000, depending on which side of the market is considered. In general we

have slightly more different prices for the supply side. In total there were about 31000 bids on distinct prices within the considered period.

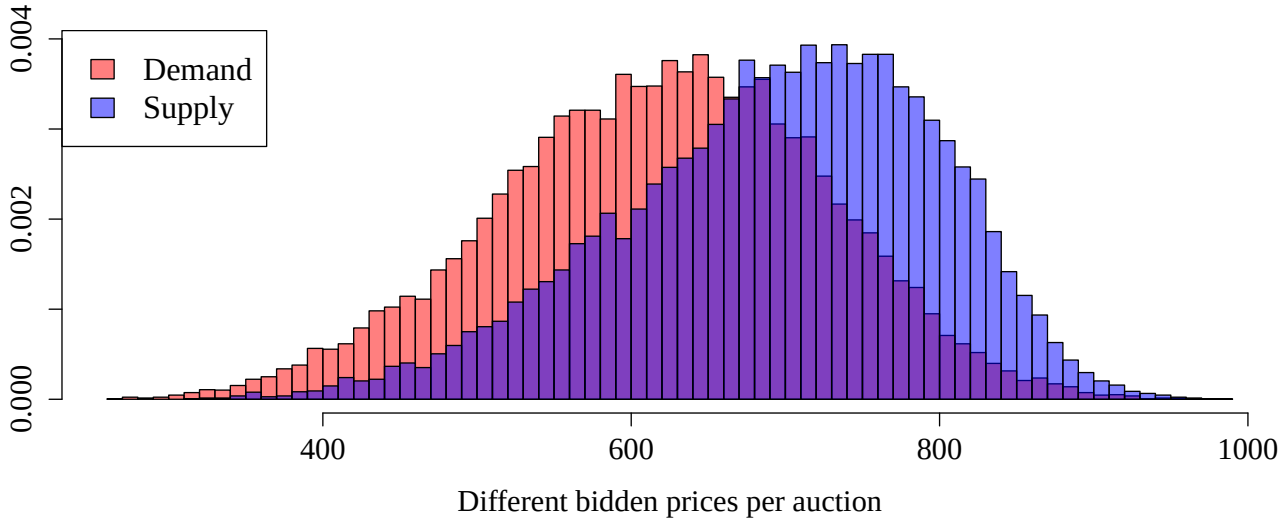
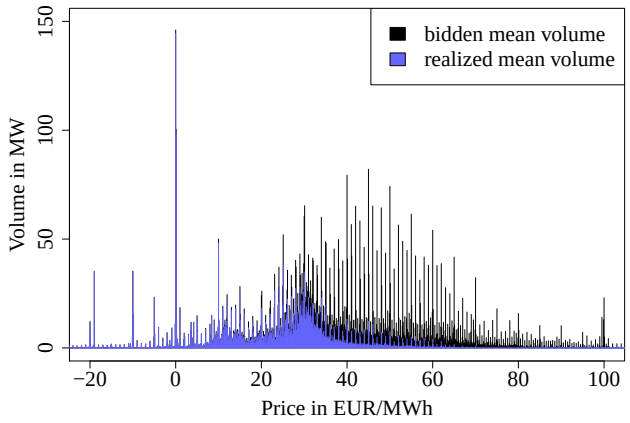


Figure 2: Amount of different bidden prices for the auctions.

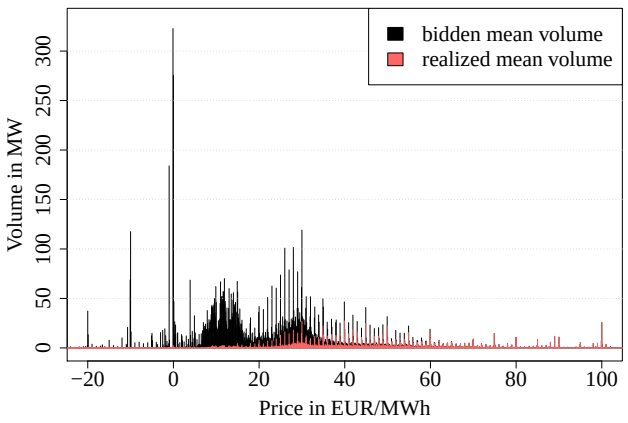
When these bids are aggregated the well-known price curves for a certain hour can be constructed, which maps a certain amount of supply or demand to a certain price, even if there was no specific bid for that price and volume combination. The resulting price curves are step functions, as there can not be sold more than bidden at a certain price. According to the operational rules of the EPEX the market clearing price is determined by connecting the mentioned price points with a specific algorithm, which considers also special orders like "All-or-None Block Orders". Nevertheless, it is assumed by the EPEX that the relation of two different bidden price and quantity combinations of one market participant is linear. Therefore and to simplify the used algorithm we will use linear interpolation between to different price and quantity combinations. So the market clearing price will be calculated by the resulting intersection of both price curves. Consequently there is sometimes a small but rather negligible difference to the true market price. Note that due to the interpolation the resulting price and volume might be not a bidden one. This fact is managed by certain rules for the traders, so that the amount of volume of the market clearing price must be delivered.

As mentioned before, most of the prices in $\mathbb{P}$ where not bidden at all. Nevertheless, given the price grid $\mathbb{P}$ we can compute the mean bidden volumes $\mathbb{V}(P)$ given a $P \in \mathbb{P}$. Figure 3 shows the average bidden volume of supply and demand volumes $\mathbb{V}_S(P)$ and $\mathbb{V}_D(P)$ with $P \in \mathbb{P}$ within the price range of -20 to 100. Additionally we highlighted the realized amount. We can observe that for the supply side almost all bidden low prices were realized whereas for high prices this is contrary. This relation is reversed for the demand side. We also observe some patterns in the bidding behavior of some traders. For example, we have quite large volumes at a price of 0, but very small amounts for e.g. 0.3 or -0.3. Furthermore we have spikes at multiples of 5, so e.g. 70.0 has a larger value than 69.0 or 71.0. This shows that agents seem to prefer to bid on round numbers. This may give indication towards the assumption that at least a noticeable amount of trading is done by human decision and not based on algorithmic trading rules.

The large bidden volumes at a certain price leads to a price cluster around this price. The most intense price cluster can be found at a value of zero, which can be retrieved from Figure 3. Even in the small time frame shown in Figure 1 there are four realized electricity price values very close to zero. For the auction at 12:00 we see that the high possibility for a price of zero is mainly driven by the supply side. Figure 1 also shows that for the auction at 12:00 there is another price cluster at -65, which is again driven by the supply side. And again, for the realized price we can observe two values very close to -65, e.g. the realized price for 13:00 is exactly -65.02.



(a) bidden and realized volumes: supply



(b) bidden and realized volumes: demand

Figure 3: Average bidden volumes $\mathbb{V}_S(P)$ and $\mathbb{V}_D(P)$ for $P \in \mathbb{P}$ in the price range -20 to 100.

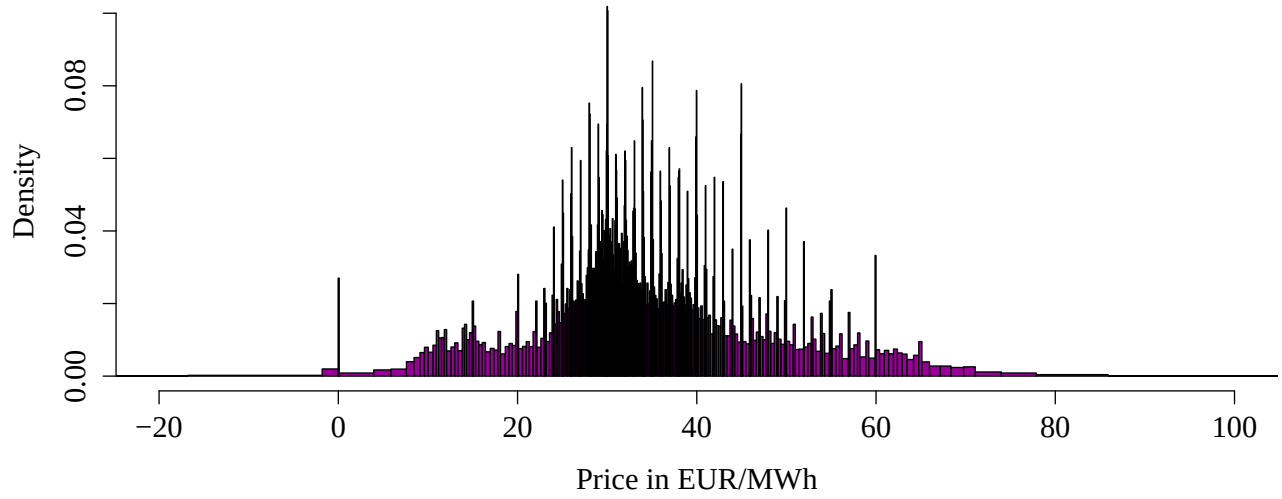


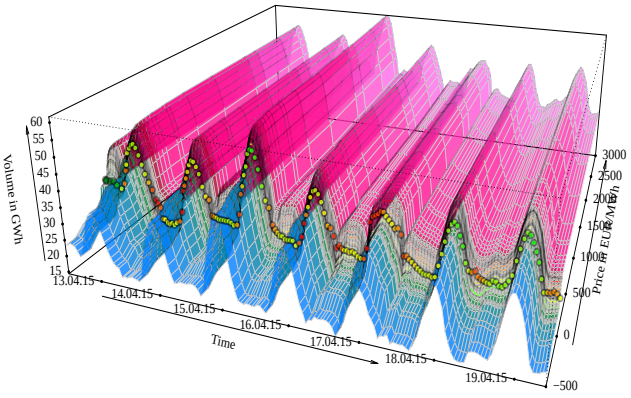
Figure 4: Histogram of realized prices (market clearing prices) in the price range -20 to 100.

In Figure 4 we see the histogram of all market clearing prices. To visualize the price cluster effect it is important to choose tiny histogram breaks. Here every tiny rectangle represents about 0.33% of the probability mass. We observe a very spiky histogram that exhibits some properties of the bidden prices in Figure 3. For example we see clearly a price cluster at zero. Here the relative frequency that the market clearing price is between -0.5 and 0.5 is with 0.634% relatively large. In contrast, the relative frequency to get a price in the neighboring intervals of the same size from -1.5 to -0.5 and 0.5 to 1.5 is only 0.079% and 0.056%, so about 10 times smaller. Other price clusters can again be found at all full integers between 10 and 60, where those clusters that are divisible by 5 are more distinct. In general the density of the electricity price is complicated due to its multi-modal shape. The modi are at the mentioned price clusters. As far as we know there is no model in electricity price modeling that at least tries to capture this behavior. In contrast to that, our modeling approach will partially incorporate this effect and thus try to capture the true market behavior more realistically.

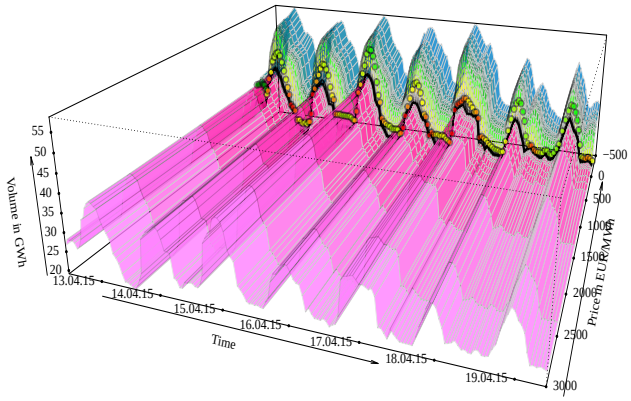
3 Model for the supply and demand curve

Modeling the supply and demand curve of electricity prices is a very complex task. Researcher who try to analyze the complex bidding structure of the supply and demand at electricity exchange usually utilize multi-agent models. ([25]) But those approaches do very rarely take into account the real time series of auction data and are therefore not eligible to give practical infor-

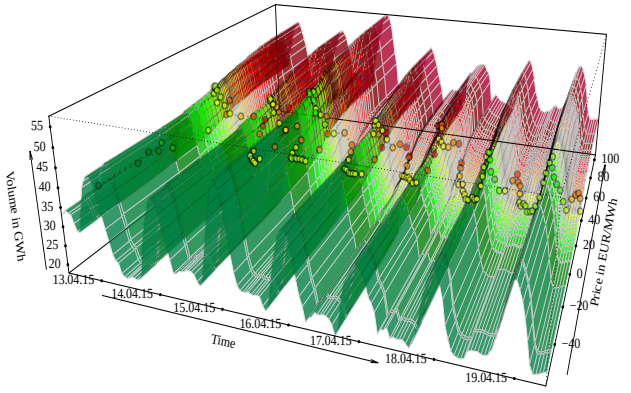
mation on short-term forecasts of the electricity price time series. This is especially interesting when taking a closer look at the price curves over time. In Figure 5 we show the time series of both price curves from 13.04.2015 to 19.04.2015 in a three-dimensional plot. To put emphasis on the price scale, which is presented at the z-axis, we added a colored legend for them which can be found in the lower two pictures of the figure. The upper two pictures show the price curves on the full price grid, whereas the the two pictures in the middle focus on the price range close to the market clearing price. Judging only by the figure it can be obtained that both, the supply and the demand curve, exhibit a seasonal pattern over time with at least daily dependence.



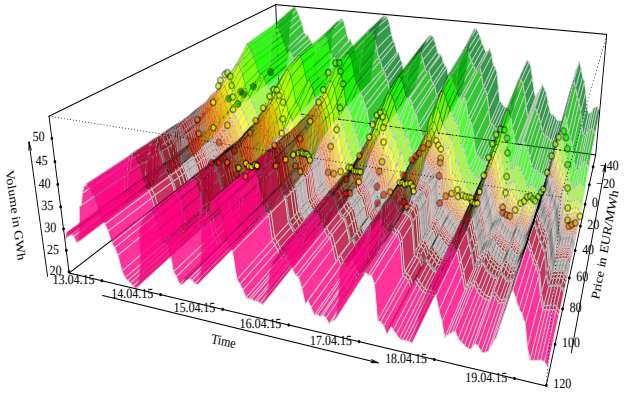
(a) Supply from -500 to 3000 in EUR/MWh



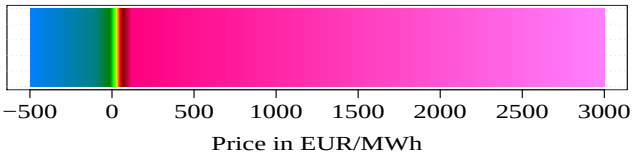
(b) Demand from -500 to 3000 in EUR/MWh



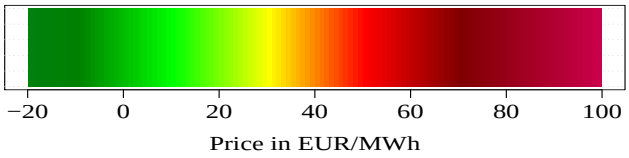
(c) Supply from -20 to 100 EUR/MWh



(d) Demand from -20 to 100 EUR/MWh



(e) Legend from -500 to 3000 in EUR/MWh



(f) Legend from -20 to 100 in EUR/MWh

Figure 5: 3d price curves from 13. April 2015 to 19. April 2015 in a price color plot.

Nevertheless, the time-series approach for modeling auction data is relatively new and has not been applied for electricity price data in a comprehensible manner. Even in general only very little papers can be found which actually incorporate real time-series of bids to gain information on prices. One application which is similar to ours at least in the sense of that it uses an econometric approach is the work of [2] who applies a functional signal plus noise time series model to a security of the FTSE100. In electricity price research a slightly related approach can be found in [10] and [3] who perform an analysis of the sale and purchase structure and integrate some of its aspects by incorporating the bid stack model. However, they assume that the demand is inelastic of the price, which is for our data obviously not the case for every point in time, as already shown.

However, to our knowledge there is not one single paper for the electricity market which actually models real price curves and use them to forecast real electricity price time-series by an econometric approach. Therefore the following sections will describe the necessary setup for such a model highly detailed and use references whenever our model idea makes use of a well-known econometric technique.

For our model we proceed in four steps:

1. To overcome the massive amount of data we will organize the bids in price classes. This will be discussed in section 3.1.
2. We provide a stochastic model to forecast the bidden volume of each price class. Section 3.2 will cover this step.
3. Given the forecasted bids within each price class we reassemble the precise bidding structure by reconstructing the classes. This will be covered in section 3.3.
4. Finally, we calculate the supply and demand curves to compute the market clearing price by the intersection of both curves. This will be also explained in section 3.3.

We will refer to our model for the sale and purchase curves of the electricity price as X-Model throughout the paper. We choose the letter X, as it symbolizes visually the intersection of the supply and demand curve.

3.1 Price classes for bids

As mentioned there are 35001 possible volumes on the full price grid $\mathbb{P}$. Theoretically we could model every of these processes but this is almost unfeasible due to computational burdens. Therefore we show how to choose and apply a simple dimension reduction procedure to the price formatting process that is computational manageable and still balances the related loss of information. Therefore, we merge the 35001 prices in $\mathbb{P}$ into a smaller amount of classes. For the bids within a price class we will assume later on that they behave similarly over time.

For creating the price classes we consider the mean bidden volume $\mathbb{V}_S(P)$ and $\mathbb{V}_D(P)$ as a measure of the importance for the price P for the supply and demand side of trading. We accumulate these volumes and calculate

$$\mathbb{V}_{S,\text{cum.}}(P) = \sum_{\substack{p\in\mathbb{P}\\ p\le P}} \mathbb{V}_S(p) \quad \text{and} \quad \mathbb{V}_{D,\text{cum.}}(P) = \sum_{\substack{p\in\mathbb{P}\\ p\ge P}} \mathbb{V}_D(p) \tag{1}$$

for the supply and demand side. Note that it is important to divide the cases for supply and demand. The functions $\mathbb{V}_{S,\text{cum.}}$ and $\mathbb{V}_{D,\text{cum.}}$ represent the price curves by linear interpolation as already seen in Figure 1, where cum. represents the fact that the values are accumulated.

Then we set an amount of volume V_* which will give the average amount of volume that should be represented by every price class. Therefore we define an equidistant volume grid $\mathbb{V}_* = \{iV_* | i \in \mathbb{N}\}$. Using $\mathbb{V}_*$ we can define the upper and lower values of the price classes by

$$\mathbb{C}_S = \mathbb{V}^{-1}_{S,\text{cum.}}(\mathbb{V}_*) = \{\mathbb{V}^{-1}_{S,\text{cum.}}(iV_*) | i \in \mathbb{N}\} \quad \text{and} \quad \mathbb{C}_D = \mathbb{V}^{-1}_{D,\text{cum.}}(\mathbb{V}_*) = \{\mathbb{V}^{-1}_{D,\text{cum.}}(iV_*) | i \in \mathbb{N}\}$$

where $\mathbb{V}^{-1}_{S,\text{cum.}}$ and $\mathbb{V}^{-1}_{D,\text{cum.}}$ are the (generalized) inverse functions of $\mathbb{V}_{S,\text{cum.}}$ and $\mathbb{V}_{D,\text{cum.}}$. Then the associated volumes at time t to the class prices $\mathbb{C}_S$ and $\mathbb{C}_D$ are given by

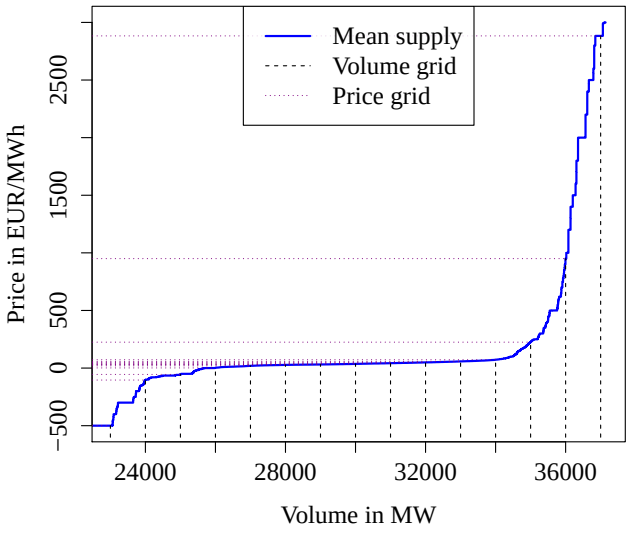
$$X^{(c)}_{S,t} = \sum_{P\in\mathbb{P}_S(c)} V_{D,t}(P) \quad \text{for} \quad c \in \mathbb{C}_S \quad \text{with}$$

$$\mathbb{P}_S(c) = \{P \in \mathbb{P} | P > \max\{p \in \mathbb{C}_S | p < c\}, P \le \min\{p \in \mathbb{C}_S | p \ge c\}\}$$

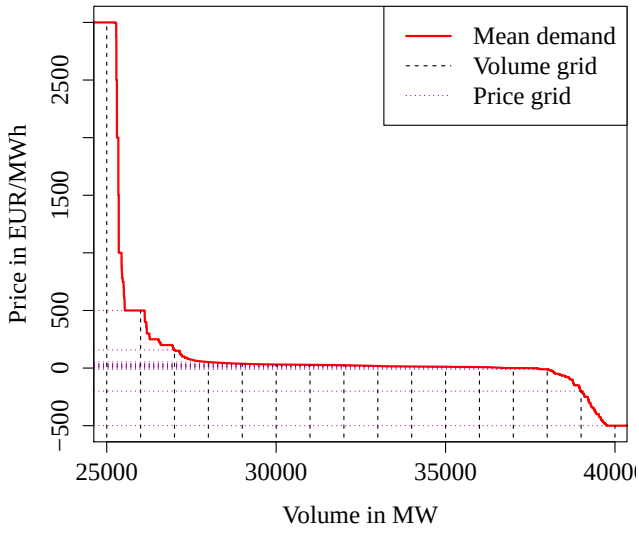
$$X^{(c)}_{D,t} = \sum_{P\in\mathbb{P}_D(c)} V_{D,t}(P) \quad \text{for} \quad c \in \mathbb{C}_D \quad \text{with}$$

$$\mathbb{P}_D(c) = \{P \in \mathbb{P} | P \ge \max\{p \in \mathbb{C}_D | p \le c\}, P < \min\{p \in \mathbb{C}_D | p > c\}\}$$

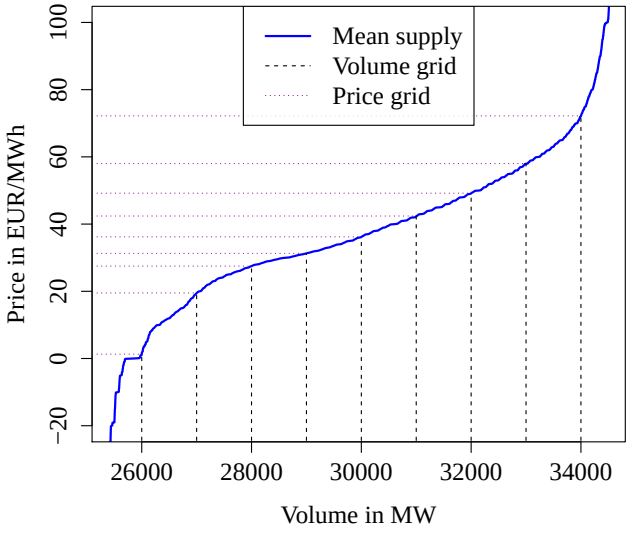
where $V_{S,t}(P)$ and $V_{D,t}(P)$ are the bidden supply and demand volume at price P and time t and c represents an arbitrary price of a specific price class. Here $\mathbb{P}_S(c)$ and $\mathbb{P}_D(c)$ are all prices that belong to the same price class as c.



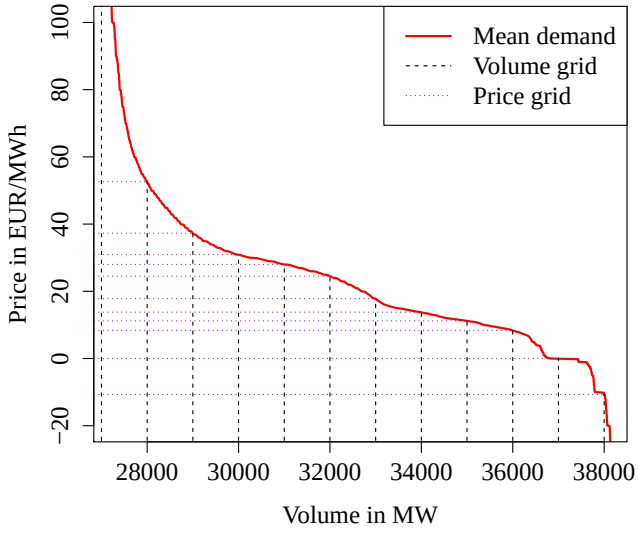
(a) Mean supply on range -500 EUR/MWh to 3000 EUR/MWh



(b) Mean demand on range -500 EUR/MWh to 3000 EUR/MWh



(c) Mean supply on range -20 EUR/MWh to 100 EUR/MWh

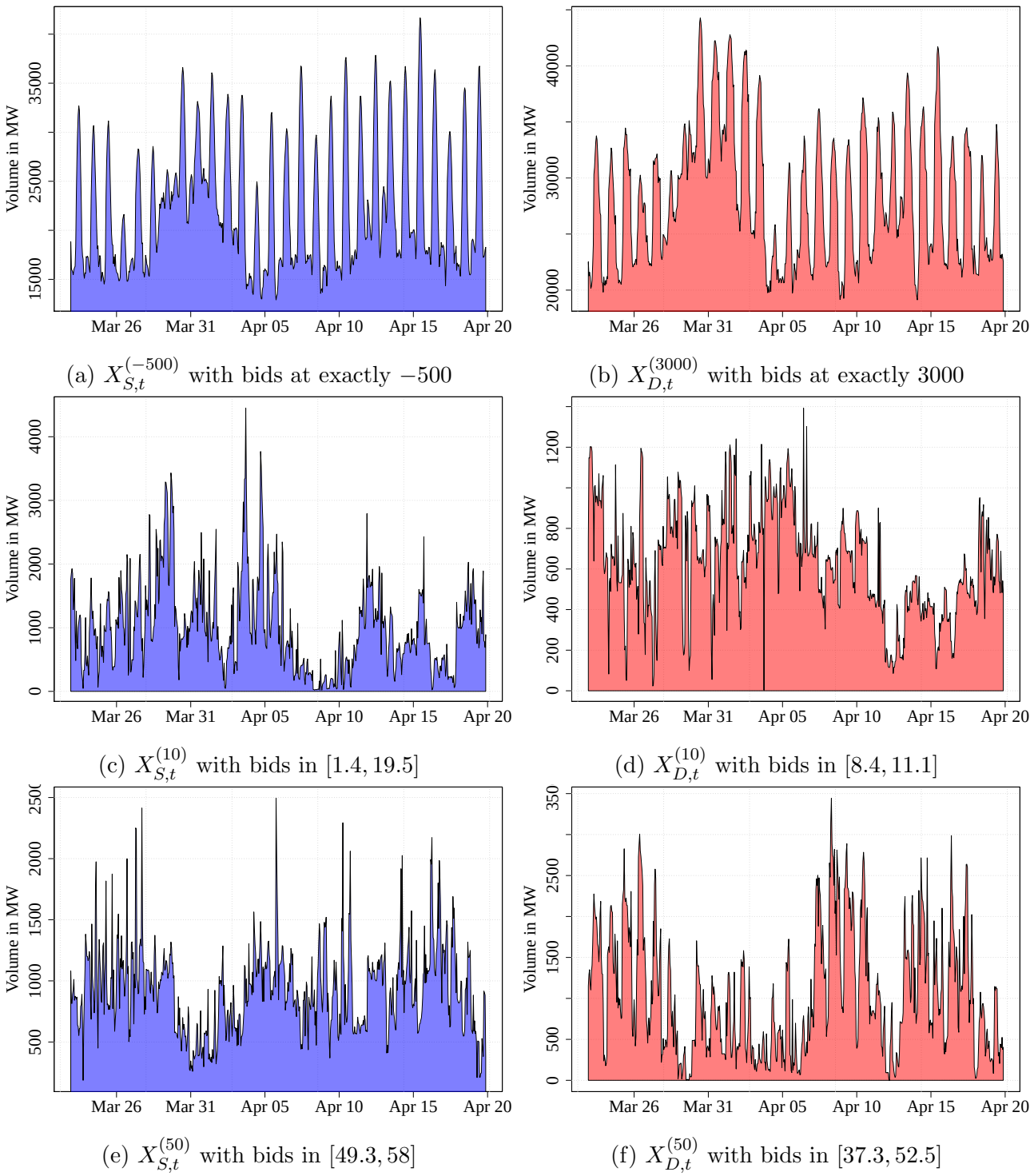


(d) Mean demand on range -20 EUR/MWh to 100 EUR/MWh

Figure 6: Mean supply and demand curves $\mathbb{V}_{S,C}$ and $\mathbb{V}_{D,C}$ on two selected price ranges with volume grid $\mathbb{V}_*$, price class grids $\mathbb{C}_S$ and $\mathbb{C}_D$ for $V^* = 1000$.

Figure 6 visualizes the classifying procedure for a volume of $V^* = 1000$. For our modeling approach later on we decided to stick with a volume size for classifying of 1000, as it provides us a manageable size of classes to estimate. However, other amounts of volume are definitely thinkable. Given our data we receive a total of $M_S = 16$ classes for the supply side with $\mathbb{C}_S = \{-500.0, -103.9, -55.1, \ldots, 3000\}$. This means for instance that $X_{S,t}^{(-500)}$ gives the amount of volume bidden on the supply side at exactly -500 at time t and $X_{S,t}^{(-103.9)}$ the amount between inclusively -499.9 and -103.9 at time t. For the demand side we also get $M_D = 16$ price classes with the price grid $\mathbb{C}_D = \{-500.0, -200.0, -10.7, \ldots, 3000\}$.

As an example we show several bid processes in Figure 7 in a short time perspective. Note that the illustrated bids $X_{S,t}^{(-500)}$ and $X_{D,t}^{(3000)}$ are very important in practice, as they represent large volumes. Moreover, both bids are favored by some market participants as they will always be realized. We observe that they have a more distinct seasonal structure than the common



(a) $X_{S,t}^{(-500)}$ with bids at exactly -500

(b) $X_{D,t}^{(3000)}$ with bids at exactly 3000

(c) $X_{S,t}^{(10)}$ with bids in $[1.4, 19.5]$

(d) $X_{D,t}^{(10)}$ with bids in $[8.4, 11.1]$

(e) $X_{S,t}^{(50)}$ with bids in $[49.3, 58]$

(f) $X_{D,t}^{(50)}$ with bids in $[37.3, 52.5]$

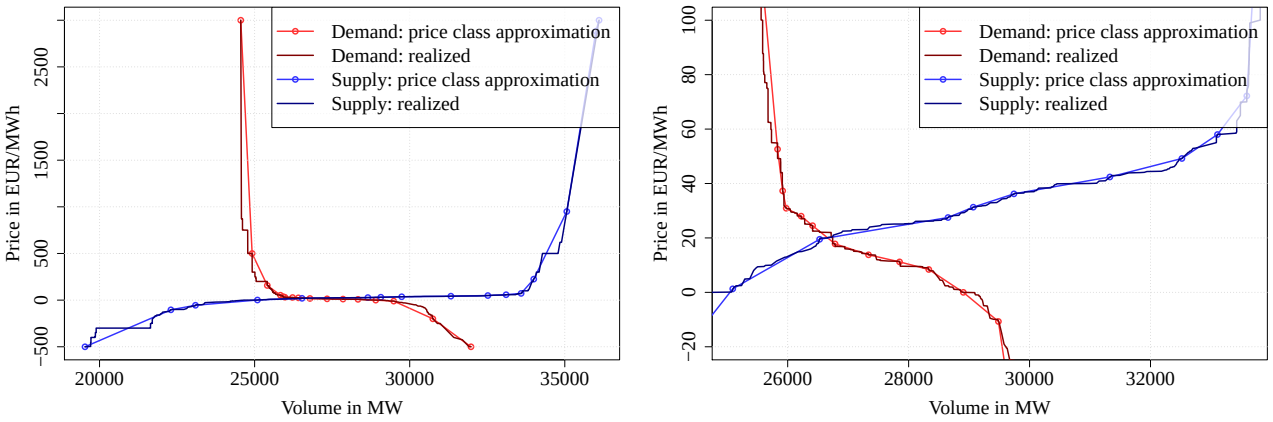
Figure 7: Offered volumes for supply and demand for four observed weeks in 2015.

bids. Due to the way we construct our classes, $X_{S,t}^{(-500)}$ and $X_{D,t}^{(3000)}$ are independent of the choice of volume V^*.

Every other bid process, e.g. $X_{S,t}^{(10)}$, $X_{D,t}^{(10)}$, $X_{S,t}^{(50)}$ and $X_{D,t}^{(50)}$, in Figure 7 has a a more complex structure. But many of them exhibit also a daily and weekly seasonal pattern.

Given the price classes $\mathbb{C}_S$ and $\mathbb{C}_D$ we can construct the corresponding price curves that approximate the real price curves. We can easily derive the price curves by accumulating $X_{S,t}^{(c)}$ in $c \in \mathbb{C}_S$ and $X_{D,t}^{(c)}$ in $c \in \mathbb{C}_D$. Figure 8 plots the class-based constructed price curves and the corresponding volume of a selected day, e.g. 19.04.2015 08:00. The first plot shows the price curves covering the full price range from -500 and 3000. In order to achieve a more detailed view on the price area around the market clearing price of that day and hour, we also plotted the price curves in the price range of -20 to 100. Comparing our approximation of a curve by

classes and the true price curve shows that both lines usually lie very close to each other, with minor deviations possible especially at very high or low prices.



(a) Range -500 EUR/MWh to 3000 EUR/MWh (b) Range -20 EUR/MWh to 100 EUR/MWh

Figure 8: Price curve example of 19th April 2015 08:00

3.2 Time series model for bid classes

Now we provide a model for the bids $X_{S,t}^{(c)}$ and $X_{D,t}^{(c)}$ of the price classes. Therefore, we introduce $X_{d,h}^{S,c}$ and $X_{d,h}^{D,c}$ as the bidden supply and demand volume of price class $c \in \mathbb{C}_S$ resp. $c \in \mathbb{C}_D$ at day d and hour h. For the well-known issue with the clock change due to daylight saving time we decided to interpolate the missing hour in March with the two hours around the missing hour and use the average of the double hours in October so that there are 24 observable prices each day. Thus, the volume processes $X_{d,h}^{S,c}$ and $X_{d,h}^{D,c}$ are well defined.

As mentioned above the processes $X_{d,h}^{S,-500}$ and $X_{d,h}^{D,3000}$ play an important role. But we also consider the impact of other possible sources that might influence the bidding behavior. In detail we use the EPEX market clearing price and volume of Germany and Austria of previous auctions, the planned electricity power generation in Germany of conventional power plants with more than 100 MW power as well as the planned wind and solar power feed-in. The last three processes are provided by the EPEX transparency database. Hence, we assume that market participants have access to this database or similar information and base their bids at least partially on those time series. Especially the impact of wind and solar energy on the electricity prices itself is well known (see e.g. [5], [14] or [27]). Therefore we introduce $M_X = 5$ additional processes denoted by $X_{\text{price},t}$, $X_{\text{volume},t}$, $X_{\text{generation},t}$, $X_{\text{wind},t}$ $X_{\text{solar},t}$ that represent the additional information that is available at the time where the auction will take place. A sample of the considered processes is given in Figure 9.

Similarly to $X_{d,h}^{S,c}$ and $X_{d,h}^{D,c}$ we introduce the slighly transformed processes $X_{d,h}^{\text{price}}$, $X_{d,h}^{\text{volume}}$, $X_{d,h}^{\text{generation}}$, $X_{d,h}^{\text{wind}}$ and $X_{d,h}^{\text{solar}}$ at day d and hour h. Note that the planned generation as well as the projected wind and solar power is known for one day in advance so we can use e.g. $X_{d+1,h}^{\text{generation}}$ to predict $X_{d+1,h}^{S,c}$.

The considered model is a simple regression approach and similar to the basic autoregressive model as used in [26] or [17] for modeling the electricity price. But we will use it in a more flexible way for the bidden volume processes of the price classes. For example, [26] allow for a linear dependency of $X_{d,h}^{\text{price}}$ to $X_{d-1,h}^{\text{price}}$, $X_{d-2,h}^{\text{price}}$ and $X_{d-7,h}^{\text{price}}$ as well as dummies on Sunday, Monday and Saturday. However, the choice of lags 1, 2 and 7 as well as the selection of the weekday dummies is the same for all 24 hours. We will allow for much more flexibility in the model, as the structure of the data is far more complex. This applies to both, the autoregressive lag structure and the weekday structure.

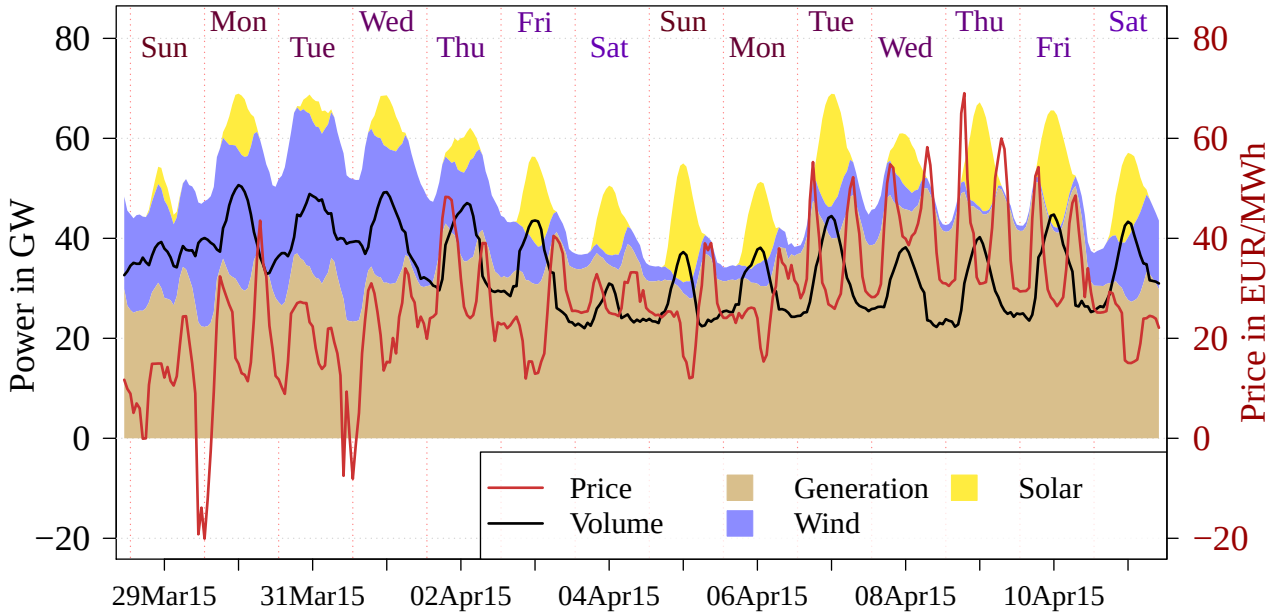
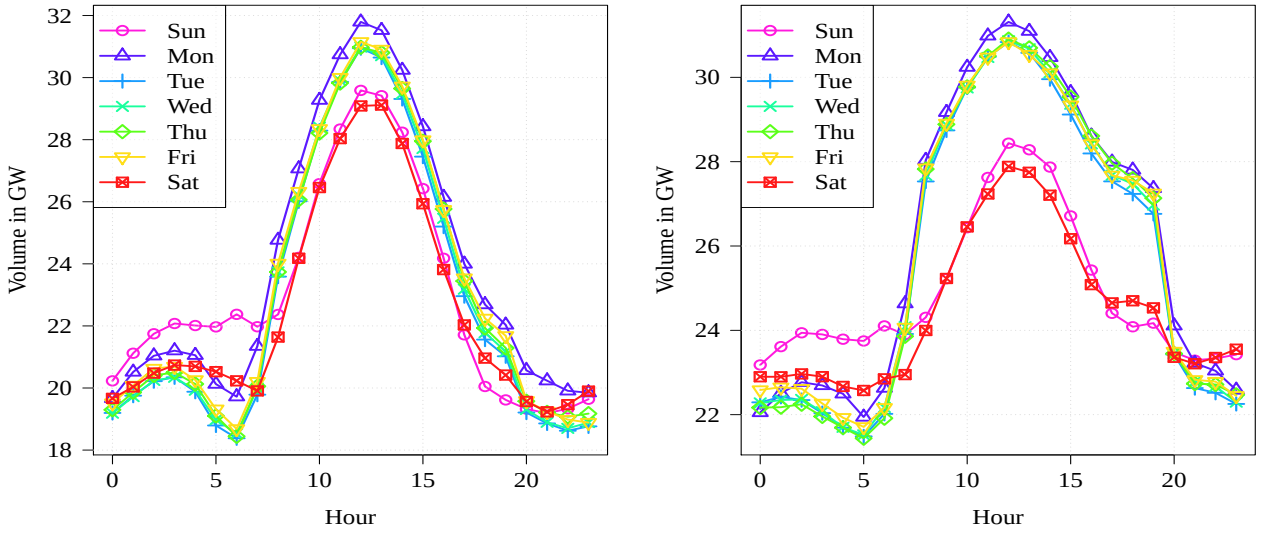


Figure 9: Sample of $X_{\text{price},t}$, $X_{\text{volume},t}$, $X_{\text{generation},t}$, $X_{\text{wind},t}$ $X_{\text{solar},t}$ from 29.03.2015 to 11.04.2015.

In Figure 10 the weekly sample mean of the bidden volume processes $X_{d,h}^{S,-500}$ and $X_{d,h}^{S,3000}$ for our full sample time range is given. There we can see that the daily seasonal structure seems



(a) Supply $X_{d,h}^{S,-500}$ (b) Demand $X_{d,h}^{D,3000}$

Figure 10: Weekly mean bidden volumes for each day of the week and hour of the day.

to depend on the day of the week, as it is typical for the electricity market clearing price or the electricity load (see e.g. [27]). So we see that the Saturdays and Sundays have a clearly different behavior than the other weekdays. But from 0:00 to 6:00 the Saturday seems to leave the typical pattern of a Sunday. Furthermore, we recognize that for the demand side the hours from 8:00 to 19:00 are clearly on an higher level during the working days. This is interesting, as it exactly matches the peakload standard block order at EPEX.

For modeling the day of the week impact we define the weekday indicators

$$W_k(d) = \begin{cases} 1 & , \mathcal{W}(d) < k \\ 0 & , \mathcal{W}(d) \geq k \end{cases}$$

where $\mathcal{W}(d)$ is a function that gives a number that corresponds to the weekday of day d. We use without loss of generality $k = 1$ for a Monday, for a Tuesday $k = 2$ up to $k = 7$ for a Sunday.

To fully present the considered time series model, it is necessary to introduce the object $\boldsymbol{X}_{d,h} = (X_{1,d,h}, \ldots, X_{M,d,h})' = ((X^{S,c}_{d,h})_{c\in\mathbb{C}_S}, (X^{D,c}_{d,h})_{c\in\mathbb{C}_D}, X^{\text{price}}_{d,h}, X^{\text{volume}}_{d,h}, X^{\text{generation}}_{d+1,h}, X^{\text{wind}}_{d+1,h}, X^{\text{solar}}_{d+1,h})'$. As the planned processes (generation, wind and solar) are known one day in advance they are represented with the day $d+1$ in the object $\boldsymbol{X}_{d,h}$. Note that the dimension of $\boldsymbol{X}_t$ is $M = M_S + M_D + M_X = 16 + 16 + 5 = 37$, but we have to model and forecast only the first $M_S + M_D = 32$ components for each hour h. Subsequently, we assume that $(\boldsymbol{X}_{d,h})_{d\in\mathbb{Z}}$ is a stationary process. However, we do not impose a time series model to $\boldsymbol{X}_{d,h}$ directly, but to its zero mean process $\boldsymbol{Y}_{d,h} = \boldsymbol{X}_{d,h} - \boldsymbol{\mu}_h$ with $\boldsymbol{\mu}_h = \mathbb{E}(\boldsymbol{X}_{d,h})$. We estimate the mean $\boldsymbol{\mu}_h$ by the corresponding sample mean.

Now for each hour h the considered time series model of $\boldsymbol{Y}_{d,h} = (Y_{1,d,h}, \ldots, Y_{M,d,h})'$ is constructed that it can potentially depend linearly on $\boldsymbol{Y}_{d-k,h}$, but also on a different hour $\boldsymbol{Y}_{d-k,j}$ with $j \neq h$ and the introduced weekday dummies. The considered time series model for $Y_{m,d,h}$ for each hour h and $m \in \{1, \ldots, M_S + M_D\}$ is given by

$$Y_{m,d,h} = \sum_{l=1}^{M} \sum_{j=1}^{24} \sum_{k\in\mathcal{I}_{m,h}(l,j)} \phi_{m,h,l,j,k} Y_{l,d-k,j} + \sum_{k=2}^{7} \psi_{m,h,k} W_k(d) + \varepsilon_{m,d,h} \tag{2}$$

with parameters $\phi_{m,h,l,j,k}$ and $\psi_{m,h,k}$, $\mathcal{I}_{m,h}(l,j)$ as index sets of lags and $\varepsilon_{m,d,h}$ as error term. We assume that the error process $(\varepsilon_{m,d,h})_{d\in\mathbb{Z}}$ is i.i.d. with constant variance $\sigma^2_{m,h}$. The introduced parameters $\phi_{m,h,l,j,k}$ will model the linear autoregressive impact and $\psi_{m,h,k}$ the day of the week effect.

The choice of index sets $\mathcal{I}_{m,h}(l,j)$ is crucial for the full model, as they specify the possible model structure. In general it holds true that larger sets $\mathcal{I}_{m,h}(l,j)$ increase the likelihood of overfitting, even though this likelihood is limited due to our used regularized estimation technique. However, if the index sets were chosen to small, we might miss important features in the data. Thus, we should always choose $\mathcal{I}_{m,h}(l,j)$ of reasonable size. This size is determined by the user and can be chosen freely or be backed up by fundamental data analysis, e.g. the correlation structure. Please note that this procedure only determines the possible lag structure and not the final lag structure as it only defines the set of lags which our estimation algorithm will consider. The coefficients that correspond to these lags can have zero impact because of the estimation procedure. For this paper we decided on

$$\mathcal{I}_{m,h}(l,j) = \begin{cases} \{1,2,\ldots,36\} & , m = l \text{ and } h = j \\ \{1,2,\ldots,8\} & , (m = l \text{ and } h \neq j) \text{ or } (m \neq l \text{ and } h = j) \\ \{1\} & , m \neq l \text{ and } h \neq j \end{cases}.$$

Thus, the process $Y_{m,d,h}$ of class m at day d and hour h can depend on the values of the past 36 days of class m at hour h. In contrast, $Y_{m,d,h}$ for a specific class m and a specific hour h is only allowed to depend on the value of another process at another hour one with a maximum lag of one. In all other cases a maximum lag of eight is possible. To illustrate this setting Figure 11 shows the possible dependency structure of $Y_{m,d,h}$ for an exemplary volume class of $h = 2$. The left hand side of the figure shows a specific price class, named as target class. The blue rectangle symbolizes the hour which was modeled, e.g. the hour 2:00. Every green rectangle gives information, if this lag is considered for modeling that hour. Red rectangles indicate that the lag is not considered as it lays outside of our lag definition, gray rectangles indicate that this data is not available as it is future information. Therefore, the target price class for one hour can be dependent on every other hour for that same price class up to eight days back in time, and on the same hour for the same price class up to 36 days back in time. The possible dependencies on other price classes is provided in the illustration in the middle. The allowed dependencies on planned regressors (generation, wind and solar) is depicted on the right hand side of the figure. The color scheme applies as well to those classes. It is worth mentioning that the model for hour 2:00 of a specific day and price class can be dependent on other hours of

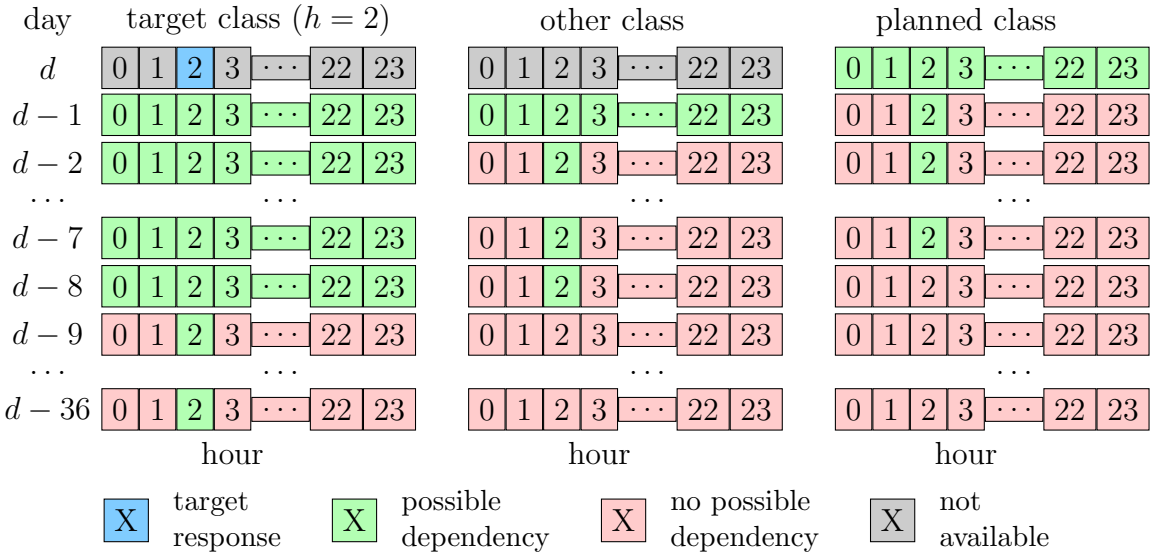


Figure 11: Illustration of the dependency structure for a target price class at hour 2.

the regressors for planned generation of that same day, as those information is indeed available before the auction starts. Besides that it can only depend on their lagged values for exactly hour 2:00 of the previous seven days. The shift in dependence on historical values is due to our definition of the regressors.

For the estimation of the parameters in (2) we use a method of high-dimensional statistics, namely the lasso estimation procedure, introduced by [21]. It was also used in a context of electricity price forecasting by [27].

The lasso estimator is a penalized least square estimator, thus we require the regression representation of model (2). Therefore, let the multivariate ordinary least squares representation of (2) be:

$$Y_{m,d,h} = \mathbb{X}_{m,d,h}\boldsymbol{\beta}_{m,h} + \varepsilon_{m,d,h}, \tag{3}$$

with the $p_{m,h}$-dimensional vector of regressors $\mathbb{X}_{m,d,h} = (\mathbb{X}_{m,d,h,1}, \ldots, \mathbb{X}_{m,d,h,p_{m,h}})'$ and $\boldsymbol{\beta}_{m,h} = (\boldsymbol{\beta}_{m,h,1}, \ldots, \boldsymbol{\beta}_{m,h,p_{m,h}})'$ as $p_{m,h}$-dimensional parameter vector. For the considered lasso estimation procedure it is also important that the regressors in (3) are standardized. Thus, we introduce with

$$\widetilde{Y}_{m,d,h} = \widetilde{\mathbb{X}}_{m,d,h}\widetilde{\boldsymbol{\beta}}_{m,h} + \widetilde{\varepsilon}_{m,d,h}, \tag{4}$$

the standardized version of equation (3). Here $\widetilde{Y}_{m,d,h}$ and the elements of $\widetilde{\mathbb{X}}_{m,d,h}$ are scaled so that they have a variance of 1. Given the scaled parameter vector $\widetilde{\boldsymbol{\beta}}_{m,h}$ we can easily reproduce $\boldsymbol{\beta}_{m,h}$ by rescaling. We estimate the scaled parameter vectors $\widetilde{\boldsymbol{\beta}}_{m,h}$ given n observable days by using the lasso estimator $\widehat{\widetilde{\boldsymbol{\beta}}}_{m,h}$:

$$\widehat{\widetilde{\boldsymbol{\beta}}}_{m,h} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{p_{m,h}}}{\arg\min} \sum_{d=1}^{n} (\widetilde{Y}_{m,d,h} - \widetilde{\mathbb{X}}_{m,d,h}\boldsymbol{\beta})^2 + \lambda_{m,h} \sum_{j=1}^{p_{m,h}} |\boldsymbol{\beta}_j| \tag{5}$$

where $\lambda_{m,h} \geq 0$ is a penalty parameter. Note that for $\lambda_{m,h} = 0$ we receive the common ordinary least square estimator and for sufficiently large $\lambda_{m,h}$ we get $\widehat{\widetilde{\boldsymbol{\beta}}}_{m,h} = \mathbf{0}$.

As estimation algorithm we use the coordinate descent approach of [9] which is a fast estimation procedure. The algorithm solves the lasso problem on a given grid $\Lambda_{m,h}$ of $\lambda_{m,h}$ values. This grid $\Lambda_{m,h}$ is usually chosen to be exponential decaying. Given a grid $\Lambda_{m,h}$, we select our optimal tuning parameter $\lambda_{m,h}$ by minimising the popular Baysian information criteria (BIC) which performs conservative model selection. However, other information criteria might be possible.

Given the estimated parameters $\widehat{\widetilde{\boldsymbol{\beta}}}_{m,h}$ for $\widetilde{\boldsymbol{\beta}}_{m,h}$, we calculate the lasso estimator $\widehat{\boldsymbol{\beta}}_{m,h}$ for $\boldsymbol{\beta}_{m,h}$ by rescaling. With $\widehat{\boldsymbol{\beta}}_{m,h}$ which contains the estimates for $\widehat{\phi}_{m,h,l,j,k}$ and $\widehat{\psi}_{m,h,k}$ we can compute a day-ahead point forecast by

$$\widehat{Y}_{m,n+1,h} = \sum_{l=1}^{M} \sum_{j=1}^{24} \sum_{k \in \mathcal{I}_{m,h}(l,j)} \widehat{\phi}_{m,h,l,j,k} Y_{l,n+1-k,j} + \sum_{k=2}^{7} \widehat{\psi}_{m,h,k} W_k(n+1).$$

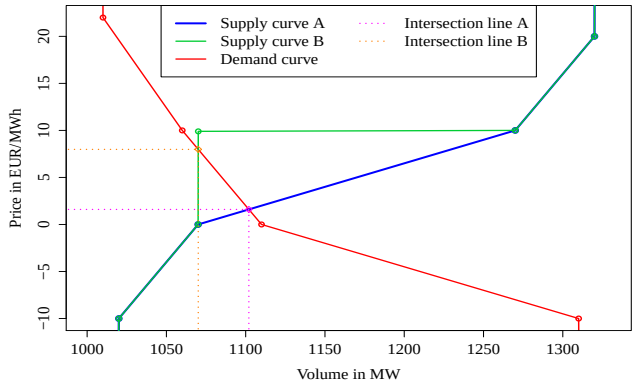
If we have the predicted values $\widehat{Y}_{1,n+1,h}, \ldots, \widehat{Y}_{M_S+M_D,n+1,h}$ we obtain the bidden volume forecasts $\widehat{X}_{1,n+1,h}, \ldots, \widehat{X}_{M_S+M_D,n+1,h}$ by adding the sample means. Using bootstrap methods we can compute B bootstrap samples $\widehat{X}^b_{m,n+1,h}$ for $b \in \{1, \ldots, B\}$ to receive a probabilistic estimate of $X_{m,n+1,h}$.

3.3 Reconstructing bids and price curves

After computing the forecast $\widehat{X}_{m,n+1,h}$ for each class $m \in \{1, \ldots, M_S+M_D\}$ and hour h we model the apportionment of the forecasted bidden volume $\widehat{X}_{m,n+1,h}$ for each price class. This will be useful for computing both point and probabilistic forecasts of the price curves. Especially, for the probabilistic forecast it is important to understand the bidding structure within a price class as we can use it for simulation methods. However, for forecasting the overall behavior, e.g. if we just want to see if there is a large probability for high prices, the reconstruction of the bids is not that relevant. We show that the reconstruction of the bids is relevant for the local price behavior, especially to explain price cluster.

For example, if we forecast the sale volume of the price class ranging from -5.1 EUR/MWh to 5.0 EUR/MWh to be 1000 MW, we have to redistribute this volume over the different price levels within that class, e.g. -5.1, -5.2, ..., 5.0, so that the real bidding behavior is captured well. In this example it is very likely that a significant amount of the 1000 MW is bidden at 0.0 EUR/MWh, as already explained in the previous section. Furthermore we have to take into account that many prices are not bidden at all. This is important due to the considered linear interpolation method for creating the price curves. So even a tiny bid of 0.1 MW can have a relative big impact on the electricity price. This holds for both, bids on the supply and demand side. As this procedure is crucial for our analysis, we briefly discuss this issue in a toy example for a minor change in the supply bidding structure. Therefore we consider two scenarios, A and B, for the supply curve and keep the demand curve constant. The scenario A differs only marginal from the scenario B in the bidding structure. In A there are 100 MW offered at 10 EUR/MWh, whereas in B 99.9 MW were bidden at 10 EUR/MWh and 0.1 MW are bidden at 9.9 EUR/MWh. The detailed assumed bids and the corresponding price curves are given in Figure 12. There we can observe that the supply curve in scenario B looks more rectangular. Indeed, also judging by our dataset market participants on the sale and purchase side seem to aim for a price curve that is close to a step function. In our short example, scenario A results in a market clearing price of 1.60 at a volume of 1102.0, whereas the market clearing price in B is 7.98 at a volume of 1070.1. The market clearing price is 6.38 EUR/MWh higher. This shows exemplarily, that minor change in the bidding structure can cause a severe price change, especially in price areas with only some bids, e.g. very large or very small (negative) prices. Even though this is a toy example it is surprisingly real. We can assume that the described behavior is known by at least some market participants, as we can observe that some agents try to strategically chose their bids to achieve the rectangular shape of the function.

To take into account whether a certain price is traded or not, we have to model the probability of that event. We will refer to this approach as "reconstructing" throughout our paper. Thus, let $X^S_{d,h}(P)$ and $X^D_{d,h}(P)$ denote the bidden volume for the supply and demand at price $P \in \mathbb{P}$. Further remember that $\mathbb{P}_S(c)$ and $\mathbb{P}_D(c)$ are the prices within the price class of price c. We can



(a) Supply and Demand curves with the corresponding market clearing prices.

Supply, Scenario A

Price	-500	-10	0	10.0	20	3000
Volume	1000	20	50	200.0	50	70

Supply, Scenario B

Price	-500	-10	0	9.9	10.0	20	3000
Volume	1000	20	50	0.1	199.9	50	70

Demand

Price	3000	22	10	0	-10	-500
Volume	1000	10	50	50	200	20

(b) Bidding structure of the supply and demand curves.

Figure 12: Toy example for two supply szenarios A and B

express the volumes $X^{S,c}_{d,h}$ and $X^{S,c}_{d,h}$ of the price classes by

$$X^{S,c}_{d,h} = \sum_{P \in \mathbb{P}_S(c)} X^S_{d,h}(P) \quad \text{and} \quad X^{D,c}_{d,h} = \sum_{P \in \mathbb{P}_D(c)} X^D_{d,h}(P).$$

Let $\pi^S_{d,h}(P)$ and $\pi^D_{d,h}(P)$ be the probabilities that $X^S_{d,h}(P)$ and $X^D_{d,h}(P)$ respectively is greater than zero, so there is actually a bid at this price. We assume these probabilities for the bids are constant over time. We simply estimate $\pi^S_{d,h}(P)$ and $\pi^D_{d,h}(P)$ by the relative frequency $\widehat{\pi}^S_{d,h}(P)$ and $\widehat{\pi}^D_{d,h}(P)$ in the given sample.

Furthermore, we assume proportionality within the bidden prices in the price classes with respect to the mean volume $\mathbb{V}_S(P)$ and $\mathbb{V}_D(P)$. Then we can express $X^S_{d,h}(P)$ and $X^D_{d,h}(P)$ by

$$X^S_{d,h}(P) = R_S(P) V^S_{d,h}(P) \quad \text{with} \quad V^S_{d,h}(P) = \frac{\mathbb{V}_S(P)}{\sum_{Q \in \mathbb{P}_S(c)} R_S(Q) \mathbb{V}_S(Q)} X^{S,c}_{d,h} \tag{6}$$

$$X^D_{d,h}(P) = R_D(P) V^D_{d,h}(P) \quad \text{with} \quad V^D_{d,h}(P) = \frac{\mathbb{V}_D(P)}{\sum_{Q \in \mathbb{P}_D(c)} R_D(Q) \mathbb{V}_D(Q)} X^{D,c}_{d,h} \tag{7}$$

where $V^S_{d,h}(P)$ and $V^D_{d,h}(P)$ are the bidden volumes that are related to price P, c is the class associated with price P and $R_S(P) \sim \text{Ber}(\pi^S_{d,h}(P))$ as well as $R_D(P) \sim \text{Ber}(\pi^D_{d,h}(P))$ are Bernoulli random variables with probabilities $\pi^S_{d,h}(P)$ and $\pi^D_{d,h}(P)$. We assume that the Bernoulli random variables $R_S(P)$ and $R_D(P)$ are independent from each other over the full price grid.

As we have estimates for the probabilities of the Bernoulli random variables $\pi^S_{d,h}$ and $\pi^D_{d,h}$ and the mean bidden volumes $\mathbb{V}_D$ and $\mathbb{V}_S$ we can easily simulate $X^S_{d,h}(P)$ and $X^D_{d,h}(P)$ by equations (6) and (7) given the price class forecast of $X^{S,c}_{d,h}$ and $X^{D,c}_{d,h}$ from the time series model. These simulations can be utilized to construct probabilistic forecasts later on. If we only want to receive point estimates for $X^S_{d,h}(P)$ and $X^D_{d,h}(P)$ we recommend to set $R_S(P)$ and $R_D(P)$ to one, if $\widehat{\pi}^S_{d,h}(P)$ and $\widehat{\pi}^D_{d,h}(P)$ are greater than a certain threshold and to zero otherwise. For our purpose we will consider the probability threshold of 1/12. So in our point forecasts a price is active if it occurred at least twice a day.

Similarly to equation (1) we can calculate the volumes $X^{S,\text{cum.}}_{d,h}(P)$ and $X^{D,\text{cum.}}_{d,h}(P)$ associated with the price curves given the volumes $X^S_{d,h}(P)$ and $X^D_{d,h}(P)$ for the full price grid $\mathbb{P}$ by accumulating

$$X^{S,\text{cum.}}_{d,h}(P) = \sum_{\substack{p \in \mathbb{P} \\ p \leq P}} X^S_{d,h}(p) \quad \text{and} \quad X^{D,\text{cum.}}_{d,h}(P) = \sum_{\substack{p \in \mathbb{P} \\ p \geq P}} X^D_{d,h}(p).$$

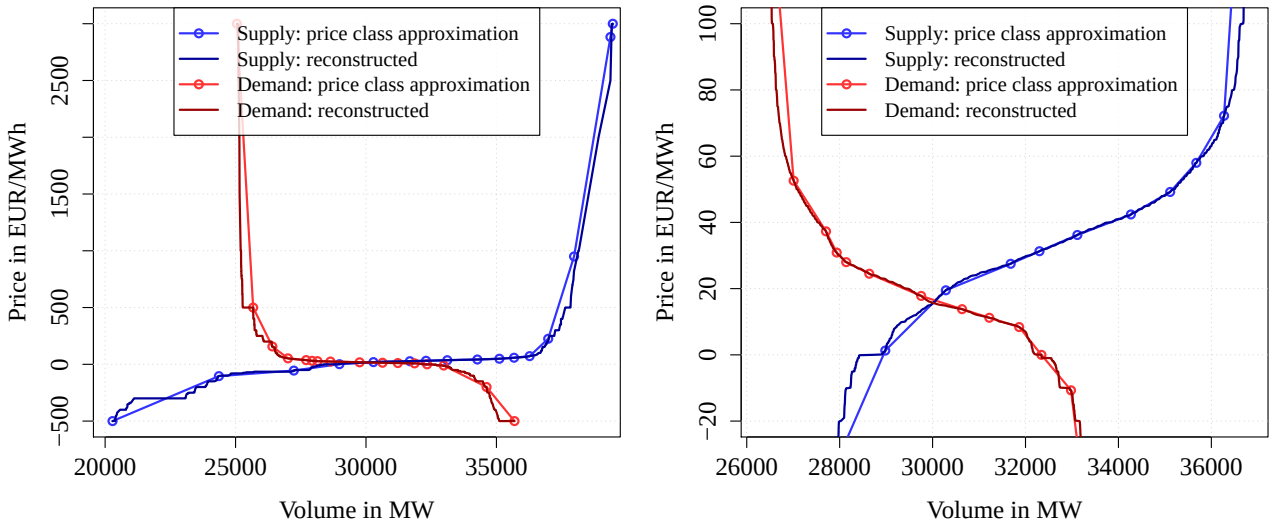
However, $X^{S,\text{cum.}}_{d,h}$ and $X^{D,\text{cum.}}_{d,h}$ do not represent the price curves itself, as they are given by the linear interpolation of all volume and price combinations that have an increase in the volume.

In other words, only points with an actual non-zero volume bid were chosen and linearly interpolated. Formally the vector of the corresponding prices of these points $\mathbb{P}_S$ and $\mathbb{P}_D$ are given by the elements of the sets $\{P \in \mathbb{P} | R_S(P) = 1\}$ and $\{P \in \mathbb{P} | R_D(P) = 1\}$. We assume that $\mathbb{P}_S = (\mathbb{P}_{S,1}, \ldots, \mathbb{P}_{S,J_S})$ is ordered monotone increasing and $\mathbb{P}_D = (\mathbb{P}_{D,1}, \ldots, \mathbb{P}_{D,J_D})$ monotone decreasing. Note that the first element of $\mathbb{P}_S$ will be always -500 and the last 3000, for $\mathbb{P}_D$ vice versa. Then we can write down the explicit formula of the supply and demand curves $S_{d,h}(V)$ and $D_{d,h}(V)$ given a volume V:

$$S_{d,h}(V) =$$
$$\begin{cases} -500 & , V \leq X^{S,\text{cum.}}_{d,h}(\mathbb{P}_{S,1}) \\ \frac{\mathbb{P}_{S,j} - \mathbb{P}_{S,j-1}}{X^{S,\text{cum.}}_{d,h}(\mathbb{P}_{S,j}) - X^{S,\text{cum.}}_{d,h}(\mathbb{P}_{S,j-1})} \left(V - X^{S,\text{cum.}}_{d,h}(\mathbb{P}_{S,j-1}) \right) & , X^{S,\text{cum.}}_{d,h}(\mathbb{P}_{S,j}) \leq V < X^{S,\text{cum.}}_{d,h}(\mathbb{P}_{S,j-1}) \\ 3000 & , V > X^{S,\text{cum.}}_{d,h}(\mathbb{P}_{S,J_S}) \end{cases},$$

$$D_{d,h}(V) =$$
$$\begin{cases} 3000 & , V \leq X^{D,\text{cum.}}_{d,h}(\mathbb{P}_{D,1}) \\ \frac{\mathbb{P}_{D,j} - \mathbb{P}_{D,j-1}}{X^{D,\text{cum.}}_{d,h}(\mathbb{P}_{D,j}) - X^{D,\text{cum.}}_{d,h}(\mathbb{P}_{D,j-1})} \left(V - X^{D,\text{cum.}}_{d,h}(\mathbb{P}_{D,j-1}) \right) & , X^{D,\text{cum.}}_{d,h}(\mathbb{P}_{D,j}) \leq V < X^{D,\text{cum.}}_{d,h}(\mathbb{P}_{D,j-1}) \\ -500 & , V > X^{D,\text{cum.}}_{d,h}(\mathbb{P}_{D,J_D}) \end{cases}.$$

The intersection of $S_{d,h}$ and $D_{d,h}$ will give us the market clearing price at day d and hour h with the corresponding volume.



(a) -500 to 3000 (b) -20 to 100

Figure 13: Example of classified and reconstructed price curves for selected price ranges.

In Figure 13 illustrates the resulting reconstructed supply and demand curve with its plain classified counterpart for a selected example. As mentioned, in the main price region around 20 to 50 with many bids the difference is marginal. But in uncommon price regions, e.g. around 0, the impact is larger. In general the reconstructed price curves look much more realistic than the grouped versions.

4 Empirical results

In order to show the results of our X-Model under real world conditions, we performed an out-of-sample analysis for the time period from 01.11.2014 to 19.04.2015. To evaluate our results, we compare our model with the results of standard models and models used frequently in the literature. Additionally, we show a detailed forecasting analysis for three days namely the

19.12.2014, 24.03.2015 and 12.04.2015. We chose those days for the following reasons. The first day is suitable to show how price clusters can be predicted. The second day and third day are these days in the selected out-of-sample data range with the largest positive and negative price spike respectively. All in all, these days are also suitable to show all important features of the model, even though they do not have by far the best point forecast performance. The detailed forecasting results of all considered days can be found in the appendix.

For estimation and forecasting we use for all days the previous $730 = 2 \times 365$ days (2 years) of data. We forecast the supply and demand curves and compute the corresponding market clearing price and volume as described in the previous section. For receiving probabilistic forecasts we perform residual based bootstrap with a bootstrap sample size of $B = 5000$. Subsequently, all prediction intervals are symmetric, in the sense that $\alpha/2$ of the probability mass is below an $(1-\alpha) \times 100\%$-prediction interval and $\alpha/2$ is above. First, we will discuss the forecasted results for the market clearing price and volume of the beforementioned three selected days. This is followed up by the results for the forecasted price curves of some hours of the 12.04.2015. Finally we will show an out-of-sample forecasting study for market clearing price over the whole forecasting period

For comparing our results regarding the probabilistic forecast, we consider two benchmarks. As a simple benchmark we take the weekly persistent model, sometimes called naive model, given by

$$X_{d,h}^{\text{price}} = X_{d-7,h}^{\text{price}} + \varepsilon_{d,h} \quad \text{with} \quad \varepsilon_{d,h} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_h^2). \tag{8}$$

Furthermore, we take a more advanced regime switching model that is in principle able to cover price spikes. The model, is very close to the one used in [12]. It is a Markov switching model and is given by

$$X_{d,h}^{\text{price}} = \mathbb{X}_{d,h} \boldsymbol{b}_{s(d,h)} + \varepsilon_{s(d,h),h} \quad \text{with} \quad \varepsilon_{s(d,h),h} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{s(d,h),h}^2) \tag{9}$$

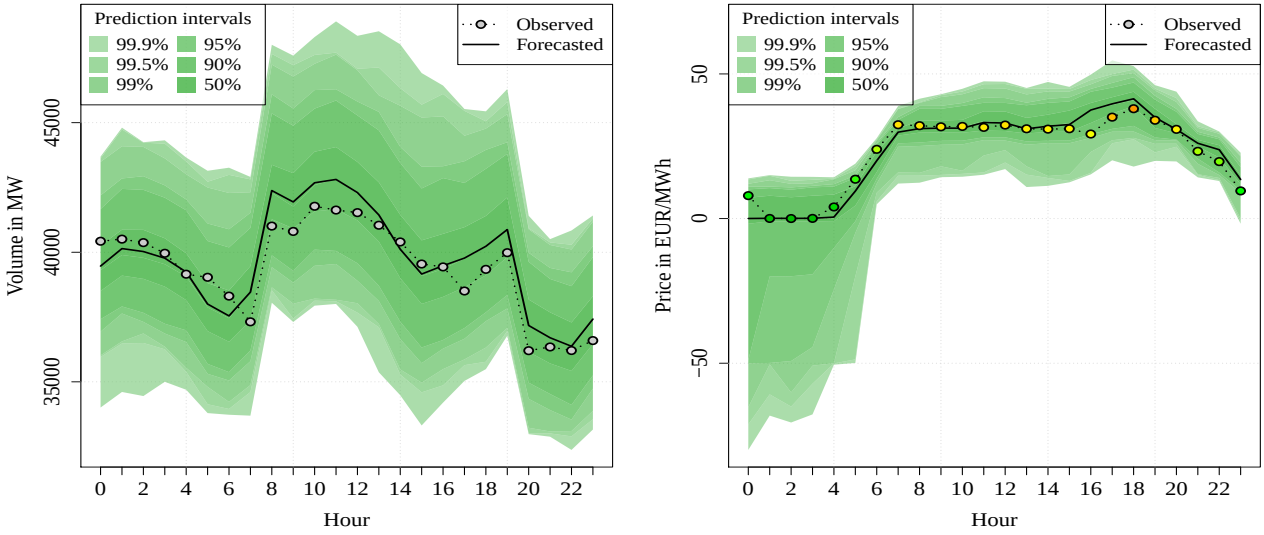
with $\mathbb{X}_{d,h} = (1, X_{d-1,h}^{\text{price}}, X_{d-7,h}^{\text{price}}, \overline{X}_{d-1}^{\text{price}}, X_{d,h}^{\text{generation}}, X_{d,h}^{\text{wind}}, X_{d,h}^{\text{solar}})$, parameter vector $\boldsymbol{b}_{s(d,h)}$, transition probabilities $p_{i,j} = P(s(d,h) = i | s(d-1,h) = j)$ and $s(d,h)$ as the latent regime at day d and hour h with $s_{\max}$ possible states. Here $\overline{X}_{d-1}^{\text{price}}$ is the mean price of the last day. Note that the solar component is not included for the hours from 0:00, 1:00, 2:00, 3:00 and 23:00 as there is no solar energy produced during night. We estimate the regime switching model with $s_{\max} = 2$ regimes by maximizing the likelihood with the EM-algorithm.

The results of the X-Model for the market clearing priceand the volume of the three chosen days are given in Figure 14. The price forecast of the two benchmarks are given in Figure 15. For the volume forecasts in 14 we see no special behavior, the prediction intervals seem to map the daily pattern well. The observations lie all clearly within the 5% prediction bands. Thus, it seems that the method provides reliable forecasting results for the volume of these days.

More interesting are the results of the X-Model for the market clearing prices, in figures 14b, 14d and 14f. There we observe the distinct non-linear behavior of the prices. For small (especially negative) prices in 14b and 14f we see clearly left skewed prediction densities. Similarly we have noticeable right-skewed prediction densities for large prices as in 14d. Therefore, the information of previous auction data seems to capture the increased likelihood of extreme price events very well.

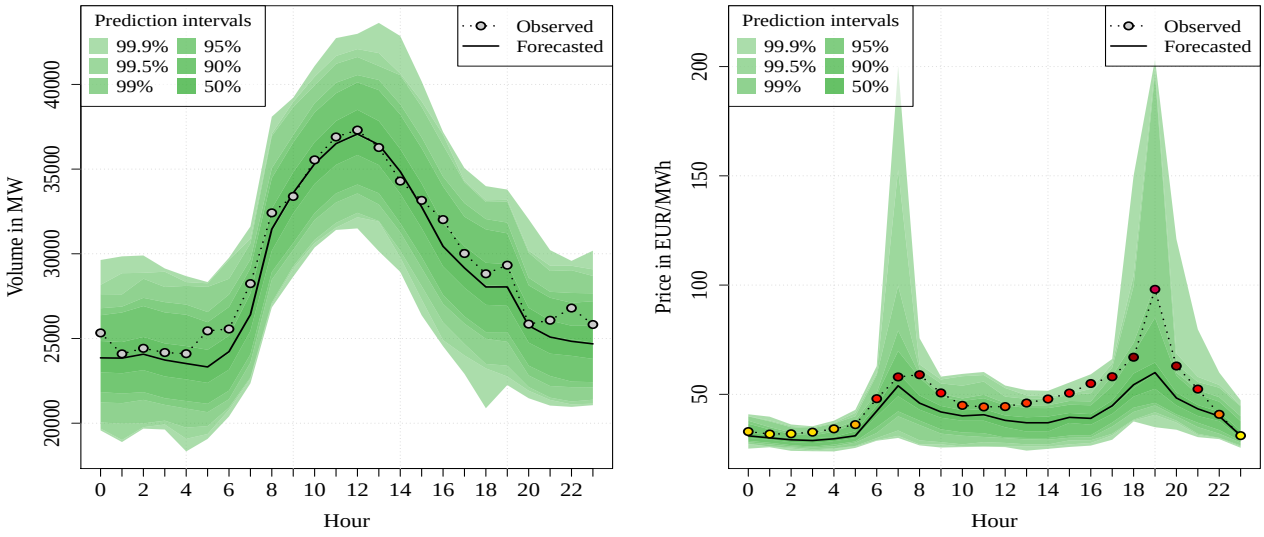
In Figure 14b we observe that the beforementioned price clustering at different integer price levels, e.g. at 0, can be modeled by this forecasting method. For the first four hours of the day the point forecast for the electricity price were extremely close to zero. So it was relatively likely to receive values at the price cluster around 0. And indeed, three of the four market clearing prices were in this price cluster, namely 0.05 at 1:00, 0.02 at 2:00 and 0.07 at 3:00.

The forecast plot 14f for the 12.04.2015 is also suitable to highlight the difference between common statistical outliers, i.e. random events that can happen, but are extremely rare, and



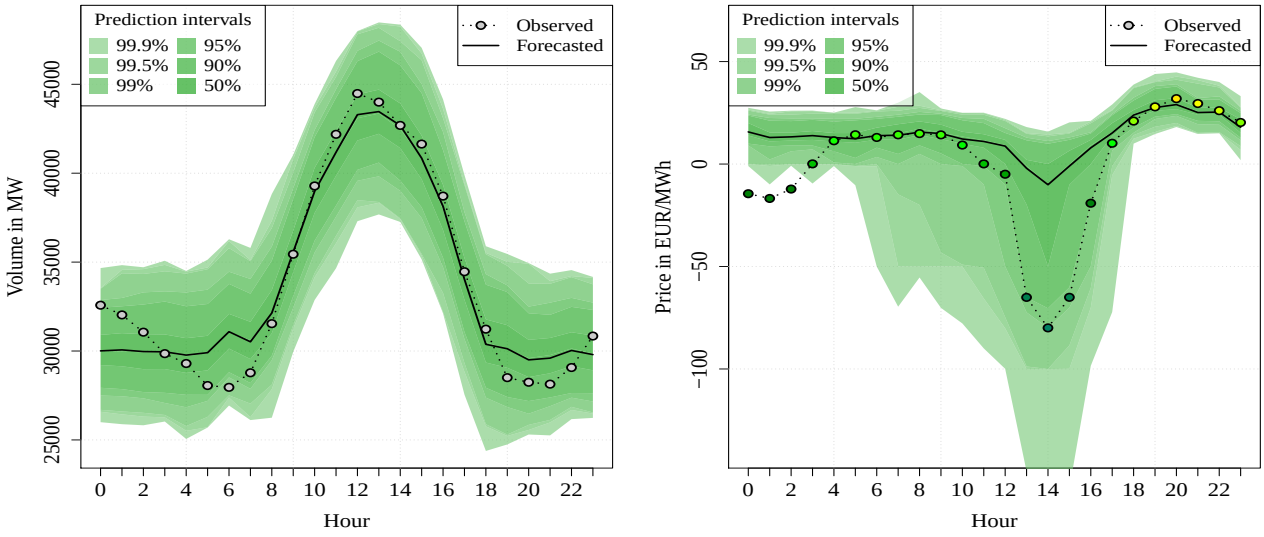
(a) Volume forecast

(b) Price forecast



(c) Volume forecast

(d) Price forecast

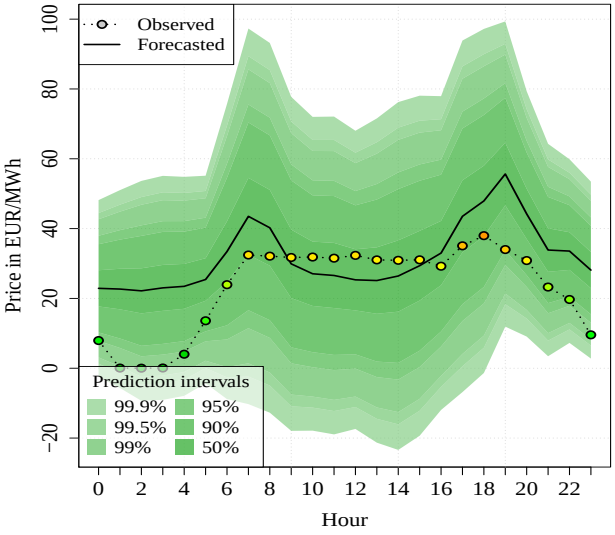


(e) Volume forecast

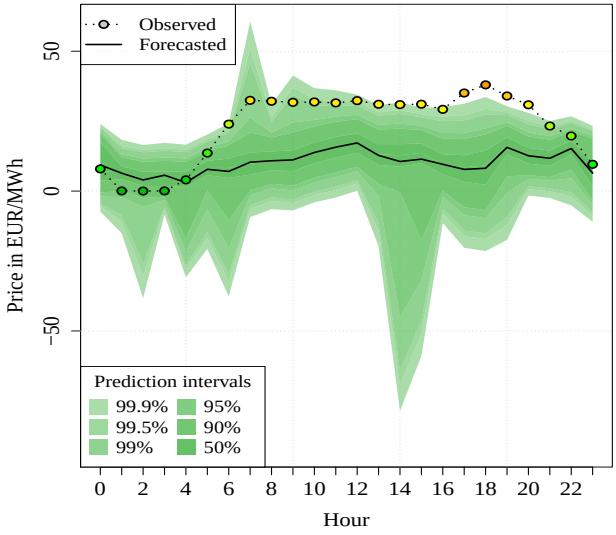
(f) Price forecast

Figure 14: Volume and Price Forecast with prediction intervals and observed values for the 19.12.2014, 24.03.2015 and 12.04.2015

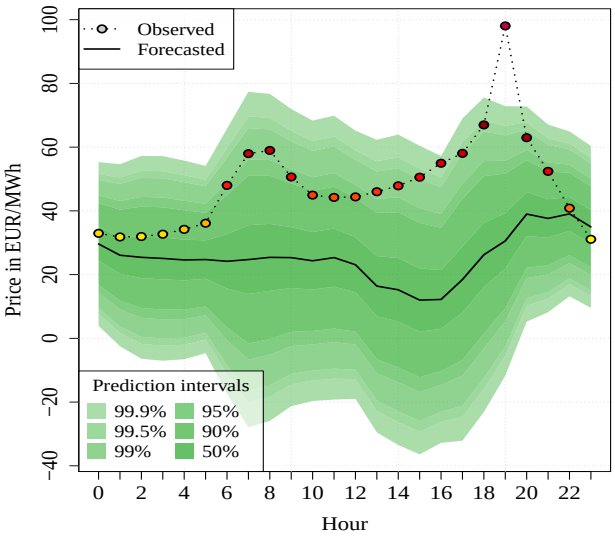
price spikes that are predictable in the sense that the probability for such an event is relatively large given the available information. The 12.04.2015 was a Sunday, one week after the Easter



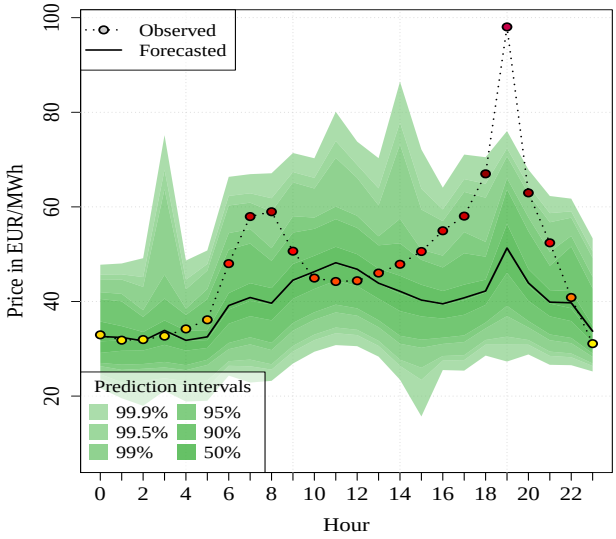
(a) Price forecast of persistent model



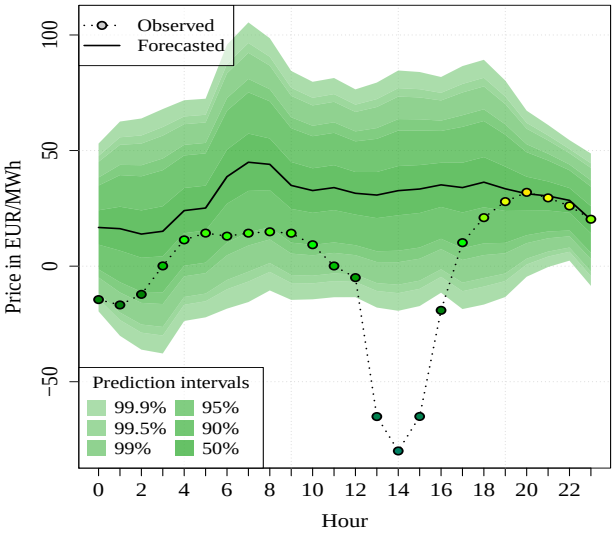
(b) Price forecast Markov-Chain-Switching



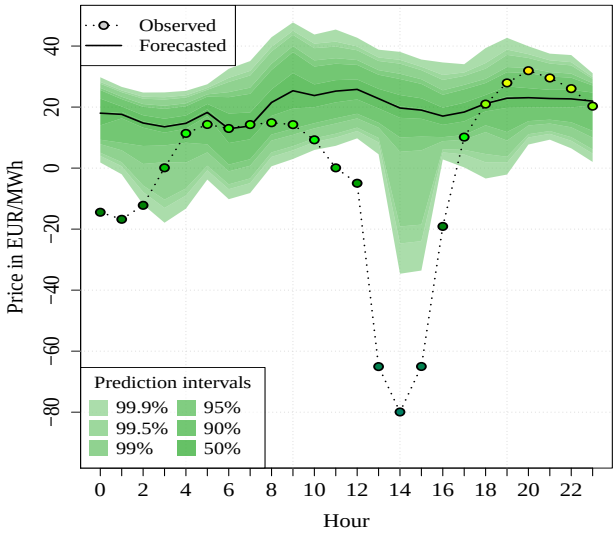
(c) Price forecast of persistent model



(d) Price forecast Markov-Chain-Switching



(e) Price forecast of persistent model



(f) Price forecast Markov-Chain-Switching

Figure 15: Price Forecast of considered benchmarks with prediction intervals and observed values for the 19.12.2014, 24.03.2015 and 12.04.2015

holidays. But the 12th April opened with a clearly negative price of -14.47 at 0:00 and reached values between -79.94 to 31.93 during the day. The prices of the past week were all in the range

of 12.00 to 69.03 with the last observation on 11.04.2015 23:00 at 22.11. Thus, it is usually very complicated to forecast a realistic likelihood for such negative prices with an autoregressive approach. However, our X-Model, which focuses on auction data, seems to have recognized the pattern within the data and provided a realistic confidence interval nonetheless. Regarding the prediction bands in 14f we see clear changes over the day. It starts quite narrow at 0:00, becomes significantly wider and more left-skewed at around 5:00. This peaks at 14:00 where the observed price also reaches its daily minimum. Afterwards the prediction intervals become smaller and more symmetric as the forecasts moves closer to common price levels. However, for the first hours of the day the negative prices are not predicted by the X-Model. Thus, we have classical outliers. The benchmark models in 15e and 15f suffer from the same problem. However, it is remarkable that the X-Model predicts a quite large probability for negative prices for the morning and afternoon hours, especially from 13:00 to 15:00. For instance, the clear negative price at 13:00 with -65.06 lies clearly within the 5% prediction intervals. Both benchmark models in Figure 15e and 15f were not able to predict these price spikes well. Many standard electricity price models only allow for errors where the shape of the density does not depend on the predicted value. As in the persistent model the shape of the prediction density is kept constant and simply gets shifted and scaled over time. Thus they are definitely not suitable to capture the real underlying behavior. In contrast, the regime switching model in 15f is in general able to cover price spikes, as the forecast density is a mixture density. We see that at 14:00 and 15:00 the prediction density becomes left skewed, which provides clear indication for a price spike. However, the magnitude is not well predicted.

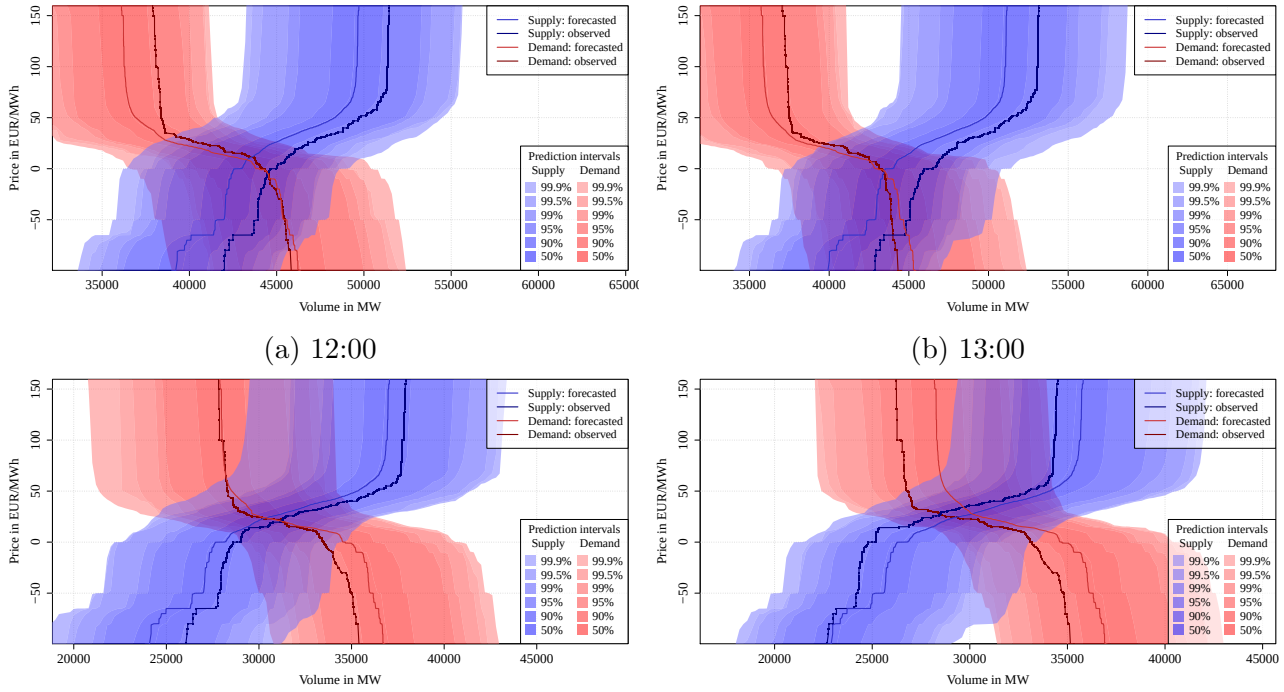
The largest weakness of all models known so far that are designed for modeling such price spikes is that they use only the information of the observed past market clearing prices and related processes like wind and solar energy. The amount of historical extreme prices, which are considered by most common models, is typically very low as they occur only rarely. Hence, such models often simply have too little datapoints to learn from the behavior at these price levels.

The X-Model on the other hand uses the bidding information from all time points in all price regions. Thus, it can learn a lot about the price behavior in every price region, even for market clearing prices, which were never realized so far.

In general, Figure 14 shows that the X-Model adopts the non-linear shape of the price curves and hands it down to the forecasts. This automatically adjusts the shape of the prediction densities.

Moreover, we are able to perform forecasts and compute prediction intervals for the full price curves. In Figure 16 we exemplarily plot the forecast for the four selected hours that we discussed in the introduction. Figures 16a and 16b show a forecast at 12:00 and 13:00 where the realized price dropped from -4.96 to -65.06. In Figures 16c and 16d we have the price curves at 19:00 and 20:00 where the market clearing price increased from 27.92 to the hightest value of that day 31.93. Remember that in the 12:00 and 13:00 case in 14f the prediction densities of the market clearing price were highly left skewed and in the 19:00 and 20:00 case relatively symmetric. Both graphs of 16 show additionally to the forecasted price curves with its prediction intervals the realized supply and demand curves of the actual auction. Note that we only show the most relevant price region between -100 and 150.

For the 13:00 case with the clear negative price, the observed demand and supply curves lie within the relatively narrow 90% prediction interval. But this does not mean that the market clearing price lies in the 90% prediction interval as well. The reason is that both confidence intervals have a complex dependence structure. In fact, the observed price in 16b lies only in the 95% prediction interval. However, we can see quite well that the predicted intersection is at a region where both, the supply and demand curve, have a relatively large absolute slope. The magnitude of the slope even increases for more negative values. This is the reason for the clearly left-skewed prediction density, because a relatively moderate increase in the supply curve or decrease in the demand curve causes relatively large price movements. In economic terms this



(a) 12:00 (b) 13:00

(c) 18:00 (d) 19:00

Figure 16: Supply and demand curves forecasts with prediction bands for the 12.04.2015 and selected hours.

coincides with a situation where both the supply and the demand side is relatively inelastic in the negative price region. This induces high price volatility.

At 19:00, where the market clearing price turned out to be relatively high, the prediction intervals look similar in general, but also have important differences in the detail. Here the demand side is still elastic at the region close to the market clearing price. However, for a price level of around 30 to 40 the demand side also seems to loose elasticity quite dramatically. The supply side is still quite elastic up to a price region of 50 to 60. Small volume shocks in the bidding structure are therefore likely to be compensated by the elastic supply resulting in a high likelihood of small price changes to occur. However, for medium to large volume shocks the intersection might be shifted out of the quite elastic area of the demand curve. This is detected by our model and indicated by a large volatility with a clear right-skewness of the confidence interval, as can be also obtained from Figure 14b.

Now we want to compare the point forecast of the proposed X-Model in the out-of-sample region from 01.11.2014 to 19.04.2015 with several common benchmarks. Even though the model is primarily designed to detect and model extreme price events with the corresponding prediction densities, it is interesting to see the performance purely based on standard error measures in comparison to other established electricity models.

Denote $\widehat{X}^{\text{price}}_{d,h}$ the predicted point forecast of a electricity price model at day d and hour h that corresponds to $X^{\text{price}}_{d,h}$. Further we denote by $\mathcal{D}$ the set of all days from 01.11.2014 to 19.04.2015, except for the 29.03.2015 which we ignore here as it is the day where the time was switched due to daylight saving time. So $\mathcal{D}$ contains in total $\#(\mathcal{D}) = 169$ days. We define the common error measures, e.g. the absolute mean absolute error (MAE_h) at hour h and the root mean square error (RMSE_h) at hour h by

$$\text{MAE}_h = \frac{1}{\#(\mathcal{D})} \sum_{d \in \mathcal{D}} |X^{\text{price}}_{d,h} - \widehat{X}^{\text{price}}_{d,h}| \quad \text{and} \quad \text{RMSE}_h = \sqrt{\frac{1}{\#(\mathcal{D})} \sum_{d \in \mathcal{D}} |X^{\text{price}}_{d,h} - \widehat{X}^{\text{price}}_{d,h}|^2}. \tag{10}$$

Both measures are suitable to compare point-forecasts of different models at a certain h. Simi-

larly to the MAE_h and RMSE_h, we define the overall MAE and RMSE by

$$\text{MAE} = \frac{1}{24\#(\mathcal{D})} \sum_{d\in\mathcal{D}} \sum_{h=0}^{23} |X_{d,h}^{\text{price}} - \widehat{X}_{d,h}^{\text{price}}| \quad \text{and} \quad \text{RMSE} = \sqrt{\frac{1}{24\#(\mathcal{D})} \sum_{d\in\mathcal{D}} \sum_{h=0}^{23} |X_{d,h}^{\text{price}} - \widehat{X}_{d,h}^{\text{price}}|^2}.$$

In general the MAE is more robust than the RMSE, as the ladder is by far more sensitive to outliers.

The first two benchmarks we consider are the persistent model (Persistent) given in equation (8) and the regime switching model as presented in equation (9) with $s_{\max} = 2$ (Regime). The next simple benchmark that we consider is a very powerful one in terms of MAE. It uses different information as our model, namely the electricity price from the Energy Exchange Austria (EXAA). This is an electricity price for Germany and Austria with the same zones for physical settlement as the German and Austrian EPEX spot price. It is traded everyday at 10:12 and the prices are known at 10:20 for market participants, which means they are especially known in advance to the EPEX auction at 12:00. [28] show that the very simple naive estimator $\widehat{X}_{d,h}^{\text{price}} = X_{d,h}^{\text{EXAA,price}}$ with $X_{d,h}^{\text{EXAA,price}}$ as EXAA electricity price at day d and hour h is very competitive. However, the EXAA benchmark model (EXAA) is basically beyond the competition, as it uses information which we did not explicitely include in our X-Model. But still, it can help to gain insights about possible improvements. Furthermore, we introduce two AR(p) based models, namely a univariate on $X_{\text{price},t}$ (AR(p)) and a 24-dimensional model with 24 simple univariate AR models on $X_{d,h}^{\text{price}}$ for each hour h (24-dim. AR). They are formally defined by

$$X_{\text{price},t} = \phi_0 + \sum_{k=1}^{p} \phi_k X_{\text{price},t-k} + \varepsilon_t \quad \text{with} \quad \varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma^2),$$

$$X_{d,h}^{\text{price}} = \phi_{h,0} + \sum_{k=1}^{p_h} \phi_{h,k} X_{d-k,h}^{\text{price}} + \varepsilon_{d,h} \quad \text{with} \quad \varepsilon_{d,h} \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma_h^2).$$

We estimate the AR models by solving the Yule-Walker equations. The optimal orders p and p_h are determined by minimizing the Akaike Information Criterion (AIC) on a grid of possible orders. For the univariate AR model we search the optimal p on $\{1,2,\dots,700\}$ which allows for dependencies of more than four weeks. For the 24-dimensional model the optimal order p_h is searched on $\{1,2,\dots,50\}$, which allows for a memory of up to seven weeks and one day.

Furthermore, we consider two more models from the literature, a wavelet based model and a more advanced time series approach. The wavelet based approached is basically the popular wavelet-ARIMA model introduced by [6]. We use Daubechies 4 wavelet decomposition and model the coefficients of the wavelet decomposition by an ARIMA(12,1,1). The second benchmark model is a time series based approach that is analyzed by [13]. We select the ARMA(5,1) model with a trend component as well as their sophisticated annual, weekly and daily seasonal components. The model is suggested as one of the best models in the comparison study by [13]. We refer to the two models as Conejo et al. and Keles et al. respectively.

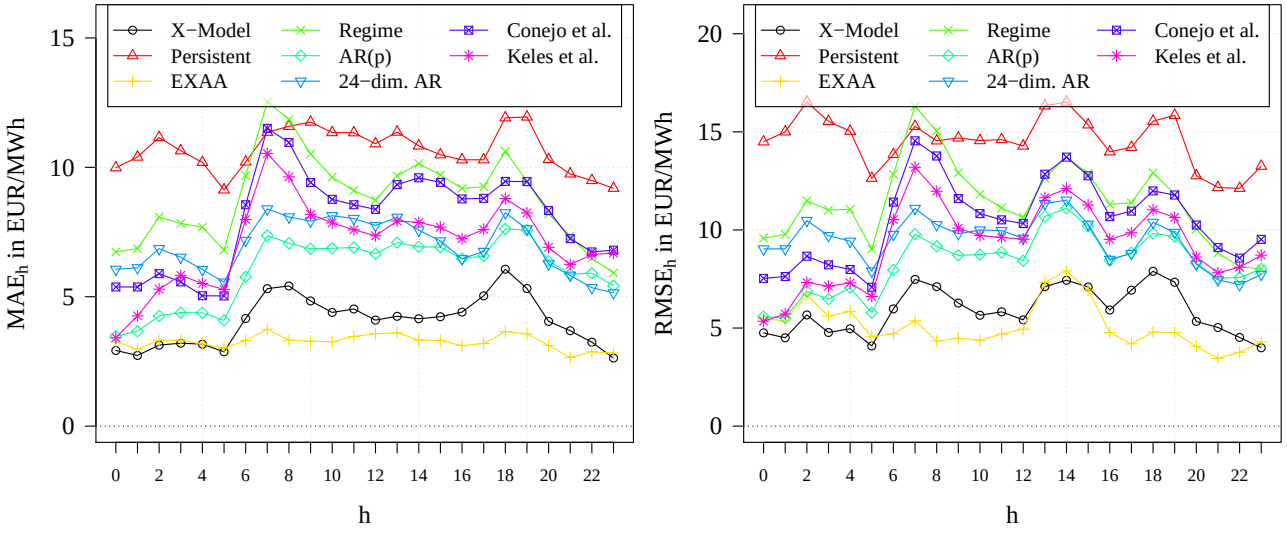
The estimated MAE and RMSE values of all considered models with their estimated standard deviations are given in Table 1. The hourly MAE_h and RMSE_h for all models are visualized in Figure 17.

There we observe that the proposed X-Model performs surprisingly well, even though is not designed to model the electricity price directly. With a MAE of 4.08 and RMSE of 5.99 it clearly outperforms all considered models, except the EXAA model with a MAE of 3.26 and RMSE of 5.23. In the night hour from 0:00 to 5:00 as well at 23:00 the X-Model seems to outperform every other model under consideration.

The out-of-sample MAE proportion of the X-Model in comparison to the persistent model is about 38.2%. The second best model which uses the same information as the X-Model is the AIC-selected univariate AR with a relative MAE proportion of 56.4%. Here the MAE is in absolute value 1.94 larger than the MAE of the X-Model.

Models	MAE (std.dev.)	% of persistent	RMSE (std.dev.)	% of persistent
X-Model	4.08 (0.068)	38.2	5.99 (0.185)	41.0
Persistent	10.66 (0.159)	100.0	14.60 (0.240)	100.0
Regime	8.83 (0.117)	82.9	11.60 (0.197)	79.5
EXAA	3.26 (0.065)	30.6	5.23 (0.303)	35.8
AR(p)	6.02 (0.097)	56.5	8.43 (0.238)	57.8
24-dim. AR	6.96 (0.103)	65.3	9.55 (0.219)	65.4
Conejo et al.	8.02 (0.112)	75.3	10.72 (0.213)	73.4
Keles et al.	7.11 (0.099)	66.7	9.53 (0.219)	65.3

Table 1: MAE and RMSE in EUR/MWh of the X-Model and several benchmark models



(a) MAE$_h$ (b) RMSE$_h$

Figure 17: MAE$_h$ and RMSE$_h$ for $h \in \{0, \dots, 23\}$ for the considered models.

5 Results and conclusion

We present a model for the day-ahead electricity spot price by directly modeling the supply and demand curves. We call our model the X-Model, as we estimate the market clearing price as the intersection of the sale and purchase curve of the German-Austrian day-ahead electricity market of the EPEX. Simple dimension reduction techniques and high-dimensional statistical methods allow us to deal with the huge amount of bid data. We group the possible bidden prices to price classes and assume a linear model for the bidden volume for every of them. Afterwards we forecast the bidden volumes in the price classes, reconstruct the sale and purchase curves and receive the corresponding market clearing price.

Our empirical results show that it is possible to model the electricity prices using such an approach in a very promising way. We can capture known stylized effects of the electricity price, like daily and weekly seasonalities, very well and are also able to model the newly elaborated stylized facts of price bids. The complex bidding structure for day-ahead prices allows us to model and predict extreme and rare price events by estimating realistic prediction densities for the market clearing price. The conducted out-of-sample study shows that the introduced model clearly outperforms standard methods and even very well performing methods of the recent literature in terms of densities as well as error measures like MAE and RMSE. Especially the ladder was stunning and remarkable to us, as the model approach is relatively simple in its core and mainly developed for the purpose of modeling extreme price events.

The provided X-Model approach opens the door to many other different applications, especially those related to policy making. One very important issue is for example the impact of

market regularizations. Many countries provide subsidies for renewable energy. This causes automatically the so called merit-order effect on the corresponding electricity markets. There are many papers (e.g. [19], [18], [5]) that aim for estimating these effects. With a sell and purchase curve based approach we can directly model the impact of the renewable energy. The only condition which must be met is the availability of data, like for the German and Austrian market. The advantage is that the sale and purchase curve based approach directly takes into account the market behavior with all its complex dependencies and non-linear properties. Common model approach are hardly able to cover such behavior.

Another important application could be the evaluation of the price effect by closing a power plant, e.g. due to a phase-out of nuclear or lignite based power plants. By proposing just a few assumptions for the bidding behavior and the fuel costs it is possible to postulate a proper model for the electricity market. This would allow the researcher to get realistic price forecasts that can be utilized from decision-makers. Note that such forecasts could be achieved for more than one day ahead. Given a proper model design even long run studies of several months and years are possible. This could be combined with different scenarios for related indicators like GDP growth or fuel costs. Moreover, the paper with its proposed model can support the dialog of two model disciplines in electricity price modeling. At the moment there are classical statistical, time series and machine learning techniques that forecast the market clearing price based on observations and related time series. The other model approaches are (multi-)agent based electricity price models, which usually ignores real auction data and analyzes the electricity market from a theoretical point of view. As this paper uses standard statistical tools to model and forecast actual bids, it may provide a good starting point for increased communication between representatives of both model disciplines.

Future research should also improve the considered model for the time series of bids. First of all more investigation could be done to optimize the way of classifying the bids. Also applying other methods of dimension reduction techniques for the bid data might grant great improvements. Another important issue concerns other relevant data used for modeling the price classes. For example we ignored so far the impact of public holidays. On holidays like Christmas Eve, Christmas Day or New Year's Day the model performs relatively weak. Here improvement is relatively easy possible. Moreover, the inclusion of market price time series of different markets like the intra-day price as well as auction results of related markets, such as those from neighboring countries, could be beneficiary for the model quality. Other useful regressors could be different fuel costs or CO_2 allowances. Also the restructuring procedure that was used for mapping the local price behavior provides a lot of space for further improvement. The probabilities that a certain price is traded or not could be modeled time-varying.

References

[1] Nicholas Bowden and James E Payne. Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models. *Energy Economics*, 30(6):3186–3197, 2008.

[2] Clive G Bowsher. Modelling the dynamics of cross-sectional price functions: an econometric analysis of the bid and ask curves of an automated exchange. Technical report, Nuffield Economics Working Paper, 2004.

[3] René Carmona, Michael Coulon, and Daniel Schwarz. Electricity price modeling and asset valuation: a multi-fuel structural approach. *Mathematics and Financial Economics*, 7(2):167–202, 2013.

[4] TM Christensen, AS Hurn, and KA Lindsay. Forecasting spikes in electricity prices. *International Journal of Forecasting*, 28(2):400–411, 2012.

[5] Johanna Cludius, Hauke Hermann, Felix Chr Matthes, and Verena Graichen. The merit order effect of wind and photovoltaic electricity generation in Germany 2008–2016: Estimation and distributional implications. *Energy Economics*, 44:302–313, 2014.

[6] Antonio J Conejo, Miguel A Plazas, Rosario Espinola, and Ana B Molina. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Transactions on Power Systems*, 20(2):1035–1042, 2005.

[7] Michael Eichler and Dennis Tuerk. Fitting semiparametric Markov regime-switching models to electricity spot prices. *Energy Economics*, 36:614–624, 2013.

[8] Alvaro Escribano, J Ignacio Peña, and Pablo Villaplana. Modelling Electricity Prices: International Evidence. *Oxford Bulletin of Economics and Statistics*, 73(5):622–650, 2011.

[9] Jerome Friedman, Trevor Hastie, Holger Höfling, and Robert Tibshirani. Pathwise coordinate optimization. *The Annals of Applied Statistics*, 1(2):302–332, 2007.

[10] Sam Howison and Michael C Coulon. Stochastic behaviour of the electricity bid stack: from fundamental drivers to power prices. *The Journal of Energy Markets*, 2(1):1–27, 2009.

[11] Joanna Janczura and Rafał Weron. Efficient estimation of Markov regime-switching models: An application to electricity spot prices. *AStA Advances in Statistical Analysis*, 96(3):385–407, 2012.

[12] Nektaria V Karakatsani and Derek W Bunn. Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International Journal of Forecasting*, 24(4):764–785, 2008.

[13] Dogan Keles, Massimo Genoese, Dominik Möst, and Wolf Fichtner. Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. *Energy Economics*, 34(4):1012–1032, 2012.

[14] Janina C Ketterer. The impact of wind power generation on the electricity price in Germany. *Energy Economics*, 44:270–280, 2014.

[15] Heping Liu and Jing Shi. Applying ARMA–GARCH approaches to forecasting short-term electricity prices. *Energy Economics*, 37:152–166, 2013.

[16] Zhen Liu, Jianming Yan, Yuren Shi, Kaiwei Zhu, and Gangqing Pu. Multi-agent based experimental analysis on bidding mechanism in electricity auction markets. *International Journal of Electrical Power & Energy Systems*, 43(1):696–702, 2012.

[17] Katarzyna Maciejowska, Jakub Nowotarski, and Rafał Weron. Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging. *International Journal of Forecasting (to appear)*.

[18] Dylan McConnell, Patrick Hearps, Dominic Eales, Mike Sandiford, Rebecca Dunn, Matthew Wright, and Lachlan Bateman. Retrospective modeling of the merit-order effect on wholesale electricity prices from distributed photovoltaic generation in the Australian National Electricity Market. *Energy Policy*, 58:17–27, 2013.

[19] Frank Sensfuß, Mario Ragwitz, and Massimo Genoese. 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):3086–3094, 2008.

[20] Derk J Swider and Christoph Weber. Extended ARMA models for estimating price developments on day-ahead electricity markets. *Electric Power Systems Research*, 77(5):583–593, 2007.

[21] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 58(1):267–288, 1996.

[22] Mariano Ventosa, Alvaro Baıllo, Andrés Ramos, and Michel Rivier. Electricity market modeling trends. *Energy Policy*, 33(7):897–913, 2005.

[23] Rafal Weron. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. John Wiley & Sons, Chichester, England, 2006.

[24] Rafał Weron. Market price of risk implied by Asian-style electricity options and futures. *Energy Economics*, 30(3):1098–1115, 2008.

[25] Rafał Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081, 2014.

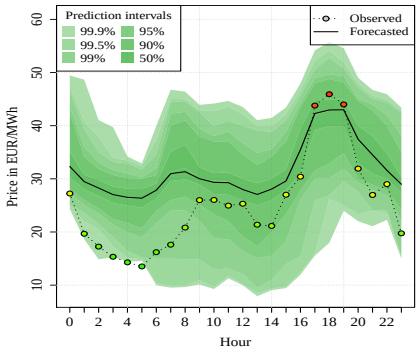
[26] Rafał Weron and Adam Misiorek. Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4):744–763, 2008.

[27] Florian Ziel, Rick Steinert, and Sven Husmann. Efficient modeling and forecasting of electricity spot prices. *Energy Economics*, 47:98–111, 2015.

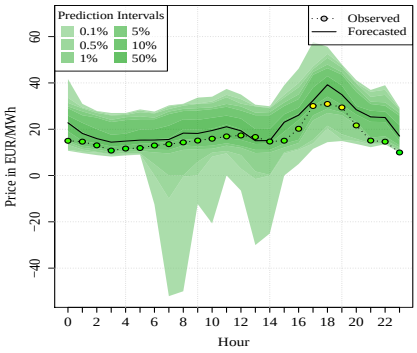
[28] Florian Ziel, Rick Steinert, and Sven Husmann. Forecasting day ahead electricity spot prices: The impact of the EXAA to other European electricity markets. *Energy Economics (to appear)*, 2015.

Appendix

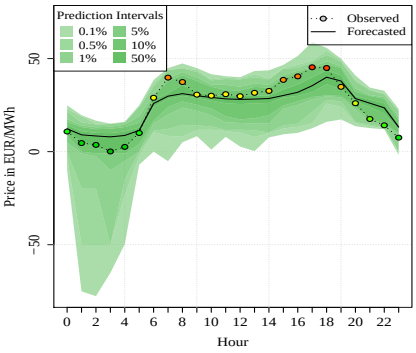
Price forecasts of the X-Model for the full out-of-sample price range:



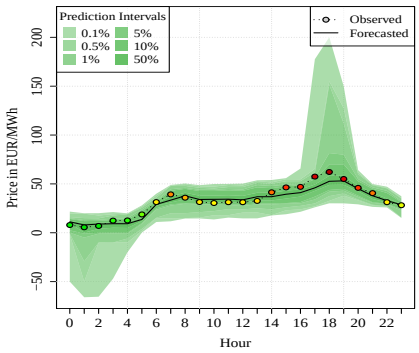
01.11.2014



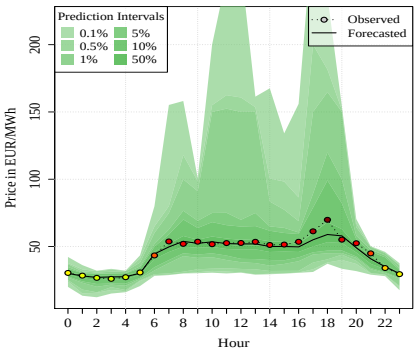
02.11.2014



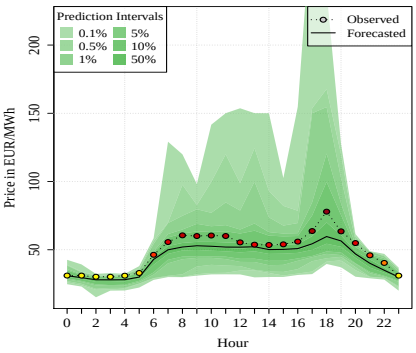
03.11.2014



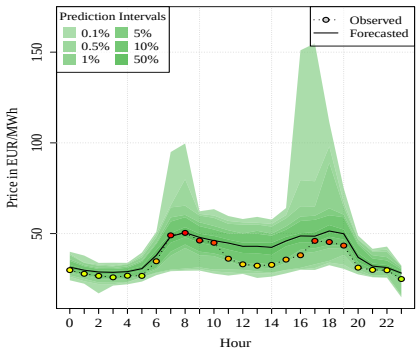
04.11.2014



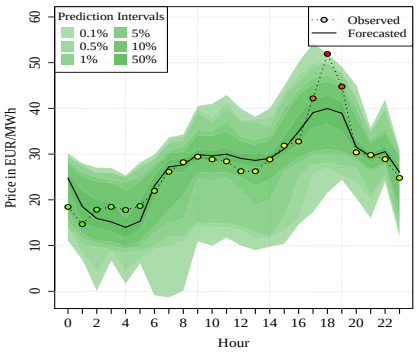
05.11.2014



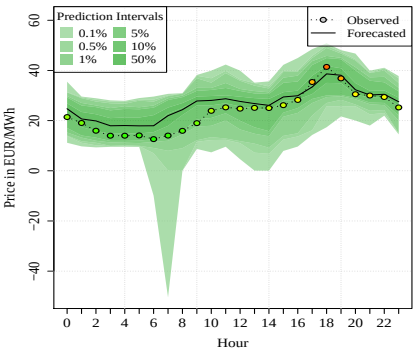
06.11.2014



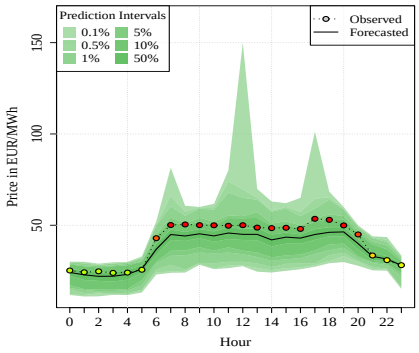
07.11.2014



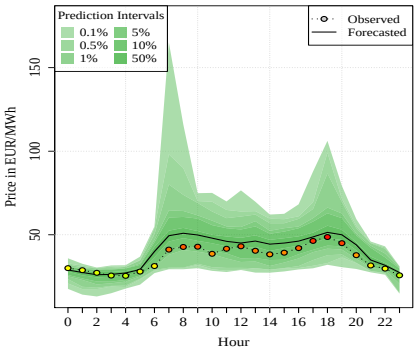
08.11.2014



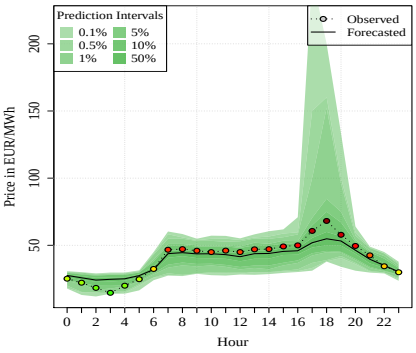
09.11.2014



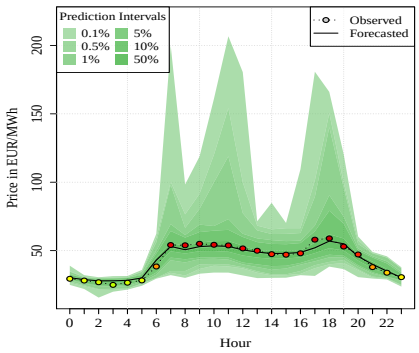
10.11.2014



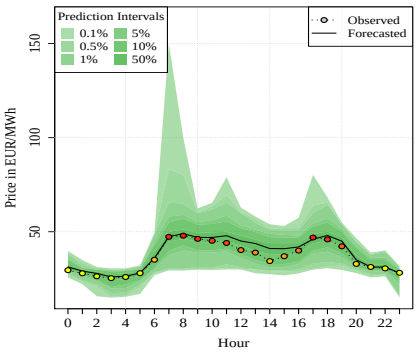
11.11.2014



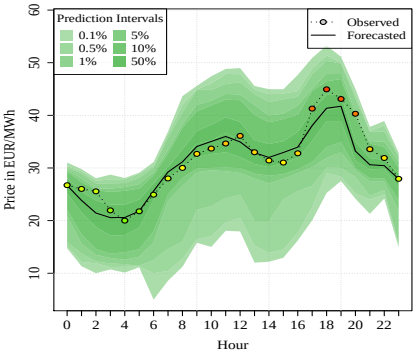
12.11.2014



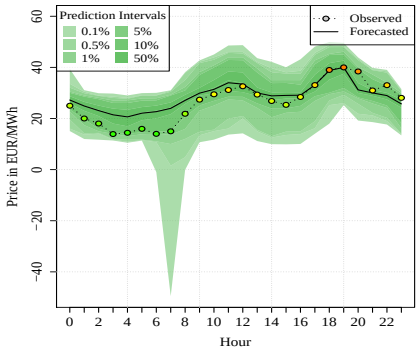
13.11.2014



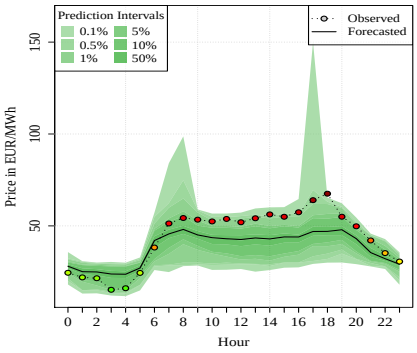
14.11.2014



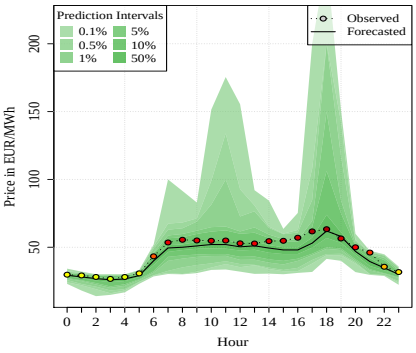
15.11.2014



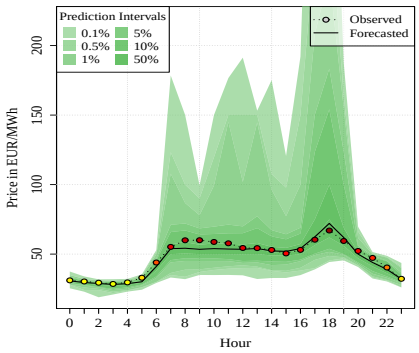
16.11.2014



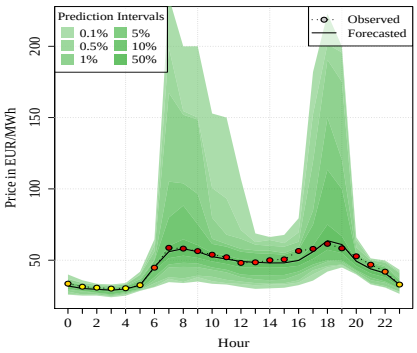
17.11.2014



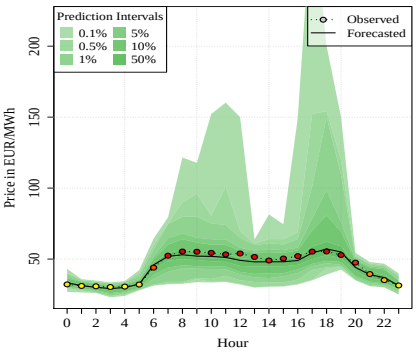
18.11.2014



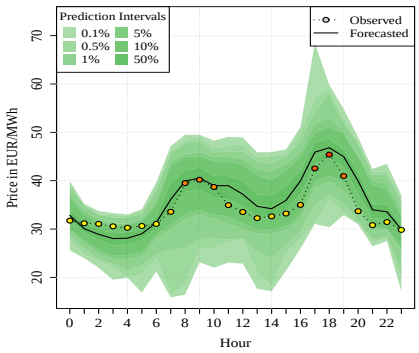
19.11.2014



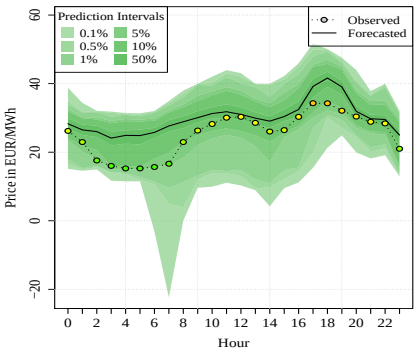
20.11.2014



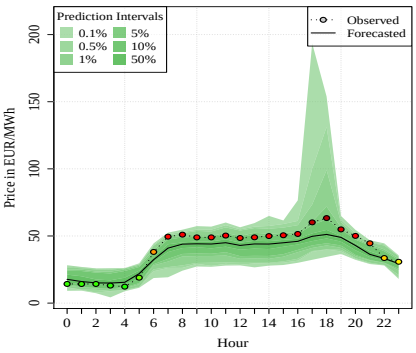
21.11.2014



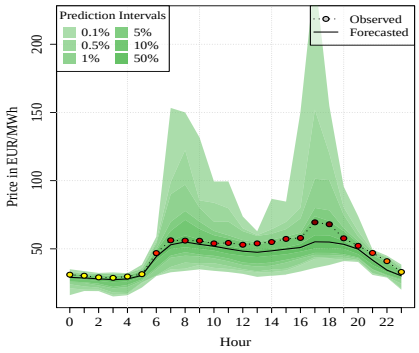
22.11.2014



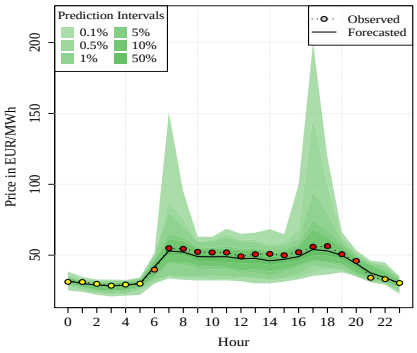
23.11.2014



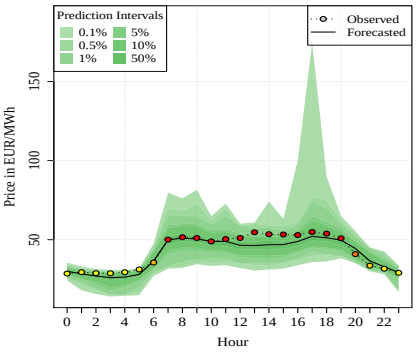
24.11.2014



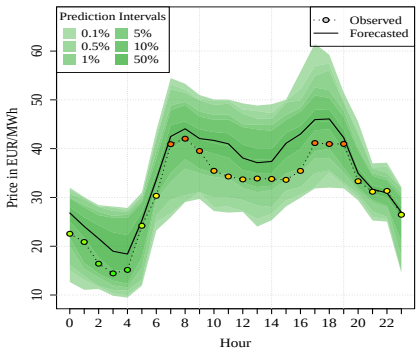
25.11.2014



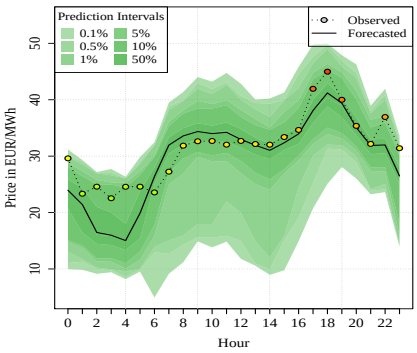
26.11.2014



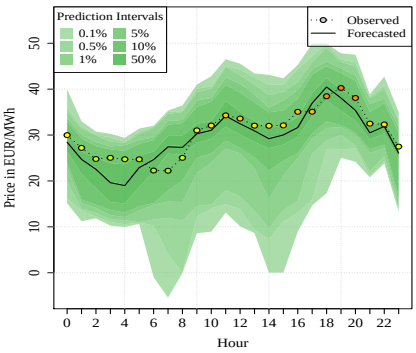
27.11.2014



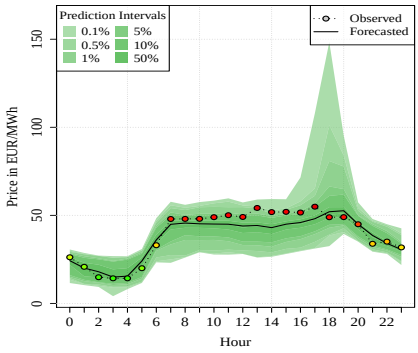
28.11.2014



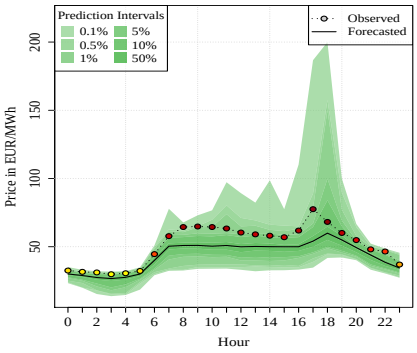
29.11.2014



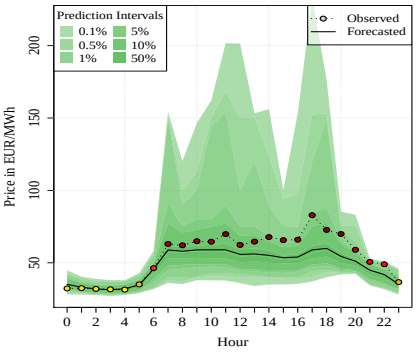
30.11.2014



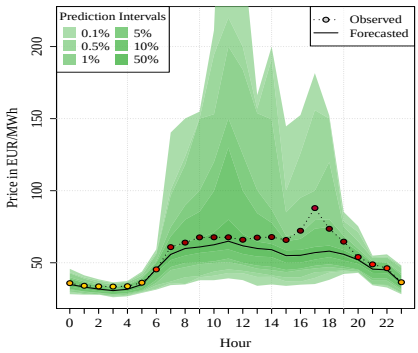
01.12.2014



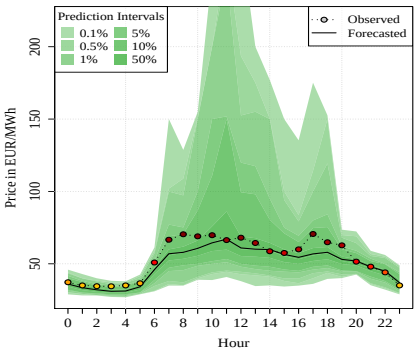
02.12.2014



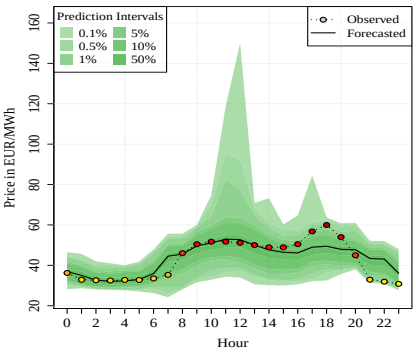
03.12.2014



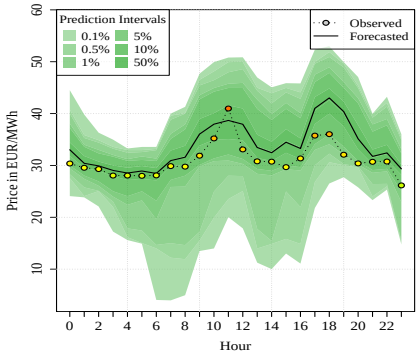
04.12.2014



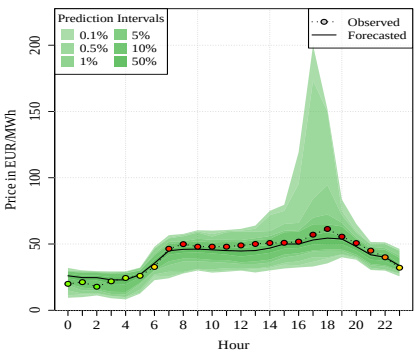
05.12.2014



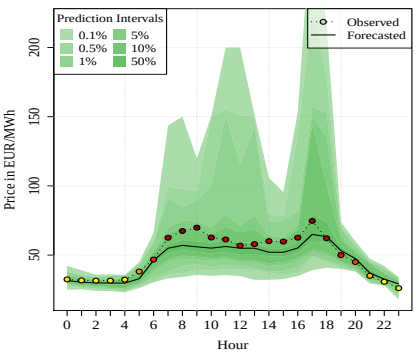
06.12.2014



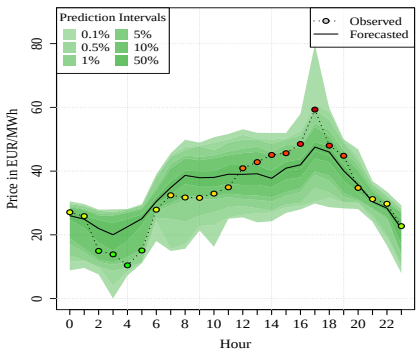
07.12.2014



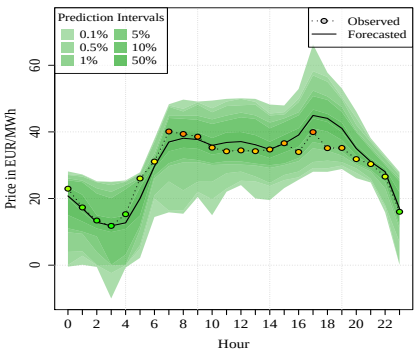
08.12.2014



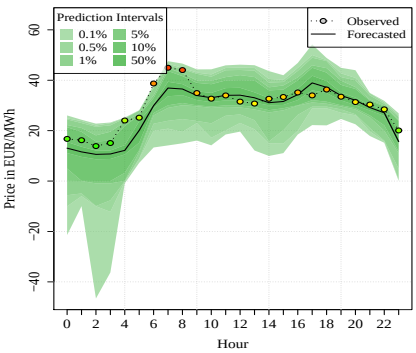
09.12.2014



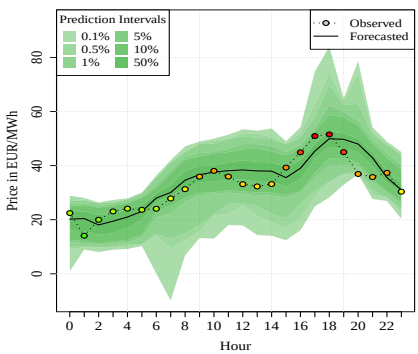
10.12.2014



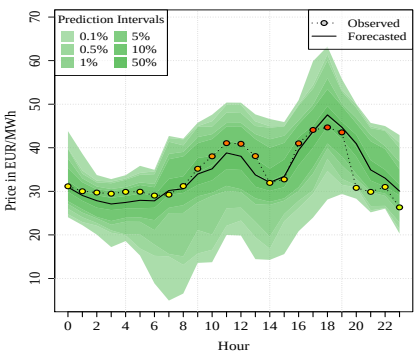
11.12.2014



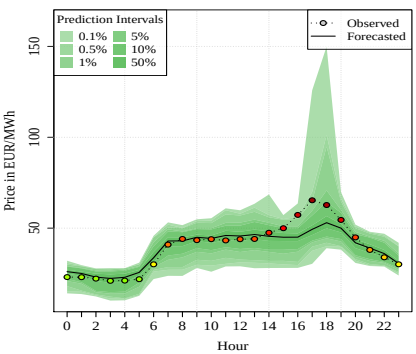
12.12.2014



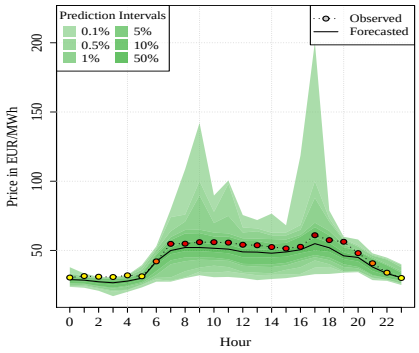
13.12.2014



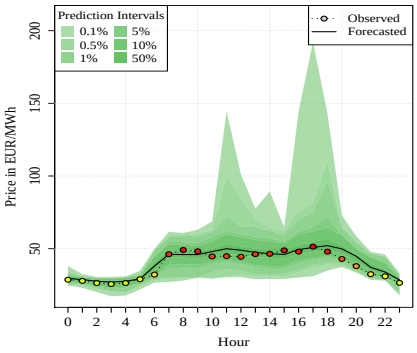
14.12.2014



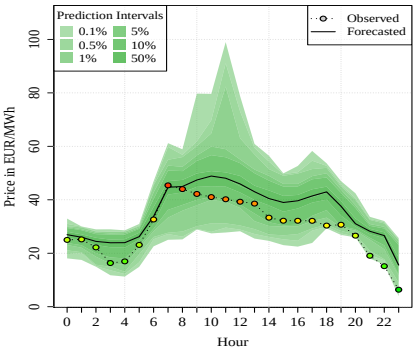
15.12.2014



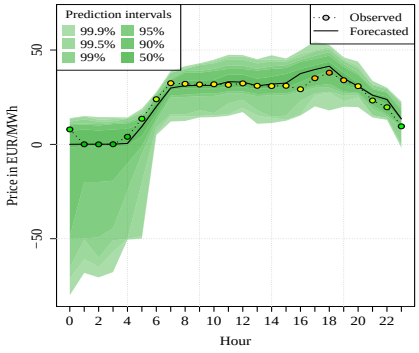
16.12.2014



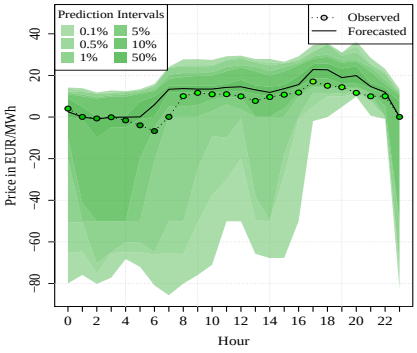
17.12.2014



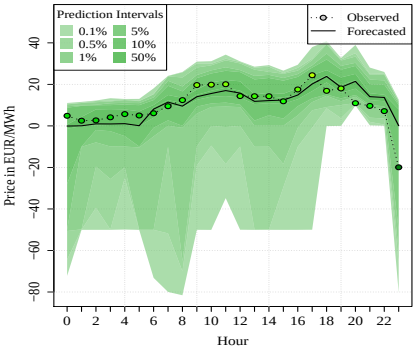
18.12.2014



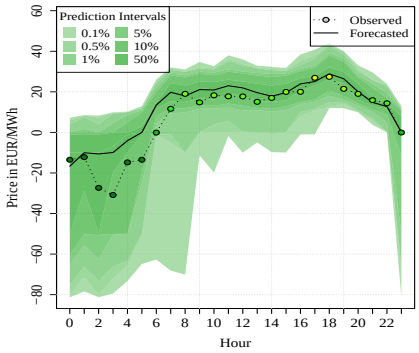
19.12.2014



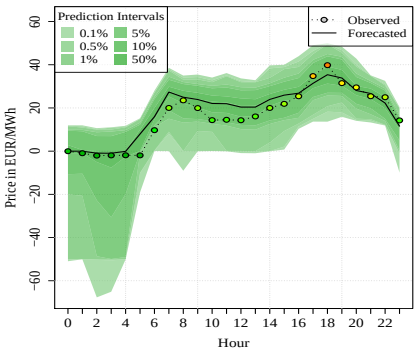
20.12.2014



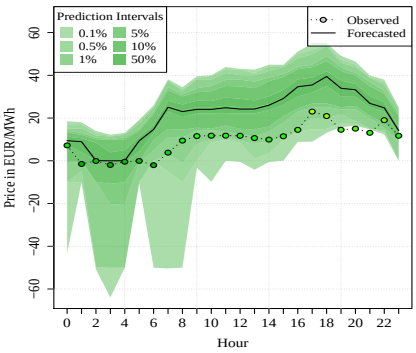
21.12.2014



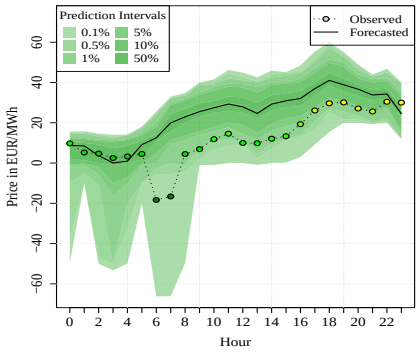
22.12.2014



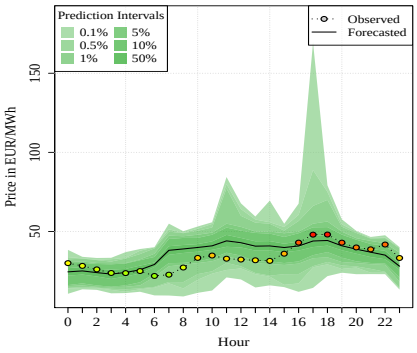
23.12.2014



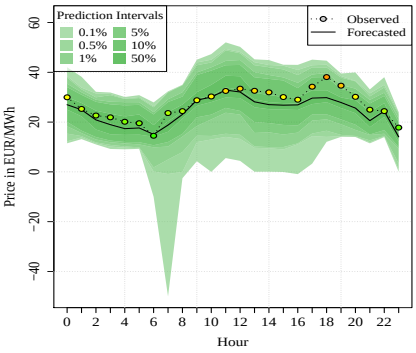
24.12.2014



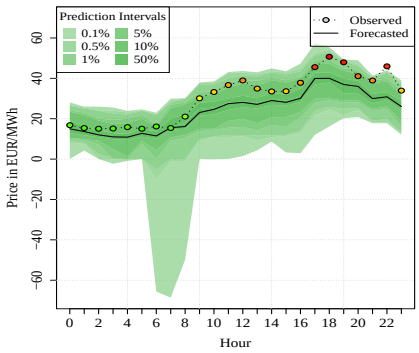
25.12.2014



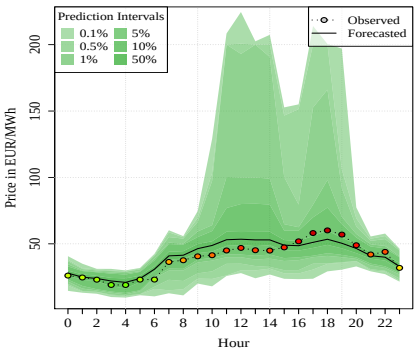
26.12.2014



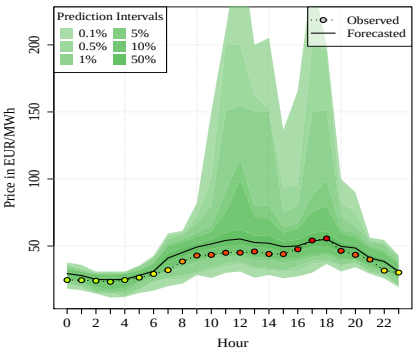
27.12.2014



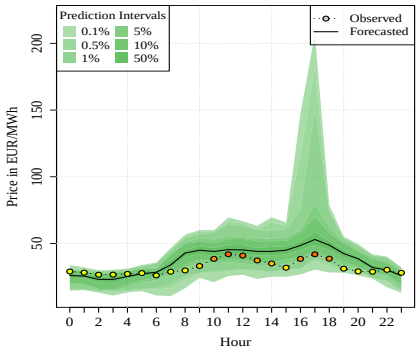
28.12.2014



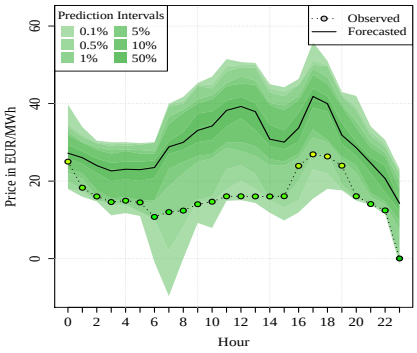
29.12.2014



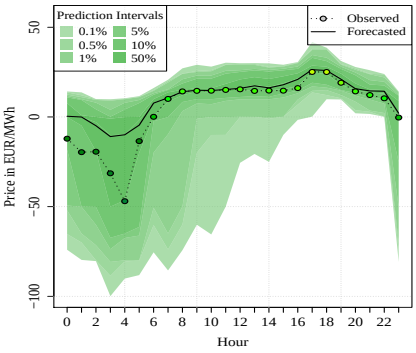
30.12.2014



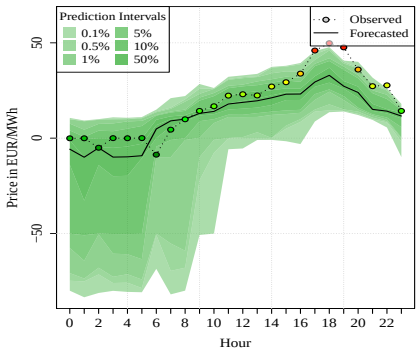
31.12.2014



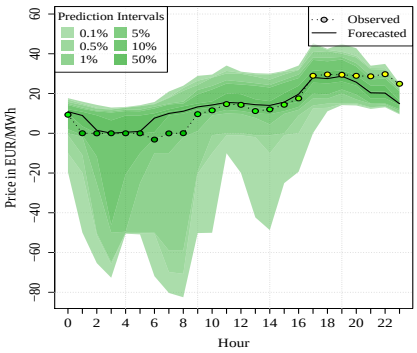
01.01.2015



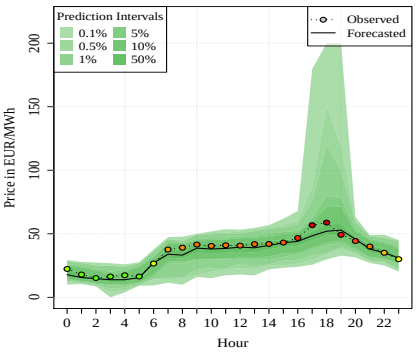
02.01.2015



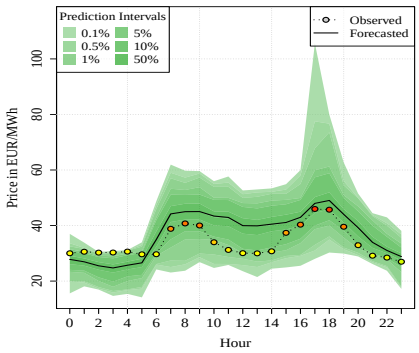
03.01.2015



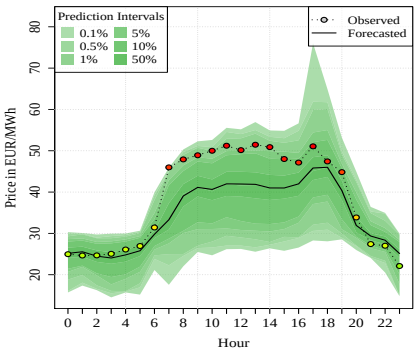
04.01.2015



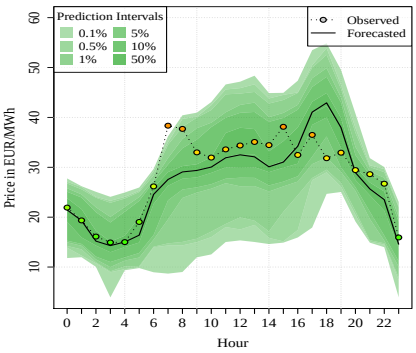
05.01.2015



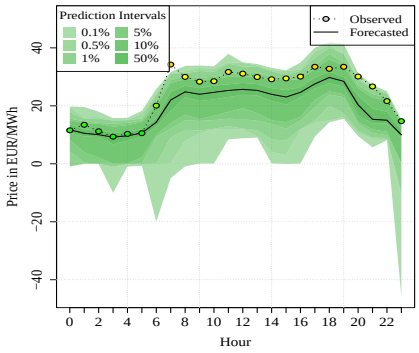
06.01.2015



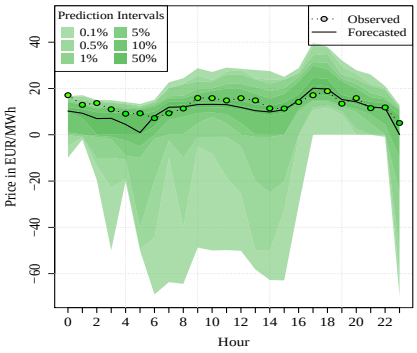
07.01.2015



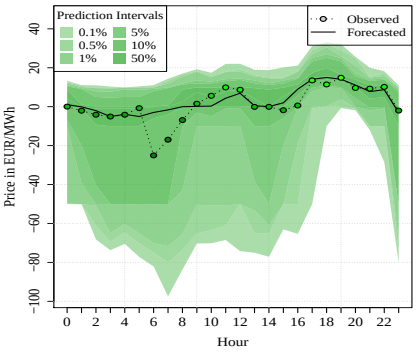
08.01.2015



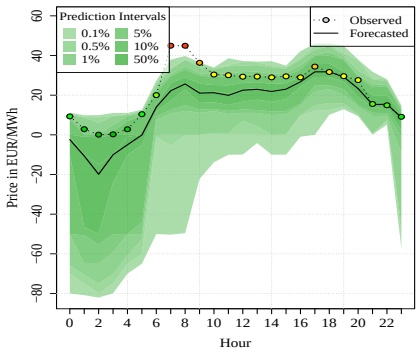
09.01.2015



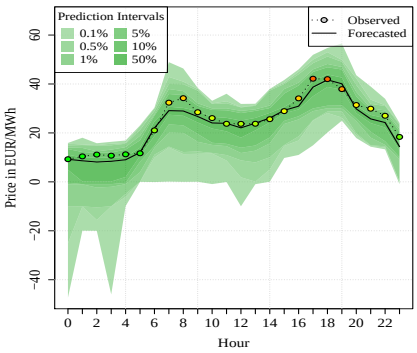
10.01.2015



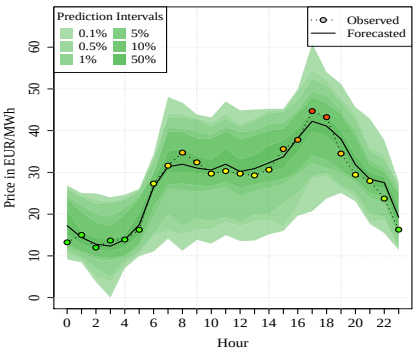
11.01.2015



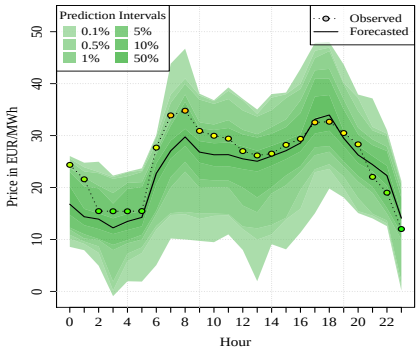
12.01.2015



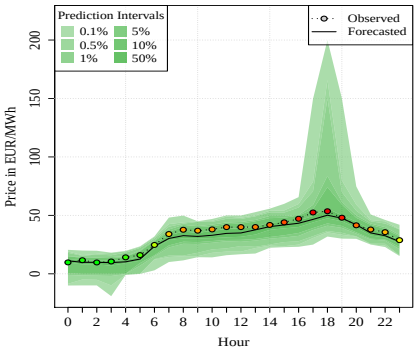
13.01.2015



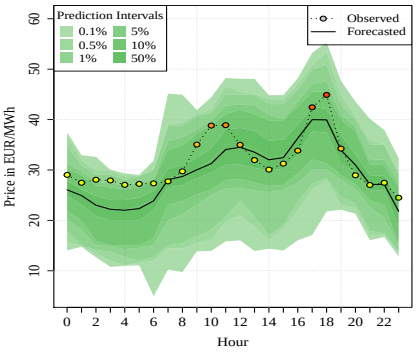
14.01.2015



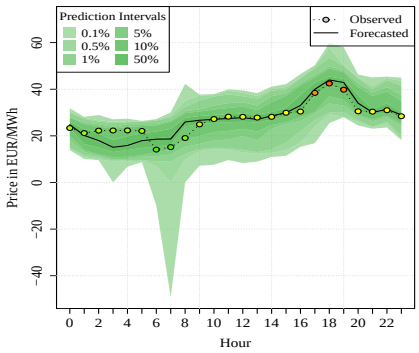
15.01.2015



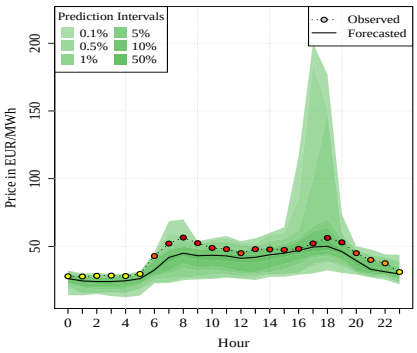
16.01.2015



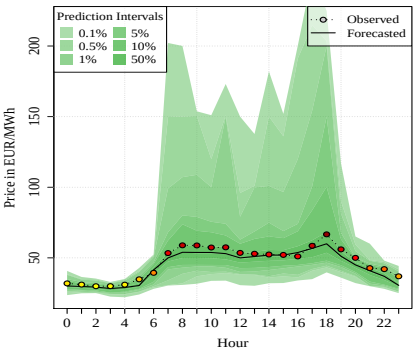
17.01.2015



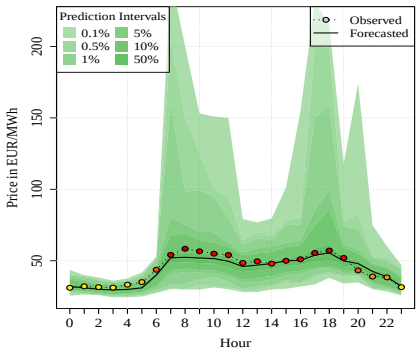
18.01.2015



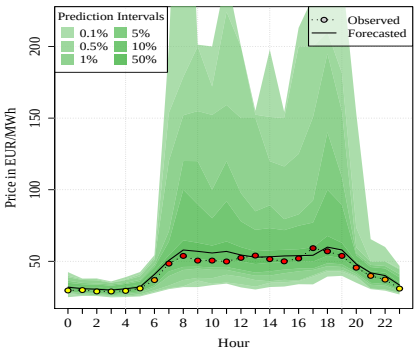
19.01.2015



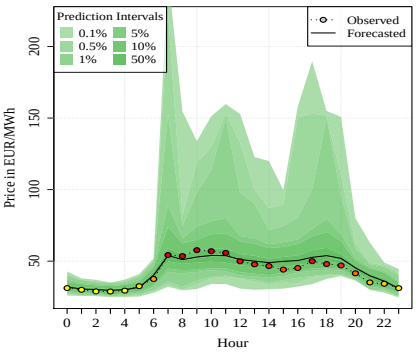
20.01.2015



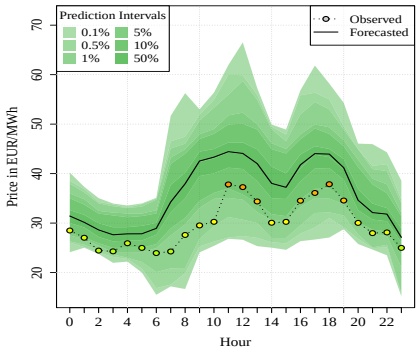
21.01.2015



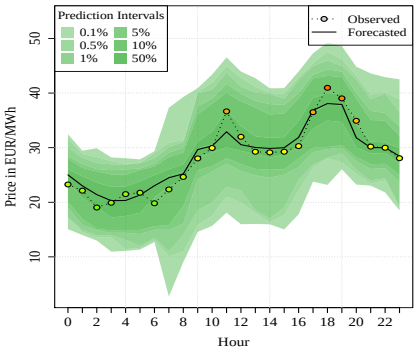
22.01.2015



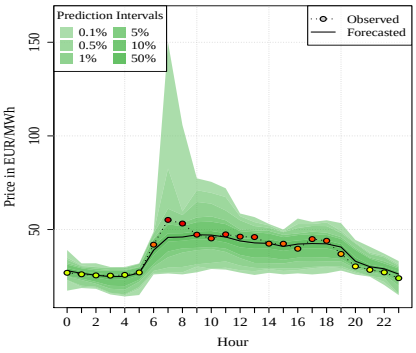
23.01.2015



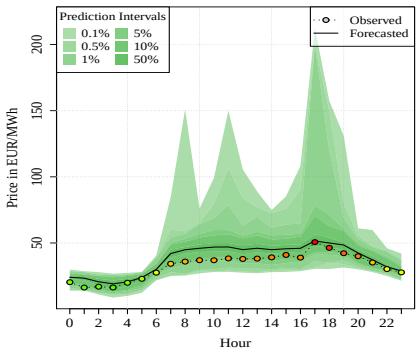
24.01.2015



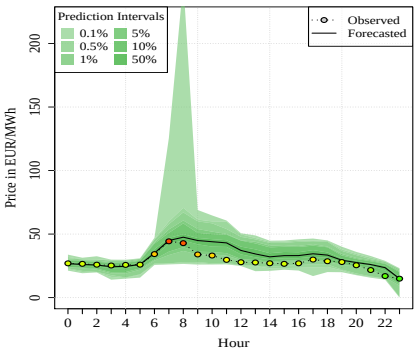
25.01.2015



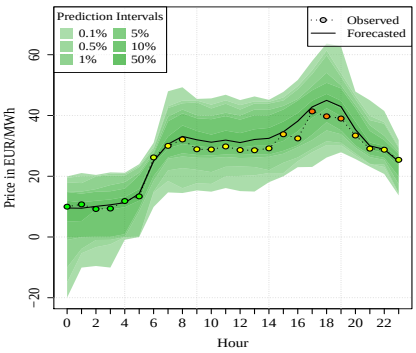
26.01.2015



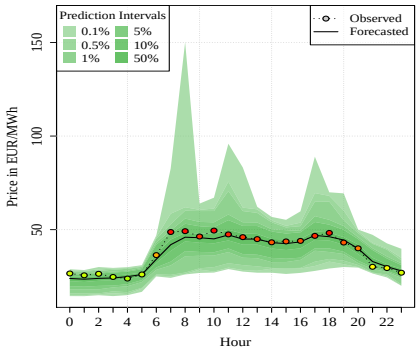
27.01.2015



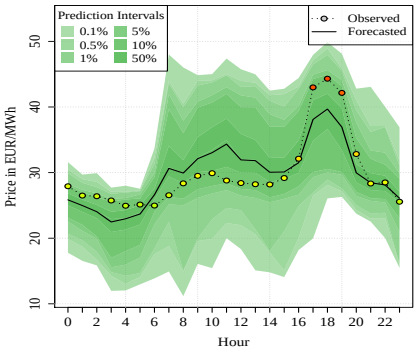
28.01.2015



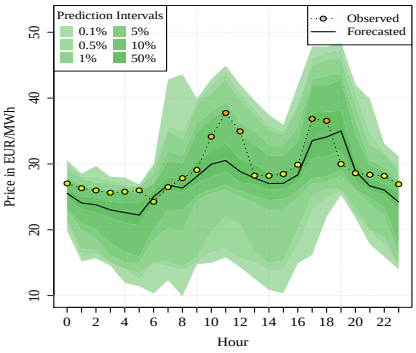
29.01.2015



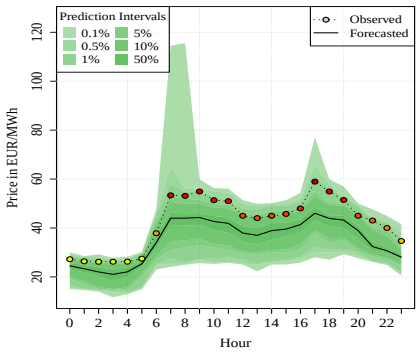
30.01.2015



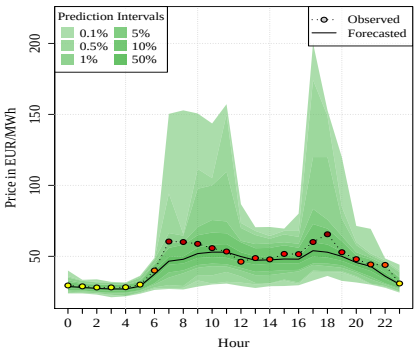
31.01.2015



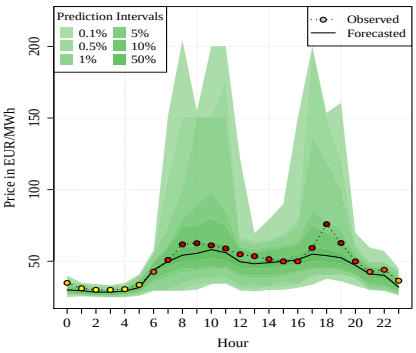
01.02.2015



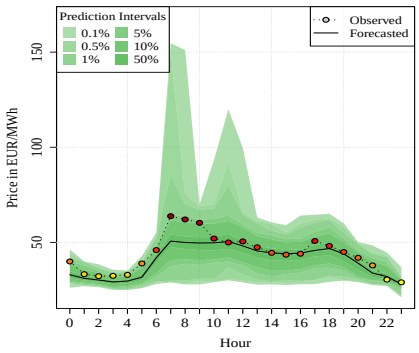
02.02.2015



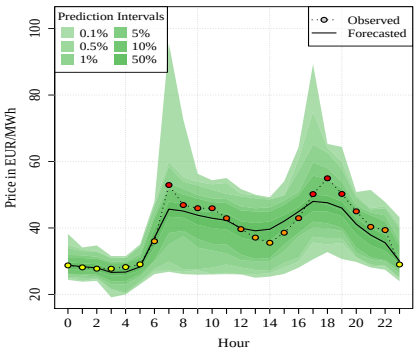
03.02.2015



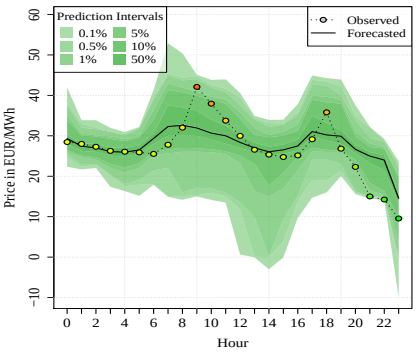
04.02.2015



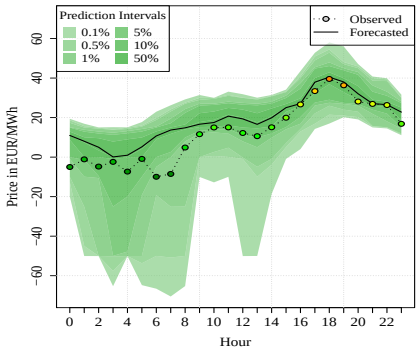
05.02.2015



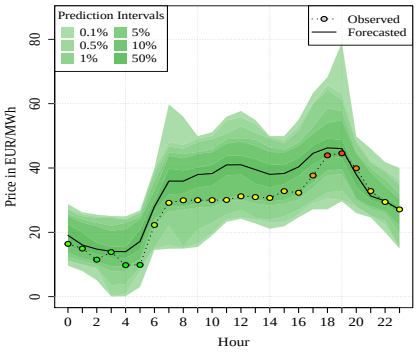
06.02.2015



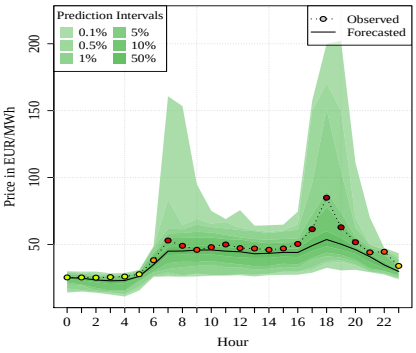
07.02.2015



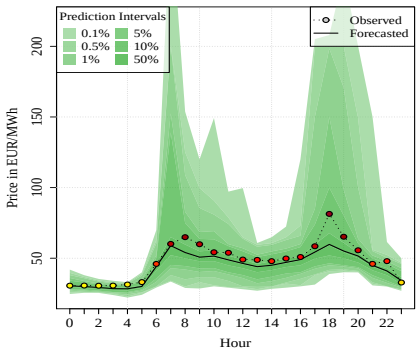
08.02.2015



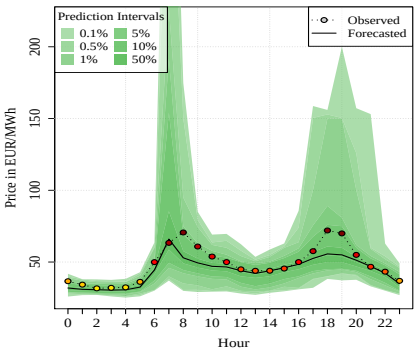
09.02.2015



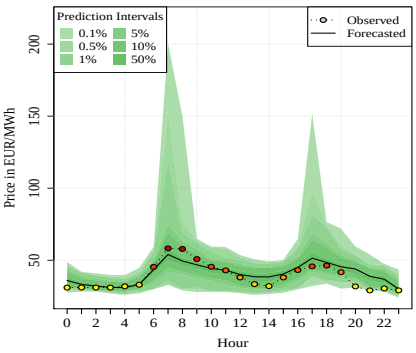
10.02.2015



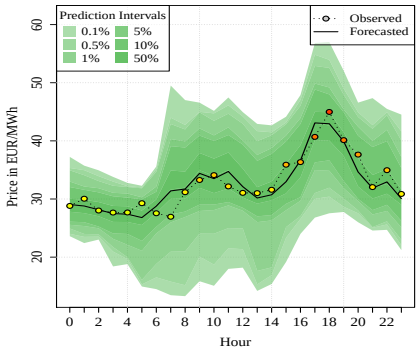
11.02.2015



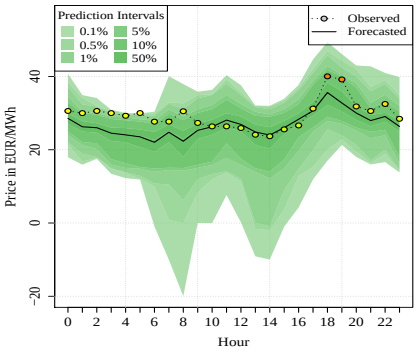
12.02.2015



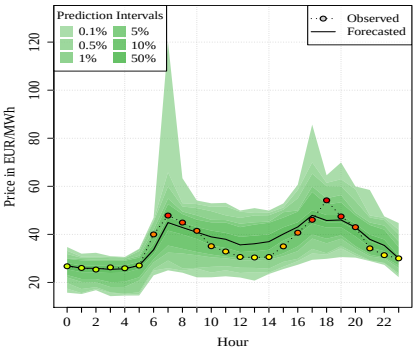
13.02.2015



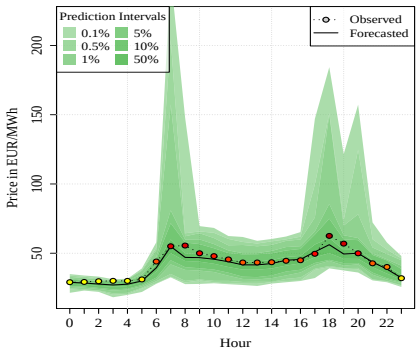
14.02.2015



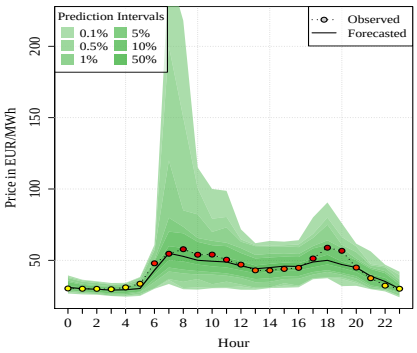
15.02.2015



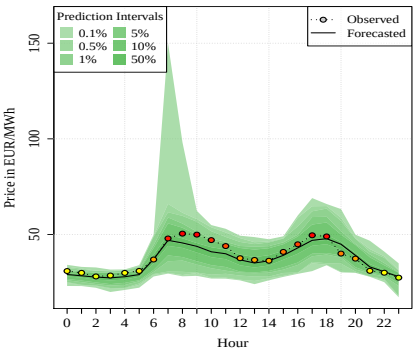
16.02.2015



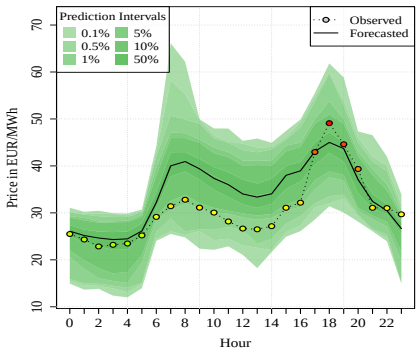
17.02.2015



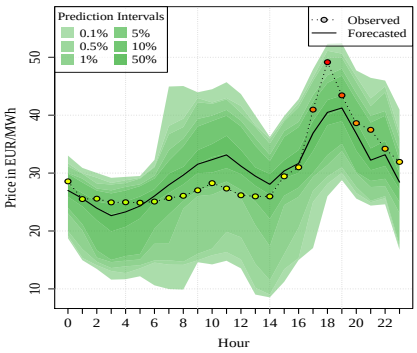
18.02.2015



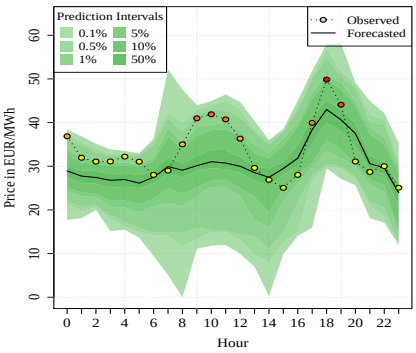
19.02.2015



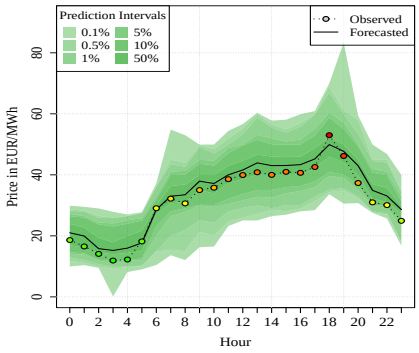
20.02.2015



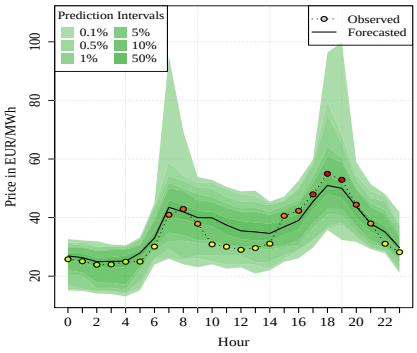
21.02.2015



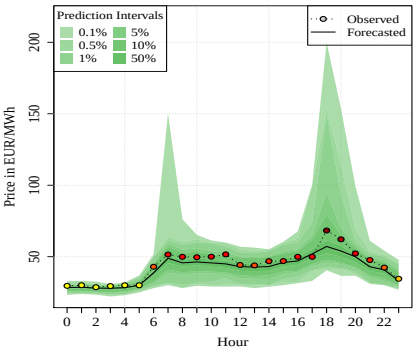
22.02.2015



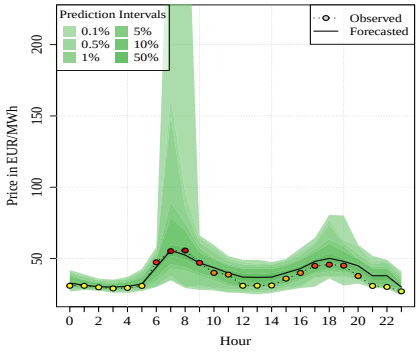
23.02.2015



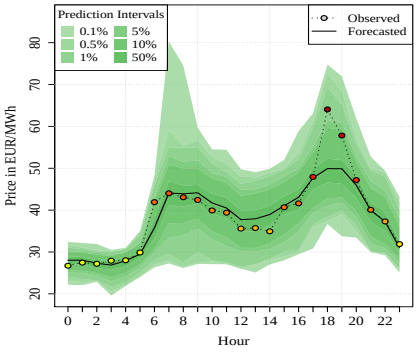
24.02.2015



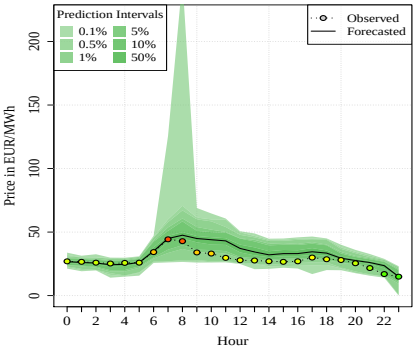
25.02.2015



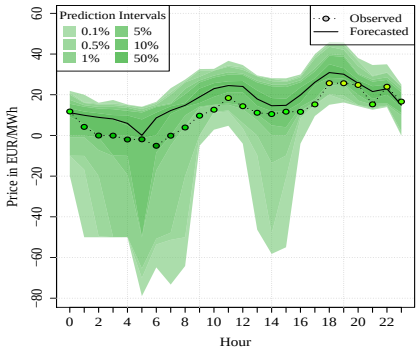
26.02.2015



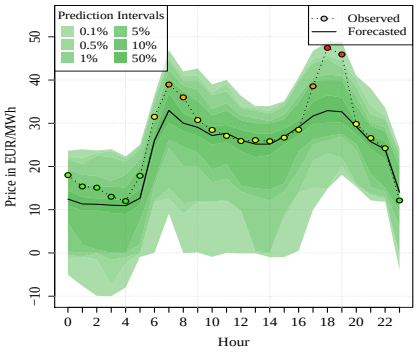
27.02.2015



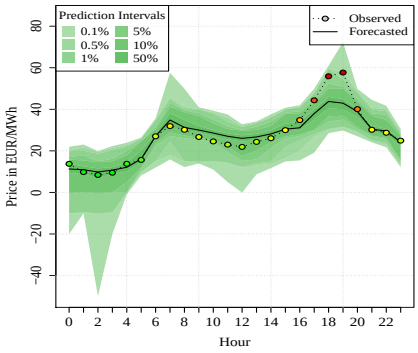
28.02.2015



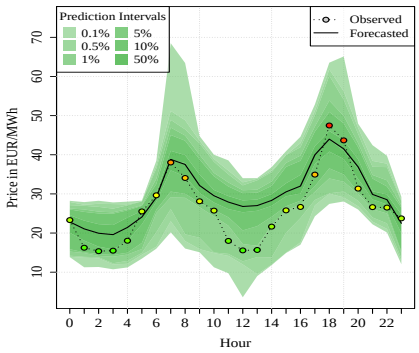
01.03.2015



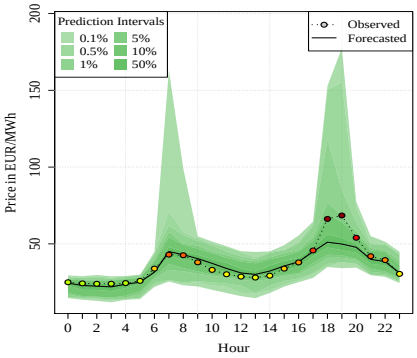
02.03.2015



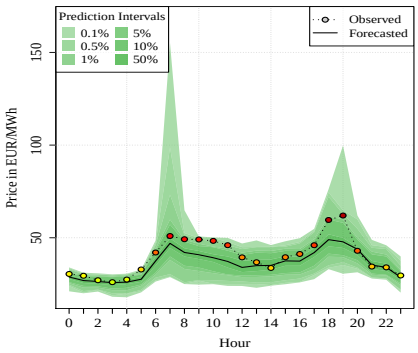
03.03.2015



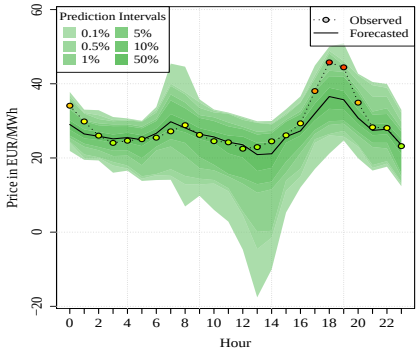
04.03.2015



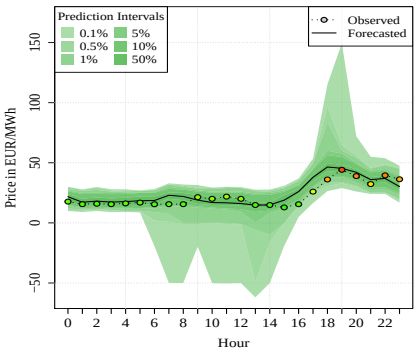
05.03.2015



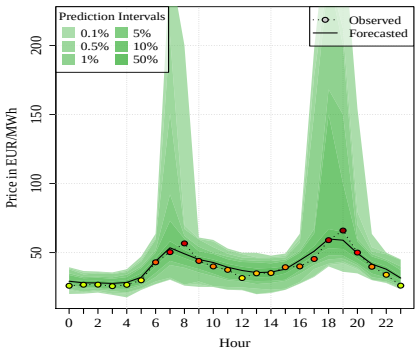
06.03.2015



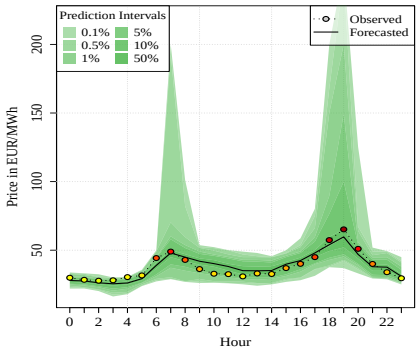
07.03.2015



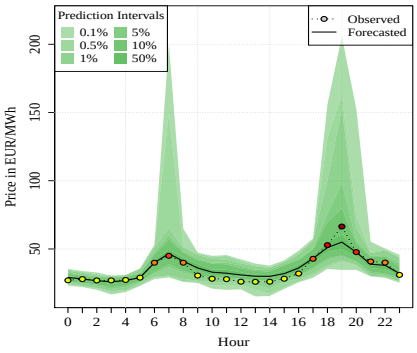
08.03.2015



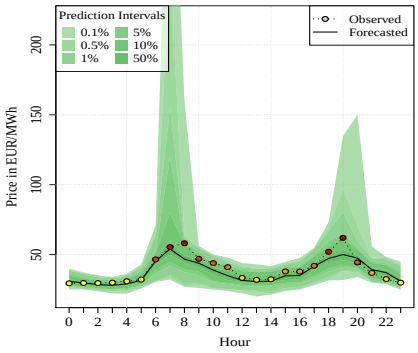
09.03.2015



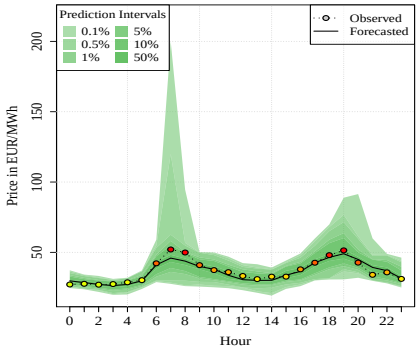
10.03.2015



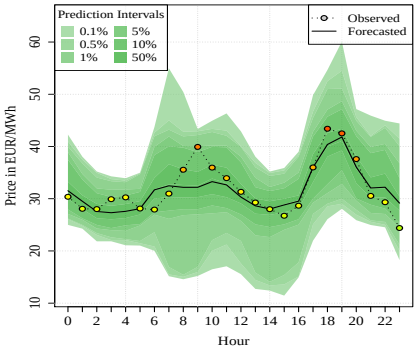
11.03.2015



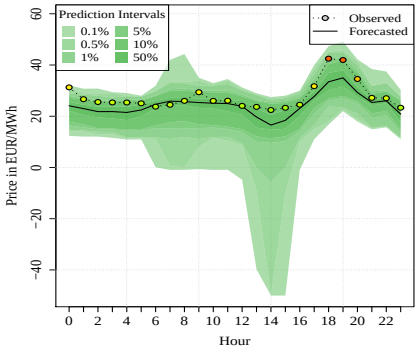
12.03.2015



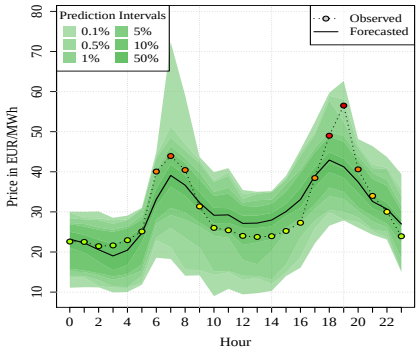
13.03.2015



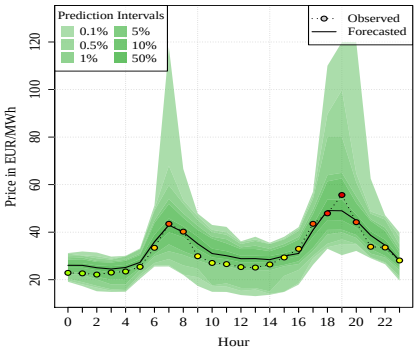
14.03.2015



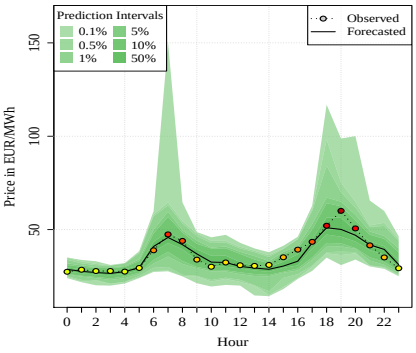
15.03.2015



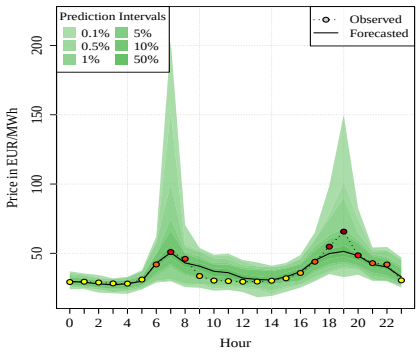
16.03.2015



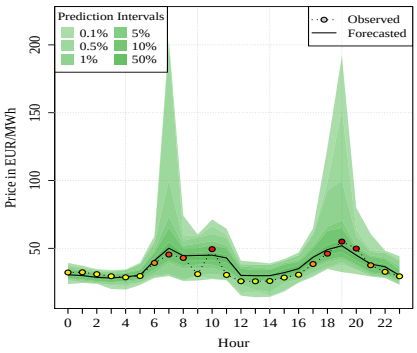
17.03.2015



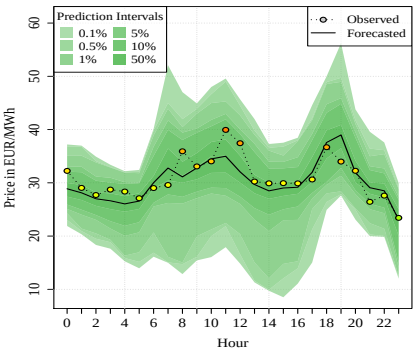
18.03.2015



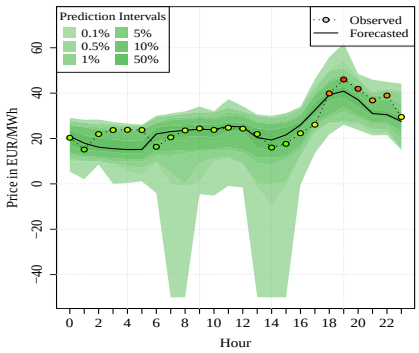
19.03.2015



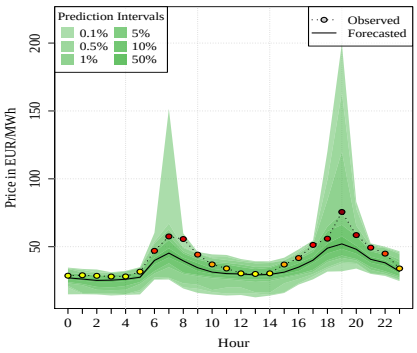
20.03.2015



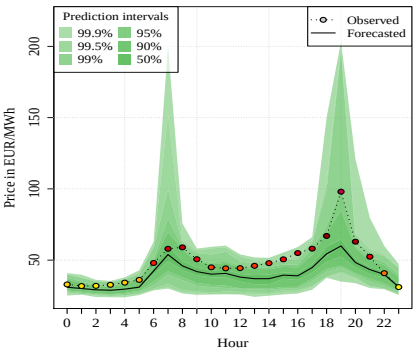
21.03.2015



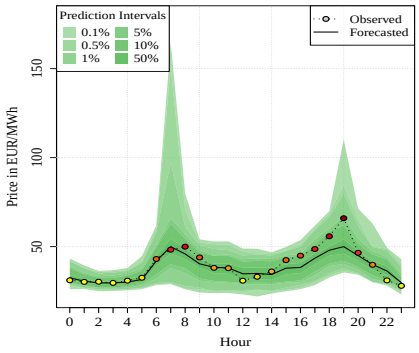
22.03.2015



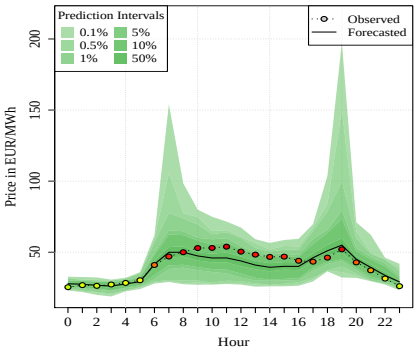
23.03.2015



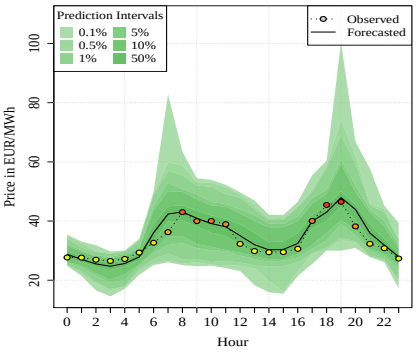
24.03.2015



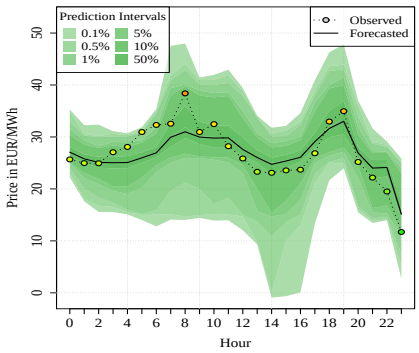
25.03.2015



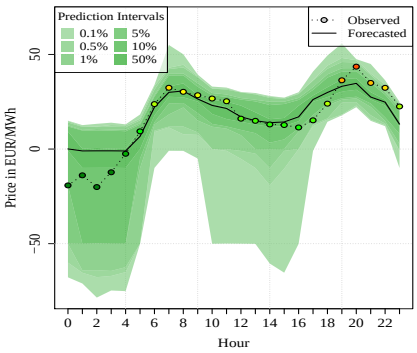
26.03.2015



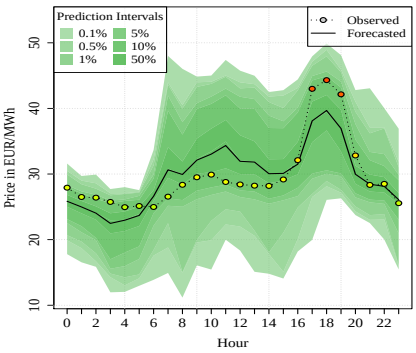
27.03.2015



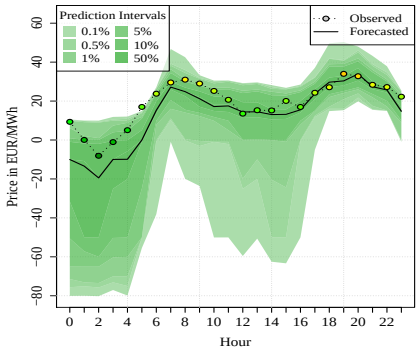
28.03.2015



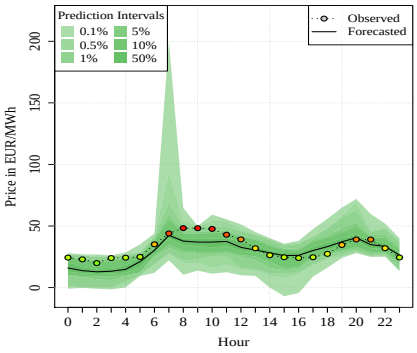
30.03.2015



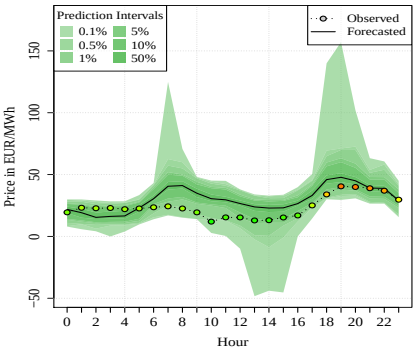
31.03.2015



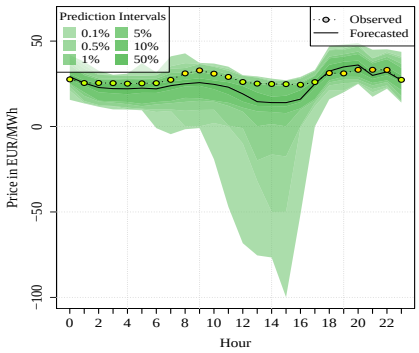
01.04.2015



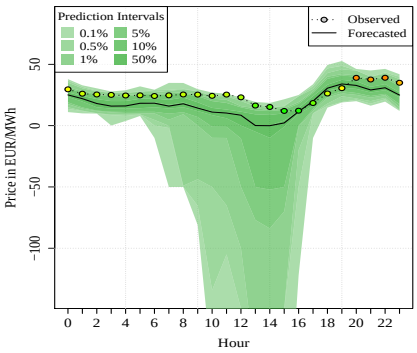
02.04.2015



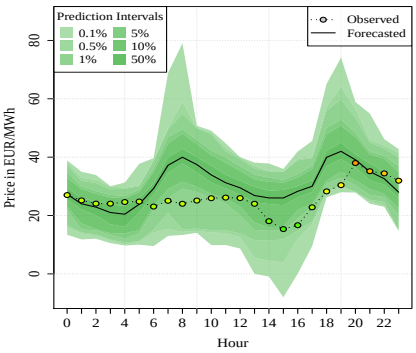
03.04.2015



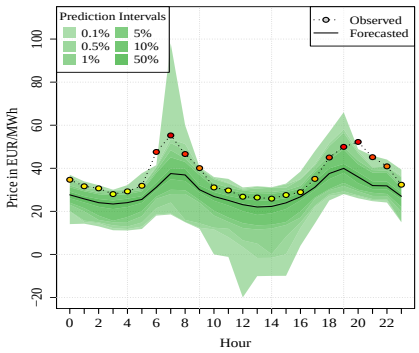
04.04.2015



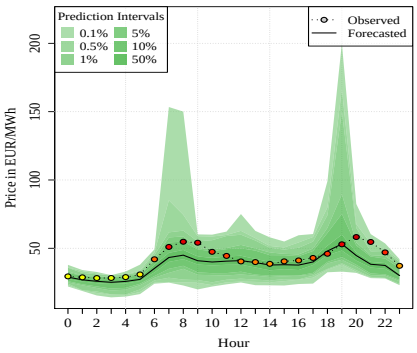
05.04.2015



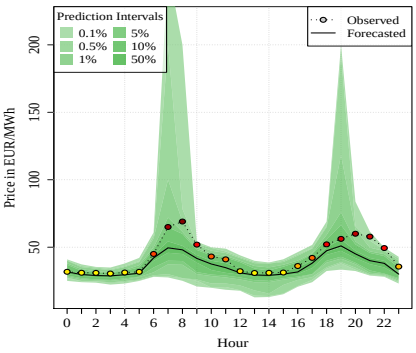
06.04.2015



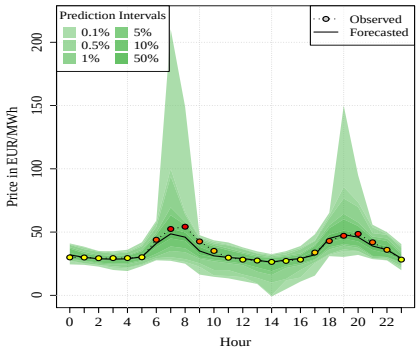
07.04.2015



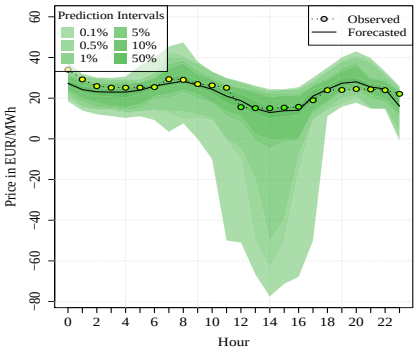
08.04.2015



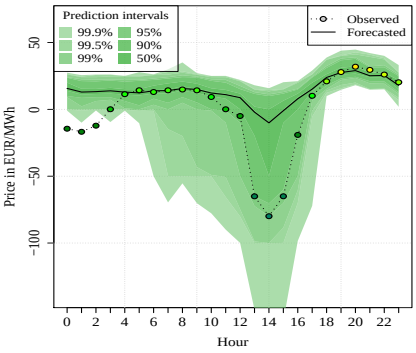
09.04.2015



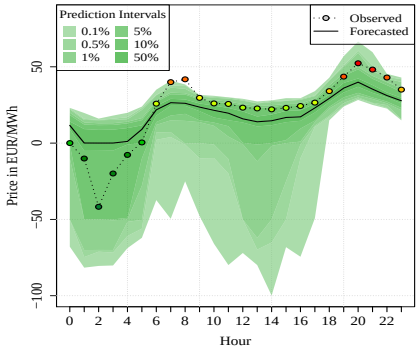
10.04.2015



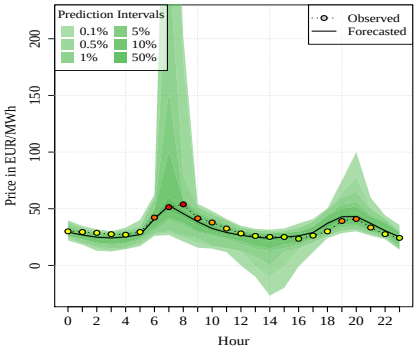
11.04.2015



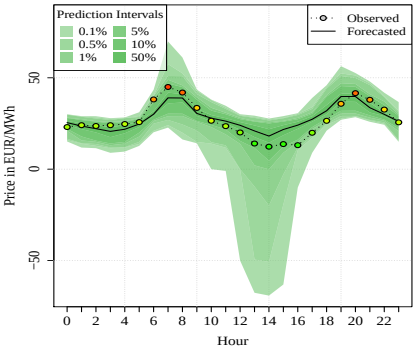
12.04.2015



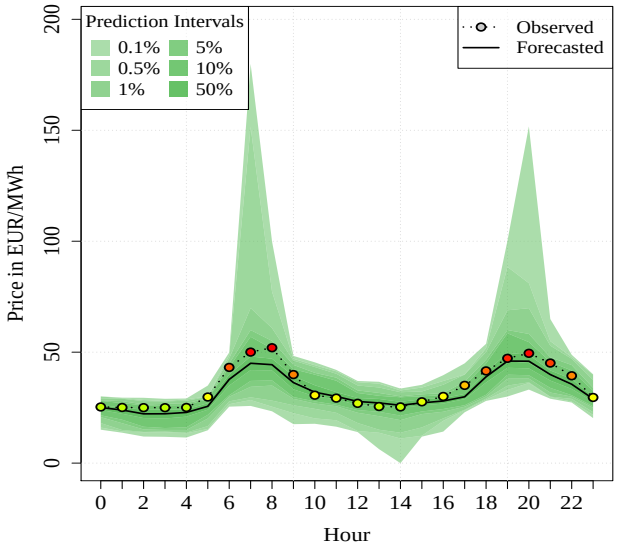
13.04.2015



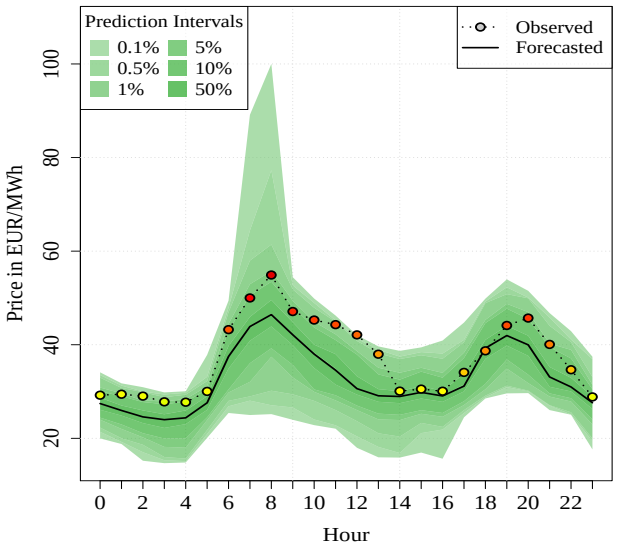
14.04.2015



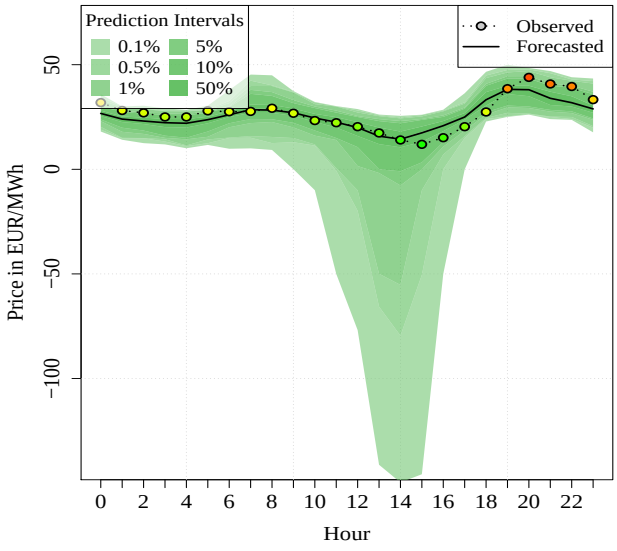
15.04.2015



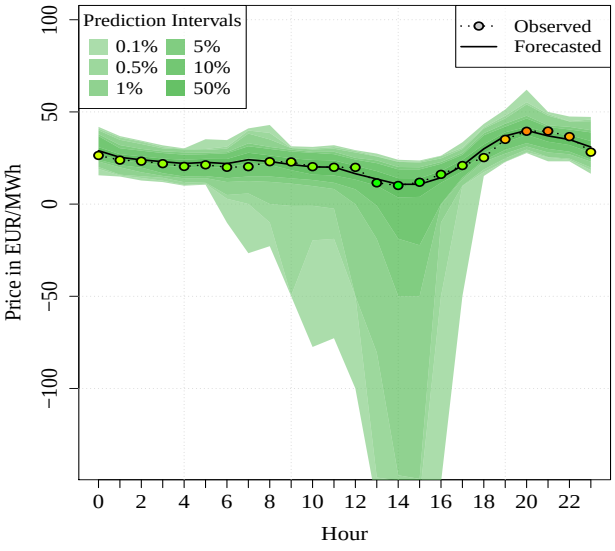
16.04.2015



17.04.2015



18.04.2015



19.04.2015