Robust Calculation and Parameter Estimation of the Hourly Price Forward Curve

Marcus Hildmann
ETH Zürich
Zürich, Switzerland
hildmann@eeh.ee.ethz.ch

Florian Herzog
swissQuant Group
Zürich, Switzerland
herzog@swissquant.ch

Dejan Stokic
swissQuant Group
Zürich, Switzerland
stokic@swissquant.ch

Jeroen Cornel
swissQuant Group
Zürich, Switzerland
cornel@swissquant.ch

Göran Andersson
ETH Zürich
Zürich, Switzerland
andersson@eeh.ee.ethz.ch

Abstract - Deregulated energy market participants mostly must use the Hourly Price Forward Curves (HPFC) to evaluate their long term energy price contracts. We propose a new framework to estimate of hourly, daily and yearly energy price profiles, based on the median estimation, instead of the widely used mean value. Hourly and daily data used for estimation are normalized in order to minimize the seasonality bias in our predictions. Given the high dimensionality of the problem, we use the LAD-Lasso model selection, as a way to prevent data overfitting. To test the proposed framework we used the German electricity spot price time-series. We discuss the effects of large negative hourly prices, residuals of daily estimation and data outliers. We show that our framework provides significantly improved estimation of HPFC compared to the simple mean estimators.

Keywords - Electricity prices modeling, hourly price forward curve, robust statistics, non-normal distributions, spot markets.

1 Introduction

In the early 1990's many power markets in Europe were transformed from monopolies of publicly owned companies into liberalized power markets. Since 1998, the power market in Germany is fully liberalized, with markets for spot and long term contracts available on energy exchanges and for over-the-counter trading. All companies, from the large power generating conglomerates to the public utility companies, sell and buy energy over these markets. However, the majority of the contracts are long term, hence they are used as power supply contracts and as hedging instruments for spot contracts. Every market participant, large producers, traders and public utility companies, must determines the price of a contract with a given profile on an hourly basis. Because electrical energy is a non-storable good, the standard forward pricing method, pricing the future F_0 by discounting the current spot price S_0 and storage costs U with interest rate r and time to maturity T as $F_0 = (S_0 + U)e^{(rT)}$ is not possible [1]. Another method is proposed by [2], with contract pricing based on the ability to produce power. Since the standard contract pricing methodology does not work and the method proposed by [2] cannot replicate seasonality and requires the full knowledge of all power plants which is not available, the method used to evaluate a long term contract is the hourly price forward curve (HPFC), which is used by all market participants as the arbitrage free instrument for pricing contracts on an hourly basis [3]. The HPFC is the starting point for every trading and hedging decision, contract risk evaluation and the basis of Monte Carlo methods for power plant evaluation.Therefore, it is the most important pricing instrument in power markets. The HPFC construction methods consist of the following elements:

1. estimation of the daily profile
2. estimation of the hourly profile
3. arbitrage free inclusion of the exchange traded forward prices

For the estimation of 1), factor models are used for seasonality, weather and other external influences, based on several regression techniques and with daily average of the spot prices as an input.
The spot market is in general a volatile market with spikes in price and with a skewed, non-normal distribution. In Germany, the high wind power generation capacity and, since September 2008, the possibility of negative prices result in even bigger challenges to achieve robustness in parameter estimation and model selection. The estimation of 2) is done through expectation estimation over different clusters of days. 3) will be done by multiplying the traded future products to the estimated profile.

In this paper, we develop a framework to estimate hourly profiles, daily base profiles, and a yearly profile via the median, see [4]. The estimation of the mean or the center of a distribution by using the average over the observed realizations is a non-robust estimation if the data is heavy-tailed or skewed. Due to the non-normal distribution of the power prices, the mean estimator needs to be robust and thus, we use the median as robust mean estimator. Price jumps highly affect the daily mean estimation, hence there are mostly one or two hours within a day with extreme, positive or negative values, which reflect only the outliers and not the whole day. This effects the calculation of all expectations and also the behavior of the regression algorithms. It will be shown that using the median as a basis for the expectation value estimation increases the robustness and also avoids the generation of negative daily averages.

For the calculations we ignore daylight saving time changes to ensure 24 hours per day over the full year. In section 4, we discuss the application of the daylight saving time to the HPFC.

The paper is organized as follows. We first discuss the HPFC model based on the mean estimator. In section 3 we introduce the HPFC model based on the median estimator and discuss the application of the LAD estimator and the LAD-Lasso model selection framework. The testing framework and a detailed discussion of the results is given in section 4. The last section 5 summarizes our results.

2 Hourly Price Forward Curve Calculation via the Mean

The HPFC is a forecast of future hourly prices including seasonality (intraday profile, weekly- and yearly), weather, environment information and the long term market expectations. To prevent arbitrage opportunities, the average price of the HPFC must be equal to the future contract at the corresponding time interval. This condition must be satisfied also for peak and offpeak hours. This is the so called arbitrage free condition.
To do so, it is necessary to calculate the profile on an hourly basis, which carries all seasonality and weather information with a mean representing the neutral element to the applied operation. Technically, all seasonality can be modeled on the highest frequency, hourly data for the price models. When we face limitations on data, i.e. unavailable hourly data or computational time issues, the model has to be divided in two elements, the daily model and the intraday model. In this paper we use the following notation: The day index for realized days is t and τ for future days. The index for the hours is j and k for the realized years. All prediction calculations are done at the last day t and the first element of the days τ. We denote by $S_{t,j}$ the spot data at day t and hour j.
The daily profile d_t is defined as

$$d_t = \bar{d}_t + \varepsilon_t, \tag{1}$$

with the expected value $\bar{d}_t$ and where ε_t defines a white noise (wn) process with zero expectation and a finite variance σ^2, i.e. $\varepsilon_t \sim wn(0, \sigma^2)$. The daily profile is the deviation from the expected value $\bar{d}_t$. The hourly profile is defined as

$$h_{t,j} = \bar{h}_{t,j} + \eta_{t,j}, \tag{2}$$

with the expected value $\bar{h}_{t,j}$ and where $\eta_{t,j}$ defines a white noise process with zero expectation and a finite variance σ^2, i.e. $\eta_t \sim wn(0, \sigma^2)$. The hourly profile is the deviation from the expected value $\bar{h}_{t,j}$. The yearly profile a_k is defined as

$$a_k = \bar{a}_k + \xi_k, \tag{3}$$

where ξ_k defines a white noise process with zero expectation and a finite variance σ^2, i.e. $\xi_t \sim wn(0, \sigma^2)$. The yearly profile is the deviation from the expected value $\bar{a}_k$.

The long term product price is the market expectation fair price for energy in this period, so the daily profile d_t and hourly profiles $h_{t,j}$ are modeled as weight for the existing traded long term future contracts F. The hourly prediction of the HPFC for future days τ and hours j is modeled as

$$\text{HPFC}_{\tau,j} = \bar{d}_\tau \bar{h}_{\tau,j} F_{\tau,j} \tag{4}$$

with $\mathrm{E}_\tau[\bar{d}_\tau] = 1$ and $\mathrm{E}_j[\bar{h}_{\tau,j}] = 1$.

2.1 Normalization of the Data

The first step is to model the seasonality and external impacts on the spot prices on daily basis. The estimation of the expected daily profile d_t is calculated as:

$$\hat{\bar{d}}_t = \frac{\sum_{j=1}^{NoH_t} S_{t,j}}{NoH_t} = \widehat{\mathrm{E}}_j[S_{t,j}], \tag{5}$$

where $S_{t,j}$ is the spot price of day t and hour j, NoH_t is the number of hours a day and $\widehat{\mathrm{E}}$ denotes the empirically estimated expectation.

The estimation of the expected yearly profile a_k is defined by the equation

$$\hat{\bar{a}}_k = \frac{\sum_{t=1}^{NoD_k} d_t}{NoD_k} = \widehat{\mathrm{E}}_t[d_t \,|t \in \mathfrak{K}_k], \tag{6}$$

where NoD_k is the number of days a year, k is the index for the years and $\mathfrak{K}_k$ is a set which contains all data of the full year. The full year means all days of the year, for example 15^{th} of March 2012 to the 14^{th} of March 2013. The sets $\mathfrak{K}_k$ form together the entire set $\mathfrak{K}$. Given the yearly profile a_k the normalized daily profile y_t is defined by

$$y_t = \frac{d_t}{a_k} = f_t + \vartheta_t \tag{7}$$

ϑ_t defines a white noise process with zero expectation and a finite variance σ^2, i.e. $\vartheta_t \sim wn(0, \sigma^2)$ and $\widehat{\mathrm{E}}[y_t] = 1$.

2.2 Estimation and Prediction of the Daily Weights

For the normalized daily profile estimation, we propose a model with a deterministic part $\hat{f}_t$ and a stochastic part ϑ.

$$\hat{y}_t = \frac{\hat{\bar{d}}_t}{\hat{\bar{a}}_k} = \hat{f}_t + \vartheta_t. \tag{8}$$

We choose a linear factor model for the deterministic element $\hat{f}_t$ of the form:

$$f_t = \alpha + \beta X_t, \tag{9}$$

where α and β are unknown parameters. The data for $X \in \mathbb{R}^w$ are assumed to be known. We estimate the parameters and the estimated expected daily factor

$$\hat{f}_t = \hat{\alpha} + \hat{\beta} X_t, \tag{10}$$

which is determined using an regression.

The ordinary least square estimator (OLS) minimizes a squared loss function which is the equivalent of the maximization of the log-likelihood function of the Gaussian distribution to determine the conditional expectations. This is the regression equivalent of the non-conditional expectation by using an sample average.

The predictions of the independent variables y_τ for future time τ are calculated by the linear prediction equation

$$\hat{y}_\tau = \hat{\alpha} + \hat{\beta} X_\tau. \tag{11}$$

To ensure the condition $\mathrm{E}_t[y_t] = 1$, the predicted daily average must be normalized by the predicted yearly profile

$$\hat{f}_{\tau\,|\tau\in\mathfrak{K}} = \frac{\hat{y}_{\tau\,|\tau\in\mathfrak{K}}}{\widehat{\mathrm{E}}_\tau[\hat{y}_\tau\,|\tau\in\mathfrak{K}]} \quad = \quad \frac{\hat{y}_k}{\widehat{\mathrm{E}}_k[\hat{y}_k]}. \tag{12}$$

2.3 Hourly Profile Calculation

To complete the price model (4), the hourly profile $h_{t,j}$ with the intraday seasonality has to be estimated. The intraday seasonality reflects the day and night cycle and behavior such as cooking around noon, illumination in the evening and so on. To ensure a statistical amount of data, hours of familiar days will be packaged together in clusters $\mathfrak{U}$ where $\mathfrak{U}_k$ is a set which contains hourly data of days which belong to a cluster of statistically similar days, i.e. days which exhibit similar behavior, such as summer weekdays. The union of sets $\mathfrak{U}_k$ form the entire set $\mathfrak{U}$.

With given clusters $\mathfrak{U}_k$, the hourly profile $h_{t,j}^{(k)}$ can be calculated by the equation

$$h_{t,j} = \frac{S_{t,j}}{d_t}, \tag{13}$$

and the corresponding hourly profile by

$$\hat{\hat{h}}_{t,j}^{(k)} = \widehat{\mathrm{E}}_t[h_{t,j}\,|t\in\mathfrak{U}_k], \tag{14}$$

where $\mathfrak{U}_k$ denotes the appropriate cluster of comparable days and satisfy the condition

$$\widehat{\mathrm{E}}_j[h_{t,j}] = 1, \tag{15}$$

the clusters are assumed as known.

Given the normalized factor estimation $\hat{f}_k$ and the hourly profile estimation $\hat{\hat{h}}_{t,j}$ the hourly residuals $\Phi_{t,j}$ are given by

$$\hat{\Phi}_{t,j} = S_{t,j} - \hat{\hat{a}}_k \hat{\hat{h}}_{t,j} \hat{f}_t. \tag{16}$$

To scale the residuals, so they are in the same dimension as the factors, the relative residuals $\phi_{t,j}$ can be calculated by removing the yearly average from the residuals

$$\hat{\phi}_{t,j} = \frac{\hat{\Phi}_{t,j}}{\hat{\hat{a}}_k} = \frac{S_{t,j}}{\hat{\hat{a}}_k} - \hat{\hat{h}}_{t,j}\hat{f}_t. \tag{17}$$

2.4 Building the HPFC

To calculate the HPFC, traded future products must be weighted by daily and hourly profiles d_t and $h_{t,j}$ which carry the estimated deterministic information. The future products must be applied to the curve in a way that the price of the curve is equal to the mean of the corresponding future during delivery time.

$$\mathrm{HPFC}_{\tau|\tau\in\mathfrak{F}_c,\,j|j\in\mathfrak{F}_c} = \hat{f}_{\tau|\tau\in\mathfrak{F}_c}\;\hat{\hat{h}}_{\tau|\tau\in\mathfrak{F}_c,\,j|j\in\mathfrak{F}_c}\;F_{\tau|\tau\in\mathfrak{F}_c}, \tag{18}$$

$\mathfrak{F}_c$ is the set of days which are the days during the delivery time of the future product c, i.e all days of the year 2016 belong to the future contract with delivery time in the year 2016. The sets $\mathfrak{F}_c$ form together the entire set $\mathfrak{F}$.

3 Estimation Calculation via the Median

In this section we propose the calculation of the expectation using a median estimator instead of the average estimator. We discuss the assumptions of mean and median estimator, define the model based on the median estimator and introduce the LAD-Lasso method as estimation and parameter selection technique for the regression problem.

3.1 Structure of the Data and Assumptions

The sample average is a bias-free estimator of the mean under the assumption of normally distributed data. Figure 2 shows the histogram of spot prices compared with the normal distribution. An estimator based on the normal distribution will be dominated by the jumps and tail events and results in a biased estimation. To handle this kind of data, we have to rely on estimators which are robust against non-normal distributed data [4]. In this paper, we use the median and the corresponding least absolute deviation (LAD) as estimators robust against non-normal distributed data.

3.2 Median as Robust Estimator of Expectations

The median $\mathrm{M}_n[x]$ is defined as

$$\widehat{\mathrm{M}}_n[x] = \begin{cases} x_{\frac{n+1}{2}} & n \text{ odd} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & n \text{ even.} \end{cases} \tag{19}$$

Median and mean absolute deviation (MAD) are based on the Laplace distribution

$$f(x|\mu, b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}} \tag{20}$$

and are the equivalents of mean and variance in the Normal distribution

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{21}$$

The maximum likelihood estimator of the Laplace distribution is defined by

$$\begin{aligned} l(x) &= \max_{\mu_t}\Big\{\prod_{t=1}^{T} f(x_t|\mu_t, b)\Big\} \\ \log_e l(x) &= \max_{\mu_t}\Big\{\log_e\Big(\prod_{t=1}^{T} f(x_t|\mu_t, b)\Big)\Big\} \\ L(x) &= \max_{\mu_t}\Big\{\sum_{t=1}^{T}\log_e\big(f(x_t|\mu_t, b)\big)\Big\} \\ L(x) &= \max_{\mu_t}\Big\{\sum_{t=1}^{T}\log_e\Big(\frac{1}{2b}\Big) - |\frac{x_t-\mu_t}{b}|\Big\}. \end{aligned} \tag{22}$$

with $\left(\frac{1}{2b}\right) = \text{const}$, μ_t denotes the conditional expecation and $b > 0$ the log-likelihood estimation becomes

$$L(x) = \max_{\mu_t} \Big\{ -\sum_{t=1}^{T} |\frac{x_t - \mu_t}{b}| \Big\}$$
$$L(x) = \min_{\mu_t} \Big\{ \sum_{t=1}^{T} |\frac{x_t - \mu_t}{b}| \Big\}, \quad (23)$$

where $x_t = y_t$. With $\mu_t = -\alpha - \beta X_t$ the LAD estimator is defined by

$$L = \min_{\alpha,\beta} \Bigg\{ \sum_{t=1}^{T} \mid y_t - \alpha - \sum_{i=1}^{N} \beta_i X_{t,i} \mid \Bigg\}. \quad (24)$$

As shown in [5], the mean $\widehat{\mathrm{E}}[x]$ and the median $\widehat{\mathrm{M}}[x]$ can be estimated by

$$\widehat{\mathrm{E}}[x] = \arg\min_{\bar{x}} \Big\{ \sum_{t=1}^{T} (x_t - \bar{x})^2 \Big\}, \quad (25)$$
$$\widehat{\mathrm{M}}[x] = \arg\min_{\bar{x}} \Big\{ \sum_{t=1}^{T} |x_t - \bar{x}| \Big\}. \quad (26)$$

3.3 Median Based Models

We will now introduce the model based on the median estimator for daily and hourly estimations. The symbol $\widehat{\mathrm{M}}_j[S_{t,j}]$ represents the calculation of the sample median, corresponding to the sample average calculation symbol $\widehat{\mathrm{E}}_j[S_{t,j}]$. We will show the changes of the model using the median estimator instead of the mean estimator for the estimation equations (5), (6), (12) and (14). The full set of equations for the models are shown in table 1 at the end of the paper.

According to the condition $\mathrm{E}_t[y_t] = 1$, for the median estimation the condition $\mathrm{M}_t[y_t] = 1$ must hold.

Equation (5) with the median estimator will be changed to

$$\hat{\hat{d}}_t = \widehat{\mathrm{M}}_j[S_{t,j}], \quad (27)$$

where $\widehat{\mathrm{M}}$ is the sample median estimation.

The yearly profile a_k is defined with the mean estimator in (6). The problem using the median estimator is given by

$$\hat{\hat{a}}_k = \widehat{\mathrm{M}}_t[d_t \,|t \in \mathfrak{K}] \quad (28)$$

where k is the index of the years. Given the yearly estimation a_k, the normalized daily estimation y_t can be defined correspondingly to the mean estimation as

$$y_t = \frac{d_t}{a_k} \quad (29)$$

with the condition $\mathrm{M}_t[y_t] = 1$.

Corresponding to the mean estimator the hourly profile $\bar{h}_{t,j}$ for given clusters $\mathfrak{U}$, is calculated by

$$\hat{\hat{h}}^{(k)}_{t,j} = \widehat{\mathrm{M}}_t[h_{t,j} \,|t \in \mathfrak{U}_k], \quad (30)$$

where $\mathfrak{U}_k$ denotes the appropriate cluster of comparable days with the condition

$$\widehat{\mathrm{M}}_j[h_{t,j}] = 1. \quad (31)$$

Concerning computational speed, implementations, problem size and dimension limitations, the median estimator is competitive with the mean estimator. Efficient algorithms for the median calculation are available for all major programming languages.

3.4 Model Selection via the LAD-Lasso

3.4.1 LAD Estimator

We introduced the median estimator as the estimator based on the Laplace distribution with the property of robustness against non-normal distributions. The consistent estimator choice for the parameter estimation problem (10), based on the median estimations of $\hat{\hat{d}}_t$ (27) and $\hat{\hat{a}}_k$ (28), is the LAD estimator (24) based on (23) and (26).

Unlike the unconstrained OLS regression problem, the LAD problem cannot be solved analytically and must be solved numerically. The optimization problem can be written as a linear programm (LP) [6] with the structure

$$\min_{e,f,\alpha,\beta} \quad \sum_{t=1}^{T} e_t + f_t \quad (32)$$
$$\text{s.t:} \quad y_t - \alpha - \sum_{i=1}^{N} \beta_i X_{t,i} = e_t - f_t \quad \forall\, t = 1 \ldots T$$
$$e, f > 0 \quad \forall\, t = 1 \ldots T$$

Because the optimization problem is of the LP type, it can be solved very efficiently even in very high dimensions. Given the nature of the LAD estimator based on a linear penalty function, this estimator, unlike the OLS estimator, is robust against non-normal distributed data.

3.4.2 Model Selection via the LAD-Lasso

After the introduction of the median and LAD estimator, we have to address the problem of overfitting and model selection. In estimation techniques for every new parameter introduced in the optimization problem, the fit will be better. The problem is that can be because of adding a parameter carrying information or because of overfitting. The idea of model selection is to identify the variables with carry deterministic information, all variables which do not carry process information are set to zero. Several methods have been developed to tackle that problem. A widely used approach is stepwise regression, but stepwise regression has been proved to be suboptimal [7] and is computationally intensive, especially in larger problem dimensions. Starting with the work of [8] with the Lasso, also known as $l1$-regularization and extended by [9], the LAD-Lasso solves the estimation and the model selection problem on the full parameter set in one optimization problem of the formulation

$$\min_{\alpha,\beta}\left\{\underbrace{\sum_{t=1}^{T}|y_t-\alpha-\sum_{i=1}^{N}X_{t,i}\beta_i|}_{\text{regression}}+\underbrace{T\sum_{i=1}^{N}\lambda_i|\beta_i|}_{\text{regularization}}\right\} \tag{33}$$

with $\lambda \geq 0$. The λ must be estimated in an a priori process [9]. The regularization term in (33) is always larger than zero for every parameter $\beta_i \neq 0$. The full optimization problem becomes smaller only if a parameter loading $\beta_i \neq 0$ decreases the regression part of the problem faster than it increases the regularization part. Therefore, only variables carrying information will be loaded in this minimization problem. This problem can also be written as a LP, see [9], and can thus be solved in high dimensions.

4 Results

4.1 *Test Framework*

We will compare the results of the different estimators in a test framework based on the German spot price time series. German spot prices contain high price jumps in positive and negative directions and are available since several years, see Figure 1. The data is raw data, we will discuss the truncation of the data, with the aim to increase robustness, at the end of the section.

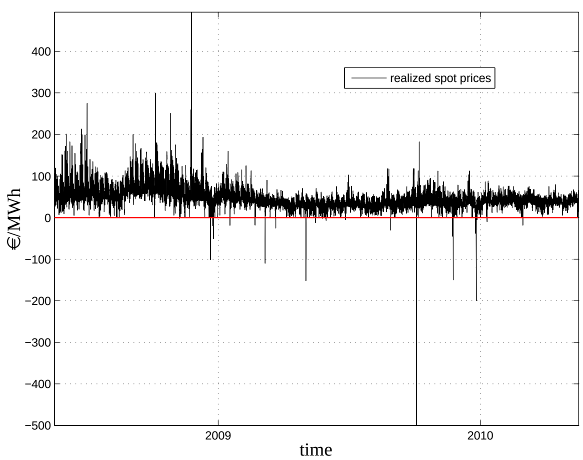

Figure 1: Realized hourly German spot price

As independent factors for the regression and model selection framework, we will use calendar information about days and months and weather information such as heating degree days (HDD), cooling degree days (CDD), average temperature and wind speed. The predictions of the weather factors will be done by the extrapolation of norm data calculated from 30 years of historical data. The reference temperature for HDD and CDD calculation depends on the country. The spot prices are provided by the European Power Exchange (Epex) and the historical weather information by Bloomberg®.

In real world applications daylight saving time is implemented. For our calculations, we assume that every day has 24 hours and ignore the daylight saving time for the estimation procedures. In the final HPFC we double the 2:00 to 3:00 hour in October and remove the 2:00 to 3:00 hour in March after the calculation to add daylight saving time. Hence this method effects only two of the 8760 hours of the year, the error on the arbitrage free condition because of this action is negligible.

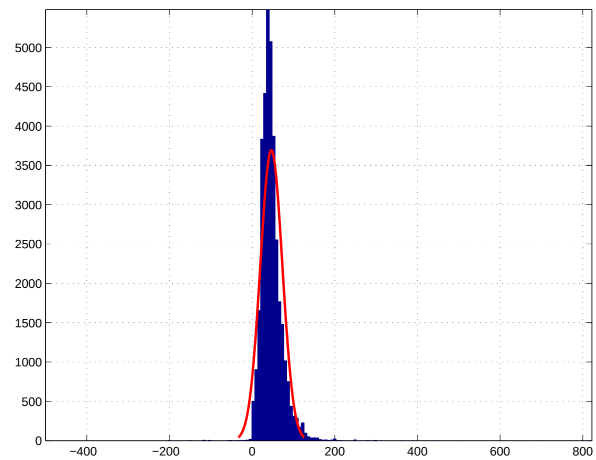

Figure 2: Histogram of hourly German electricity spot prices compared with the normal distribution

4.2 *Results*

The HPFC calculation is based on daily and hourly profiles. The daily estimation $\hat{\hat{d}}_t$ is the basis of the factor model to model seasonality and weather influences. Hence the major effect is covered by the daily estimation $\hat{\hat{d}}_t$ and the hourly estimation $\hat{\hat{h}}_{t,j}$ is modeled around the daily estimation $\hat{\hat{d}}_t$. The focus of the analysis will here be on the daily estimation $\hat{\hat{d}}_t$.
We discuss the following aspects in our analysis of the estimators:

1. Effects of large negative hourly prices on the daily estimation $\hat{\hat{d}}_t$,
2. Residuals of the daily estimation $\hat{\hat{d}}_t$,
3. Impact of outliers on the HPFC calculation,
4. Truncation of the Data.

At the end of the section, we briefly discuss the truncation of the price data to compensate the mean and OLS estimators lack of robustness. For all time series plots, the mean estimator time series are shown in black and the median estimator time series in red.

4.2.1 *Large Negative Hourly Prices*

Figure 1 shows the hourly spot prices in Germany from 2008 to 2010 with several occurrences of hours with negative prices. To discuss the effect of a few high negative prices on the estimators we will take the 4^{th} of October 2009 as an example. Figure 3 shows the histogram of the hourly spot data on the 4^{th} of October, with 20 hours with a positive spot prices and 4 hours with a negative spot prices, and the normal distribution fitted to the data. As discussed in section 3 the four negative hours, especially the -500.02 €/MWh price dominates the estimation of the OLS estimator and results in a negative daily estimation $\hat{\hat{d}}_t$

of -13.63 €/MWh because of one event under the assumption of the normal distribution. The estimation of the median estimator assuming the Laplace distribution results in a daily estimation $\hat{\bar{d}}_t$ of 17.17 €/MWh. The negative daily estimation $\hat{\bar{d}}_t$ of the OLS estimator is a miss estimation of the day, driven by mainly the one tail event. Independently from the hourly estimation $\hat{\bar{h}}_{t,j}$, the negative daily estimation $\hat{\bar{d}}_{t,j}$ results in a negative weight of the full day. Hence one negative price can be economical defended, a full negative day estimation would result in the shutdown of larger base load power plants in reality.

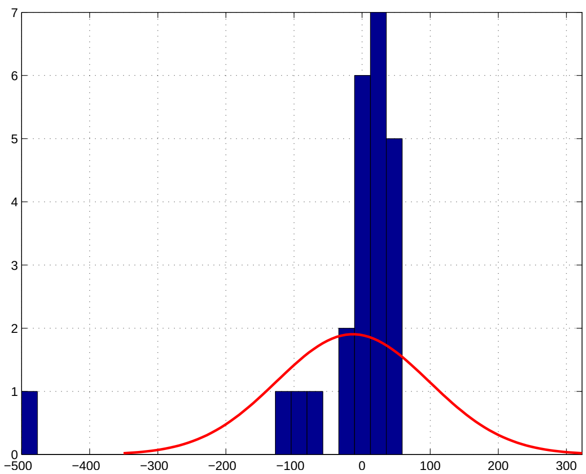

Figure 3: Histogram of German hourly spot prices on the 4^{th} of October 2009

4.2.2 *Residuals of the Estimations*

Indicators of robustness of the LAD-Lasso, against tail events, outliers and overfitting of the model selection, introduced in section 3.4.2, are the insample residuals of the estimation shown in Figure 4.
The residuals Φ_t are the part of the signal not explained by the deterministic part of the model resulting from the insample estimation, see (16). As shown in Figure 4, the negative daily estimations $\hat{\bar{d}}_t$ and the resulting negative normalized daily estimation $\hat{y}_t$ of the mean estimator are treated as noise by the LAD-Lasso model estimation framework. This is the result of the absolute value cost function which is less effected by tail events and outliers.

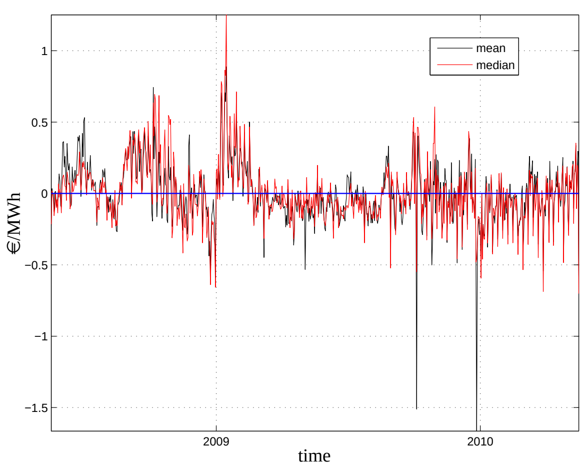

Figure 4: Daily insample residuals of the LAD-Lasso estimation

4.2.3 *Impact of Outliers on the HPFC*

Given the estimations $\hat{\bar{d}}_t$ and $\hat{y}_t$ and the discussion of the robustness of the LAD-Lasso estimator against tail events, we show the effect of the predictions on the HPFC. Qualitative checks like the correct positioning of the weekends and holidays are mostly driven by the estimation framework of the linear factor model $f_t = \alpha + \beta X_t$ and the independent variables X_t. This is mostly the task of a robust model selection estimator like the LAD-Lasso.

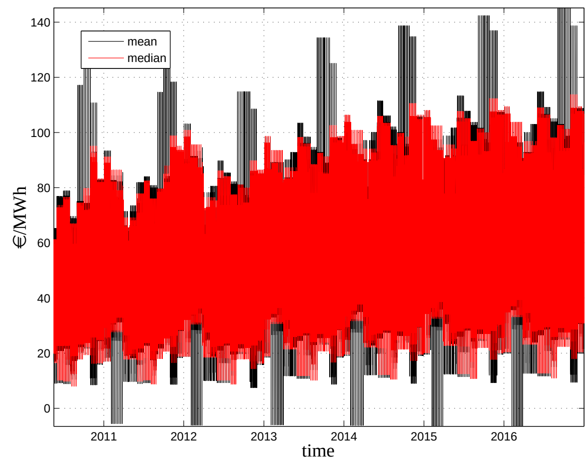

Figure 5: HPFC Calculated with mean (black) and median (red) estimation

The effect of the wrong estimations of $\bar{d}_t$ and y_t because of tail events are shown in Figure 5 with the comparison of the HPFC estimation using mean and median estimators. During September 2009 a group of French nuclear power plants were disconnected from the grid unexpectedly with the result of very high spot prices in Europe. The figure shows very high results of the price at the HPFC calculated by the mean estimator in the September month.The mean estimator completely overestimated the one high September months because of the strong impact of one exceptional month. As a result of the overestimation and the mean value constraints $\widehat{\mathrm{E}}_j[S_{t,j}] = 1$, the HPFC calculated based on the estimations $\bar{d}_t$ and y_t by the mean becomes very lacerated, which even results in prediction of negative prices in the HPFC and constant over estimations of September prices, which cannot be connected to any seasonal effect. The HPFC calculated by the median estimates of $\bar{d}_t$ and y_t results in a much more compact curve which is not influenced by the prediction of one high September spot.

4.2.4 *Truncation of the Data*

It is possible to truncate the data before the estimation to compensate the lack of robustness of mean and OLS estimators. Common actions are capping and manual weighting. Both methods result in additional parameters and thresholds, which have to be defined empirically. With the introduction of additional parameters, the danger of overfitting and insample optimization increases dramatically. The use of the median estimator does not need any data processing or additional parameters and allows this operation on pure statistical basis.

5 Conclusion

In this paper, we introduced a HPFC calculation method based on the median estimator and the LAD-Lasso as an estimator with model selection and robustness against non-normal distributed data including tail events, outliers and skewness. We showed the estimation results of the daily prediction d_t and the hourly price forward curve can be significantly improved by the median estimator compare to the mean estimator. We also showed the robustness against outliers of the LAD-Lasso for estimations of the deterministic effects in the HPFC model. The use of the median estimator allows a robust estimation without the need of any kind of data truncation, which lowers the chance of overfitting and insample optimization.

Acknowledgement

Financial support from KTI (Swiss Innovation Promotion Agency), project 10851.2, is gratefully acknowledged.

REFERENCES

[1] Hull, John, *Options, futures, and other derivatives*, 7th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2009, 2008010842 John C. Hull Ill. 1 CD-ROM Includes bibliographical references and indexes Windows 2000, NT, XP, ME, Vista.

[2] J. Hinz, L. Von Grafenstein, M. Verschuere, and M. Wilhelm, "Pricing electricity risk by interest rate methods," *Quantitative Finance*, vol. 5, no. 1, pp. 49–60, 2005, 981FE Times Cited:6 Cited References Count:22.

[3] Bloechlinger, Lea, "Power prices - a regime-switching spot/forward price model with kin filter estimation," *Dissertation School of Business Administration, Economics, Law and Social Sciences University of St. Gallen (HSG)*, 2008.

[4] Portnoy, Stephen and Koenker, Roger, "The gaussian hare and the laplacian tortoise: Computsbility of the squared-error versus absolute-error estimators," *Statistical Science*, vol. 12, no. 4, pp. 279–300, 1997.

[5] David, H. A. and Nagaraja, H. N., *Order statistics*, 3rd ed., ser. Wiley series in probability and statistics. Hoboken: Wiley Interscience, 2003, 2003050174 H.A. David, H.N. Nagaraja Includes bibliographical references (p. 367-450) and index Table of contents Online via SFX.

[6] Boyd, Stephen P. and Vandenberghe, Lieven, *Convex optimization*, 6th ed. Cambridge: Cambridge University Press, 2008, stephen Boyd, Lieven Vandenberghe Ill. Abstract / Index.

[7] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning data mining, inference, and prediction*, 2nd ed., ser. Springer series in statistics. New York, N.Y.: Springer, 2009, gFR-DNB-08,N30,0597 Trevor Hastie, Robert Tibshirani, Jerome Friedman Ill. Includes bibliographical references and index Online via SFX.

[8] Tibshirani, Robert, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[9] Wang, Hansheng and Li, Goudong and Jiang, Guohua, "Robust regression shrinkage and consistent variable selection through the lad-lasso," *Journal of Business Economics Statistics*, vol. Vol. 25, pp. pp. 347–355, 2008.

Table 1: Model comparison of mean and median based estimation

Model Element	No	Mean based model	Median based model	Name
Normalize	1	$\hat{\bar{d}}_t = \frac{\sum^{NoH} S_{t,j}}{NoH} = \widehat{\mathrm{E}}_j[S_{t,j}]$	$\hat{\bar{d}}_t = \widehat{\mathrm{M}}_j[S_{t,j}]$	daily average
	2	$h_{t,j} = \frac{S_{t,j}}{d_t}$	$h_{t,j} = \frac{S_{t,j}}{d_t}$	normalized hours
	3	$\hat{\bar{h}}_{t,j} = \widehat{\mathrm{E}}_t[h_{t,j} \mid t \in \mathfrak{U}]$	$\hat{\bar{h}}_{t,j} = \widehat{\mathrm{M}}_t[h_{t,j} \mid t \in \mathfrak{U}]$	hourly profile
	4	$\hat{\bar{a}}_k = \frac{\sum^{NoD} d_t}{NoD} = \widehat{\mathrm{E}}_t[d_t \mid t \in \mathfrak{K}]$	$\hat{\bar{a}}_k = \widehat{\mathrm{M}}_t[d_t \mid t \in \mathfrak{K}]$	yearly average
	5	$y_t = \frac{d_t}{a_k}$	$y_t = \frac{d_t}{a_k}$	normalized daily average
Daily Prediction	6	$y_t = \frac{d_t}{a_k} = f_t + \vartheta_t$	$y_t = \frac{d_t}{a_k} = f_t + \vartheta_t$	measurement equation
	7	$\hat{f}_t = \hat{\alpha} + \hat{\beta} X$	$\hat{f}_t = \hat{\alpha} + \hat{\beta} X$	regression equation
	8	$\hat{y}_\tau = \hat{\alpha} + \hat{\beta} X_\tau$	$\hat{y}_\tau = \hat{\alpha} + \hat{\beta} X_\tau$	prediction equation
	9	$\hat{f}_\tau = \frac{\hat{y}_\tau}{\widehat{\mathrm{E}}_\tau[\hat{y}_\tau]}$	$\hat{f}_\tau = \frac{\hat{y}_\tau}{\widehat{\mathrm{M}}_\tau[\hat{y}_\tau]}$	normalized factors
	10	$\hat{\Phi}_{t,j} = S_{t,j} - \hat{a}_k \hat{\bar{h}}_{t,j} \hat{f}_t$	$\hat{\Phi}_{t,j} = S_{t,j} - \hat{a}_k \hat{\bar{h}}_{t,j} \hat{f}_t$	residuum equation
	11	$\hat{\phi}_{t,j} = \frac{P\hat{h}i_{t,j}}{\hat{a}_k} = \frac{S_{t,j}}{\hat{a}_k} - \hat{\bar{h}}_{t,j} \hat{f}_t$	$\hat{\phi}_{t,j} = \frac{\hat{\Phi}_{t,j}}{\hat{a}_k} = \frac{S_{t,j}}{\hat{a}_k} - \hat{\bar{h}}_{t,j} \hat{f}_t$	relative residuals
Hourly Profile	12	$\hat{\bar{h}}^{(k)}_{t,j} = \widehat{\mathrm{E}}_t[h_{t,j} \mid t \in \mathfrak{U}_k]$,	$\hat{\bar{h}}^{(k)}_{t,j} = \widehat{\mathrm{E}}_t[h_{t,j} \mid t \in \mathfrak{U}_k]$,	hourly profile
HPFC Calculation	13	$\mathrm{HPFC}_{\tau,j} = \hat{f}_\tau \hat{\bar{h}}_{\tau,j} P_t$	$\mathrm{HPFC}_{\tau,j} = \hat{f}_\tau \hat{\bar{h}}_{\tau,j} P_t$	apply futures