Modeling electricity spot prices - Combining mean-reversion, spikes and stochastic volatility

Klaus Mayer*
Thomas Schmid
Florian Weber

Technische Universität München
Department of Financial Management and Capital Markets
CEFS Working Paper 2011-2

February 18, 2011

Abstract

Starting with the liberalization of electricity trading, this market grew rapidly over the last decade. However, while spot and future markets are rather liquid nowadays, option trading is still limited. One of the potential reasons for this is that the spot price process of electricity is still puzzling researchers and practitioners. In this paper, we propose an approach to model spot prices that combines mean-reversion, spikes and stochastic volatility. Thereby we use different mean-reversion rates for "normal" and "extreme" (spike) periods. Another feature of the model is its ability to capture correlation structures of electricity price spikes. Furthermore, all model parameters can easily be estimated with help of historical data. Consequently, we argue that this model does not only extend academic literature on electricity spot price modeling, but is also suitable for practical purposes, e.g. as underlying price model for option pricing.

Keywords: Electricity, Energy markets, Lévy processes, Mean-reversion, Spikes, Stochastic volatility, GARCH

JEL classification: G17

*Department of Financial Management and Capital Markets, Technische Universität München, Arcisstr. 21, D-80290 Munich, Germany, Tel. +49 89 289 25479, Klaus.Mayer@wi.tum.de.

List of Abbreviations

ACF	AutoCorrelation Function
EGARCH	Exponential Generalized Autoregressive Conditional Heteroscedasticity
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
RMSE	Root Mean Squared Error

List of Symbols

$\mathcal{F}$	Filtration
Γ	Affine transformation parameter
$\Lambda(t)$	Seasonality function value at time t
$\Phi(\cdot)$	Normal distribution function
$Sigma$	Correlation matrix
Ω	Probability space
α_x	Mean-reversion rate of the "normal" diffusion process
α_y	Mean-reversion rate of the "extreme" spike process
α_z	Mean-reversion rate of the unseparated process
γ	Regression coeffienct
$\varepsilon(t)$	Daily changes of Z(t) not caused by mean-reversion
$\lambda^{\pm}$	Jump intensity
$\sigma(t)$	Volatility function of the "normal" diffusion process
ζ	Spike vector
I(t)	compount Poisson process at time t
J_i	Jump height of the i-th Poisson process jump
$N^{\pm}$	Number of jumps at time t
P	Probability measure
S(t)	Spot price at time t
T	Time horizion
X(t)	"Normal" diffusion process at time t
Y(t)	"Extreme" spike process at time t at time t
Z(t)	Unseparated stochastic process at time t
$dB(t)$	Independent Increments of the Brownian motion
$dI(t)$	Independent Increments of the compound Poisson process
m(t)	Moving average filter at time t
sea(t)	Yearly seasonality at time t
w(t mod 7)	Weekly seasonality at time t

1 Introduction

Starting with the liberalization of electricity trading about one decade ago, electricity markets gained huge importance. However, electricity has several peculiarities that distinguish it from other types of commodities. The most notable difference is its non-storability - or at least the very high costs associated with its storage. This leads to several problems for the price modeling and especially for the pricing of related derivatives. Behind the background that electricity markets are quite young, the fact that research in this important field is still limited is not surprising.[1] Several recent studies address the question how electricity future prices are formed in the market (e.g. Wilkens and Wimschulte (2007), Redl et al. (2009), Botterud et al. (2010) or Furio and Meneu (2010)). Unlike futures markets, which are rather liquid, option trading is still underdeveloped in electricity markets. One of the potential reasons is that the price behavior of electricity is still puzzling researchers and practitioners. The price process in this market differs substantially from other commodity markets, with very high volatility and, even more important, spikes being far more common (at least partly as a consequence of the non-storability). Consequently, traditional models which mostly build on the assumption of Gaussian distributions are not suitable in this case. However, without a proper understanding of the price process and validated models the pricing of options is impossible. This is the main motivation of our paper: An increased understanding of the price process can increase the liquidity of option trading and the efficiency of the whole electricity market.

In commodity pricing literature, the most common approach is to model the logarithmic price through a mean-reverting process (Schwarz (1997), Lucia and Schwarz. (2002)). Like the Black-Scholes-Merton model, the mean-reverting process is based on the exponential treatment of the stochastic spot price (Black and Scholes (1973), Merton (1973)). If these models are applied for electricity, they can capture the mean-reversion of electricity prices, but fail to account for the huge and non-negligible observed spikes in this market. In order to capture the spike behavior of the electricity spot price dynamics, it is necessary to extend the model by a jump component. Merton (1976) first introduced this class of jump-diffusion models to model equity dynamics. Cartea and Figueroa (2005) apply this model to the English and Welsh electricity market and find that it offers a proper adjustment to the peculiarities of electricity markets. Geman and Roncoroni (2006) already

[1] A good overview on the literature about electricity price modeling is provided by Higgs and Worthington (2008).

discuss and try to fix the drawback of this model, namely that it only uses one unrealistically high mean-reversion rate, both for the diffusion and the jump process. However, a single mean-reversion rate for these two aspects is only of limited use because the price of electricity does exhibit spikes instead of classical jumps.[2] These spikes tend to revert very quickly, leading to a high rate of mean-reversion following a spike. In "normal" times without any spikes, the mean-reversion rate is much lower. Consequently, the use of a single mean-reversion factor results in a too slow removal of "extreme" price movements (spikes) and a too fast return to the seasonal trend in periods without "extreme" events.

This problem can be solved by separating the mean-reversion factors for the "extreme" and the "normal" process. A suitable theoretical approach for this purpose is described by Benth et al. (2003).[3] The model proposed by Benth et al. was calibrated by Kluge (2006) for the Scandinavian, British and the German electricity market. Thereby, he estimates the parameters for the diffusion process from historical data and assumes a constant volatility over time. However, this approach has several drawbacks. First, the parameters for the spike process are not estimated from the time series, but based on expert opinions. Second, this approach neglects that the volatility in electricity markets is stochastic over time. Deng (2001) compares the jump-diffusion model of Merton (1976) with constant and stochastic volatility and derives prices for different energy derivatives using the Fourier transform and shows that stochastic volatility is important. Escribano et al. (2002) provide extensive empirical tests on a wide range of markets and conclude that it is necessary to include jumps and stochastic volatility. Recent academic work by Chan and Gray (2006) and Bowden and Payne (2008) suggests that the EGARCH is the best volatility model for electricity prices. Furthermore, existing electricity price models are not able to account for the fact that spikes in different time series, e.g. electricity and gas prices, are not independent. We overcome this problem by modeling the correlation of spikes in different time series.

We contribute to the existing literature by proposing a new approach to model the electricity spot price that overcomes several drawbacks of existing models. In particular, we are - to the best of our knowledge - the first to present a self-contained model that simultenously (i) includes separate speeds of mean-reversion for the spike and the "normal" process, (ii) stochastic volatility, (iii) is estimated only with

[2]Consequently, we will denote these "extreme" events as spikes instead of jumps in the reminder of this paper.

[3]For example, Hambly et al. (2009) use this approach for a practical application, i.e. to value swing options.

historical data, and (iv) considers correlations in the spikes over different time series. Consequently, we argue that this model is suitable as price process model, e.g. in the context of option pricing, as (i) all distinct characteristics of the electricity spot price are included, and (ii) the parameters can easily be estimated with the help of historical price series.

2 Data Analysis

This section focuses on the description of the electricity price data used in this paper. Our analysis is based on data from the German electricity market, the European Energy Exchange (EEX), located in Leipzig. The data ranges from January 1^{st} 2002 to December 31^{st} 2009. The EEX distinguishes between peakload contracts, which deliver only in peak times, and baseload contracts, which deliver 24 hours. For our analysis, we use the Phelix peakload index as price indicator. However, beneath the peakload index two additional block contracts are traded at the EEX (off-peak 1 and off-peak 2). The prices of these three block contracts are the arithmetic mean of the delivery periods of the corresponding contracts which were setteled in the auction taking place on the previous day ("day-ahead market"). For the peakload index, this is the time from 8:00 to 20:00 o'clock, for the off-peak 1 contracts from 0:00 to 8:00 o' clock and for the off-peak 2 contracts from 20:00 to 24:00 o'clock. Figure 1 shows the spot price dynamics for all three price processes. In the following sections we will focus on the results for the peakload price process. The results for the other two process can be found in appendix.

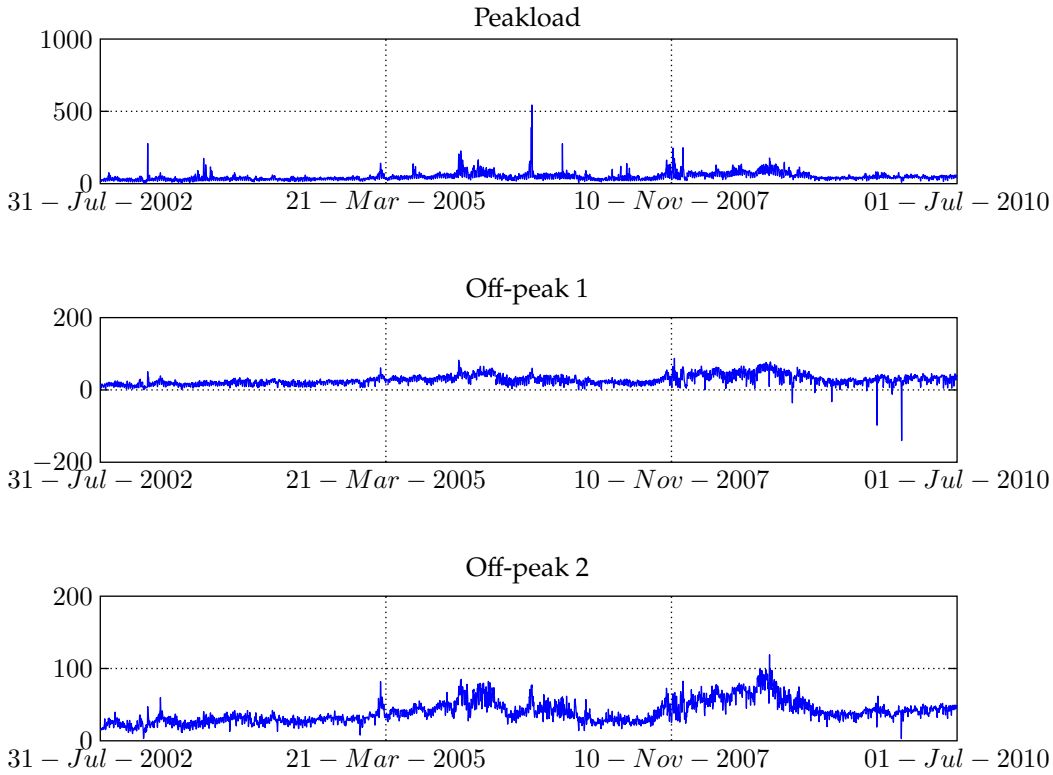

Figure 1: Peakload, off-peak 1 and off-peak 2 prices in Germany from 01/01/02-31/12/09

The Black-Scholes-Merton model (Black and Scholes (1973), Merton (1973)) assumes the prices are independently identically log-normally distributed, which implicates that the log returns of the prices follow a Normal distribution. We evaluate the validity of their assumptions in this section.

2.1 Negative prices

If the prices are indeed log-normally distributed, there should not be any negative prices. However, an additional distinctive feature of electricity is the existence of negative prices. The minimum value over the time series is $-139,96$ EUR. Especially the trading hours early in the morning, represented by the off-peak 1 contracts, are vulnerable to negative prices. This is the time of day with the lowest demand and, thus, the lowest prices. The negative prices occur due to a low demand and a (simultaneous) supply shock, for example because of unexpected high electricity transfers from wind turbines. Negative prices lead to several problems with the application of "classical" price process models which were usually developed for equity markets, where negative prices cannot occur. We solve this

problem by an affine transformation of all prices into the positive range. This does not distort the upcoming results.

2.2 Seasonality and Trend Analysis

We analyze the independence assumption for electricity prices with help of an autocorrelation test. If the data are in fact independently distributed, the autocorrelation coefficient should be close to zero. From figure 2 it can be seen that a strong level of autocorrelation exists in electricity markets (The qualitatively similar results for the off-peak 1 and off-peak 2 process are reported in figures 10 and 11). This observed autocorrelation is a result of an underlying seasonality, as discussed for instance in Pindyck and Rubinfeld (1998). In order to estimate the parameters of the price process model properly, we remove this seasonality from the return time series. After we removed the yearly and weekly seasonality, we analyze the ACF plot of the daily changes again. Without the seasonality, figure 2 shows a (strongly) significant negative autocorrelation only at lag 1 (Figures 10 and 11). However, this is not surprising since a negative autocorrelation at lag 1 is typical for a mean-reversion effect.

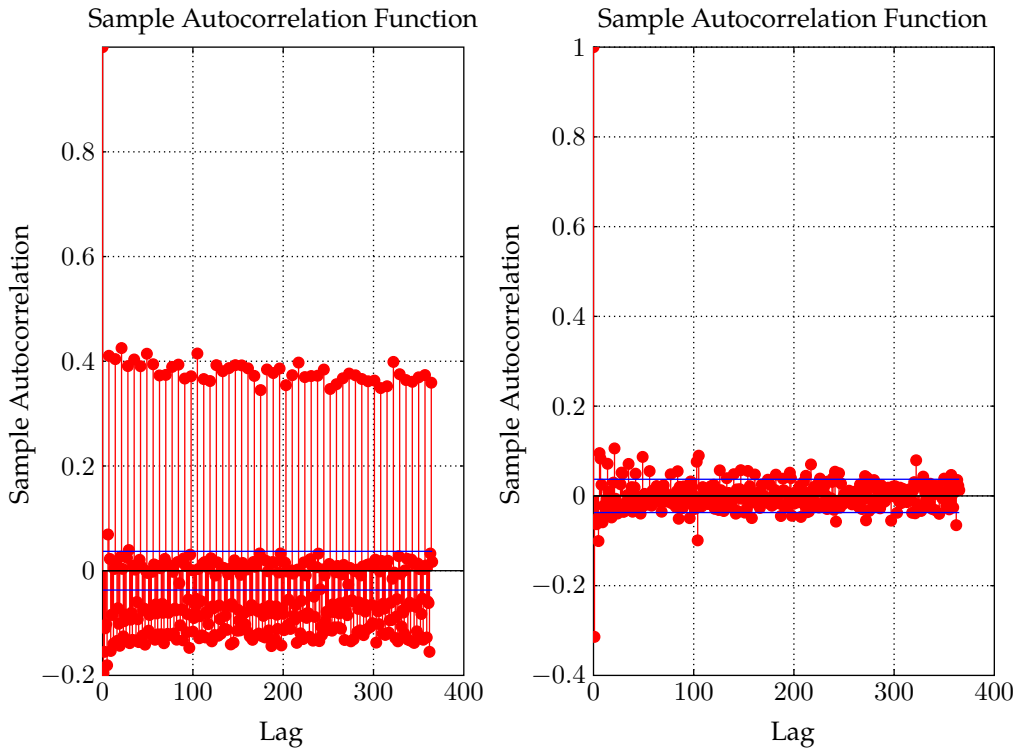

Figure 2: ACF plots of the log price and the log price without seasonality (peak-load process)

2.3 Normality test

The last assumption of the Black-Scholes-Merton model is that returns are normally distributed. Although the analysis of stock markets' data reveals a higher probability for an "extreme" event than predicted by the Normal distribution, the assumption is still embedded in most stochastic models.

However, for electricity spot prices, the deviation from normality is more extreme than e.g. for equity and most other types of commodities. Figure 3 shows normality tests for the peakload price process from January 1^{st} 2002 to December 31^{st} 2009 (The qualitatively similar results for the off-peak 1 and the off-peak 2 price process are reported in figures 8 and 9). If the empirical returns are normally distributed, we expect to observe a straight line in this figure. However, this is not the case. Contrary, we find clear indication for fat tails and hence more "extreme" events than predicted by the Normal distribution. As an example, the probability for a daily return of +/- 50 percent is virtually zero for the Normal distribution. However, we observe a non-negligible amount of such events in the time series.

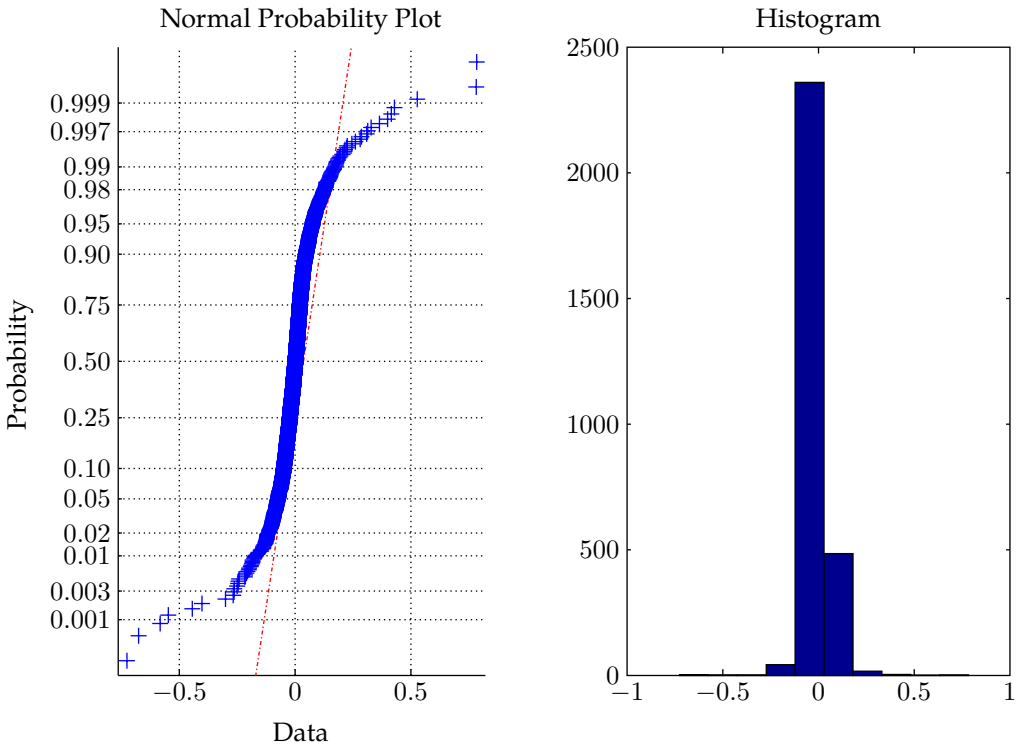

Figure 3: Normal probability plot and histogram of the unseperated peakload process

2.4 Spikes

Consequently, we argue that electricity spot prices do not follow a Gaussian distribution because too much "extreme" events occur in the time series. In a next step we extract these "extreme" events, the spikes, from the original time series by using a numerical algorithm that recursively filters returns with absolute values that are greater than 2.3 times the standard deviation of the returns of the series at that specific iteration. We end up with the "normal" return series consisting of 99 percent of the original returns.

The importance of the spikes in the electricity return series is illustrated by a simple comparison of figures 3 and 4 (for off-peak 1 and 2, refer to figures 8, 9 and 12, 13). After the extraction of the spikes from the original series, the assumption of a Gaussian distribution is more accurate. In fact, the fat tails nearly disappeared and the deviation of the empirical returns from the returns predicted by a Gaussian distribution becomes negligible.[4]

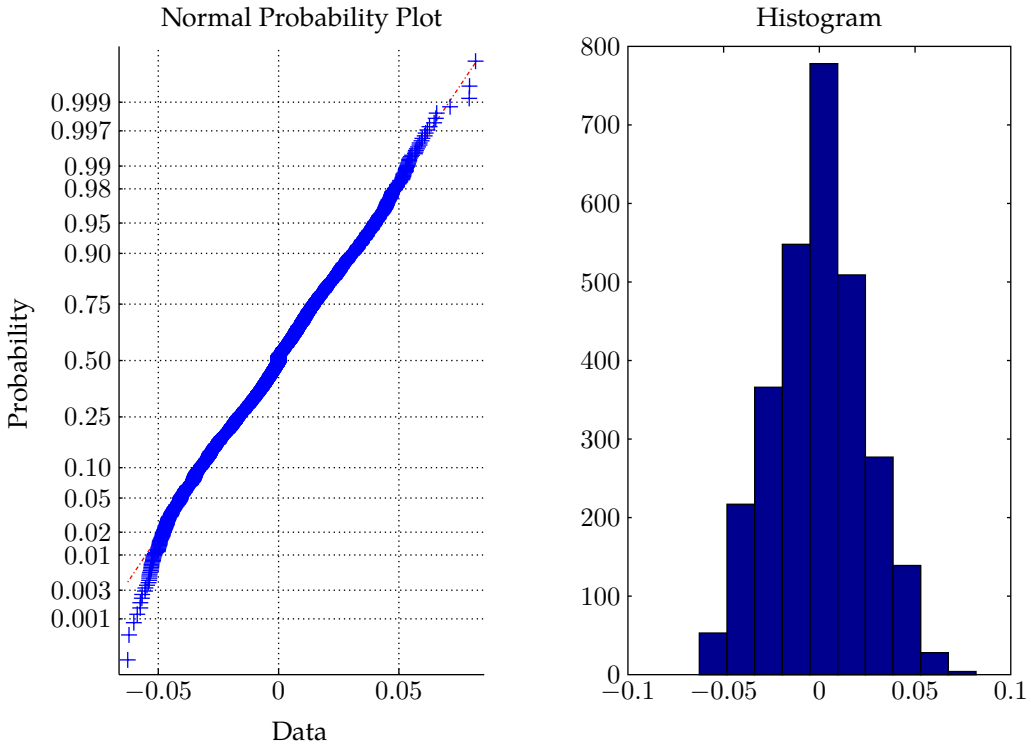

Figure 4: Normal probability plot and histogram of the diffusion part (peakload process)

[4]To be more precise, the deviation is still present, but less extreme. The assumption of a Gaussian distribution despite the fact that there exist small deviations is common, e.g. for equity prices.

Our descriptive analysis revealed that electricity prices are fluctuating around a long-term deterministic seasonal component and are characterized by price spikes and mean-reversion. Furthermore, our analysis suggest that assuming a Gaussian process is inaccurate for the electricity spot price.

Since the seminal work of Samuelson (1965), Black and Scholes (1973) and Merton (1973), most stochastic models have tried to offer a solution to the problem that empirically observed prices do not follow Gaussian distributions. This literature discussed jump-diffusion processes, stochastic volatility models, and, more recently, the use of Lévy processes.[5] While deviations from the normality assumptions are even common in equity and other commodity markets, their magnitude is far higher for electricity. Consequently, these models cannot be used to model the price process of electricity. Our model, which is described in the next section, addresses this non-normality problem by considering stochastic volatility and separate mean-reversion parameters for the spikes and the "normal" electricity spot process.

3 Theoretical framework

The data analysis of the previous section reveals three distinctive characteristics of electricity markets which should be accounted for in the model. First, negative rates have occurred over time. These is problematic for the application of geometric models. To cope with this problem, we apply an affine transformation of the time series into the positive range. An affine transformation by a factor of $\Gamma = 200$ EUR shifts all prices in the positive range and leaves ample room for future occurrence of negative prices in the subsequent price simulation. At the end of the simulation, the generated time series must be transformed back by an inverse transformation to the original price level.

The second characteristic is a seasonality component which reflects a (varying) long-term equilibrium level.Schwarz (1997) proposes a model that includes mean-reversion, and Lucia and Schwarz. (2002) extend this model by including mean-reversion and a deterministic seasonality. The third characteristic of electricity spot prices, namely the observation that the randomly fluctuation prices revert slowly back to the equilibrium level in times without "extreme" events while price spikes revert very quickly, is addressed by Benth et al. (2003). Their model in-

[5] See for example Merton (2001), Knight and Satchell (2001) or Shrevel (2004).

cludes separate rates of mean-reversion for the diffusion and the spike process. However, none of these models accounts for stochastic volatility and correlation among the spike processes over different related time series (e.g. electricity peak and off-peak price). For risk management purposes, the consideration of this correlation is of huge importance, e.g. in the context of the valuation of project financed investments. In this paper, we propose a similar model that is able to handle all these characteristics of electricity spot prices. For this, we extend the model of Benth et al. (2003) to (i) stochastic volatility, (ii) correlations of spikes, and (iii) a self-contained estimation of all relevant parameters.

We require that the standard assumptions for stochastic models hold. Formally, $(\Omega, P, \mathcal{F}, \{\mathcal{F}_t\}_{t\epsilon[0,T]})$ is a complete filtered probability space, with $T < \infty$ a fixed time horizon. If $S(t)$ denotes the spot price of electricity at time t, then we set

$$S(t) = (S(0) + \Gamma) \times \exp\left(\Lambda(t) + X(t) + Y(t)\right) - \Gamma \tag{1}$$

where $\Lambda(t)$ denotes the seasonality function value at time t, $X(t)$ and $Y(t)$ the values of two Lévy processes at time t.

Several subsequent steps are necessary to calibrate our model. *First*, we remove the seasonality component Λ of the original time series. Therefore, we set

$$\Lambda(t) = w(t \mod 7) + sea(t) \tag{2}$$

where $w(t \mod 7)$ denotes the weekly seasonality function at day k following (4) and $sea(t)$ the annual seasonality function value at time t following (5). The standard procedure to model seasonal influences is by trigonometric functions. However, the weekly behavior of the price of electricity follows a special pattern. It is significantly higher on weekdays than at the weekend. Therefore, we do not use trigonometric functions but instead a procedure proposed by Weron (2006). First, the data are smoothed using a moving average filter:

$$m(t) := \frac{1}{7}\left(S(t-3) + \cdots + S(t+3)\right). \tag{3}$$

Second, for each day the average deviation $w(t \mod 7)$ from moving average value $m(t)$ in (3) is calculated:

$$\begin{aligned} w(t \mod 7) := average\{(S((t \mod 7) + 7j) - m(k + 7j)), \\ 3 < (t \mod 7) + 7j \leq n - 3\}; \end{aligned} \tag{4}$$

The calculated values of $w(t \mod 7)$ are then normalized in a way that they add up to zero over one week. The resulting weekly seasonality $w(t \mod 7)$ is deducted from the original time series. A logarithmic transformation is performed before analyzing the long term seasonality function as well as the stochastic components. The annual seasonality, the trend and the level is included in the following function:

$$sea(t) = \ln\left(\beta_1 + \beta_2 t + \beta_3 \cos\left(\frac{2\pi(t - \beta_4)}{365}\right)\right) \tag{5}$$

Second, we include the trend factors of X and Y into β_2 of (5) . For this purpose, we assume that X and Y are Lévy processes where X is driven by a Brownian motion and Y by a compound Poisson process. Consequently, X and Y are zero level mean-reverting stochastic processes which follow the stochastic differential equations:

$$dX(t) = -\alpha_X X(t)dt + \sigma(t)dB(t) \tag{6}$$

$$dY(t) = -\alpha_Y Y(t)dt + dI(t) \tag{7}$$

where α_X and α_Y denote the mean-reversion parameter, $\sigma(t)$ the volatility of the diffusion process, $dB(t)$ the Brownian motion increments, and $dI(t)$ the increments of a compound Poisson process. Furthermore, we assume the mean-reversion parameters of (6) and (7) are constant until time T. For the volatility $\sigma(t)$, we assume that it can either be constant or stochastic. Stochastic models are either GARCH-type (Bollerslev (1986)) or EGARCH-type (Nelson (1991)) volatility models. However, the implementation of other types of volatility models is straightforward.

Third, we separate the spikes from the "normal" diffusion process. For this, we define a range to classify returns either as "normal" or "extreme", i.e. as spike. The value of this range is set recursively by removing high/low price movements as long until the distribution of the diffusion process can be classified as normal. We end up with a range of +/- 2.3 times the standard deviations of the time series. About 99% of the price movements lie in this range and are classified as "normal". Of course, the constant volatility model has a constant range (in absolute terms) to classify daily changes of σdB as spikes of the process dI. Contrary, GARCH and EGARCH volatility models have a variable range since the standard deviation is not constant over time. This variable range to classify spikes is determined each day by the multiplication of the barrier parameter with the volatility at that day. Consequently, the classification of price movements as "normal" or spike can vary between the constant volatility model and the GARCH or EGARCH volatility models. The reason for this is that in times of high (low) volatility the stochastic volatility models allow larger (smaller) movements of the process σdB without classifying them as spikes.

In a next step, we define the process for the spikes. For this, the process I in (7) is modeled with two separate compound Poisson processes for the positive and negative spikes.

$$I(t) = I^{+}(t) + I^{-}(t), \tag{8}$$

with

$$I^{\pm} = \sum_{i=1}^{N^{\pm}(t)} J_i^{\pm} \tag{9}$$

with

$$N^{\pm}(0) = 0 \tag{10}$$

$$E(N^{\pm}(t)) = \lambda^{\pm} \times t \tag{11}$$

$$Var(N^{\pm}(t)) = \lambda^{\pm} \times t \tag{12}$$

$$lnJ_i^{\pm} \sim N(\mu^{\pm}, \sigma^{\pm}) \tag{13}$$

The processes N^+ and N^- are Poisson Processes which represent the number of positive or negative spikes until time t. The jump intensities[6] λ^{pm} for the Poisson processes are estimated from the number of positive respectively negative spikes compared to all observed days. $J_i^{\pm}$ is the jump height i, which is assumed to be log-normally distributed.

4 Calibration and Results

In this section we demonstrate the calibration of all necessary model parameters. To take into account the significant differences in the electricity price during a day, we separately calibrate the parameters for the peak load and the off-peak prices, i.e. the periods before and after the peak load. Furthermore, we compare the modeled electricity price process with realized prices.

4.1 Seasonality Function

After the previously described affine transformation in the positive range we analyze the weekly seasonality in the price process S. We calculate the weekly and yearly seasonality as well as the drift according to (4) and (5) as described in section 3. Table 1 shows the seasonal function and trend parameters for the peakload price process. One distinctive feature of the off-peak 2 process to mention is that the prices already start to increase on Sunday (Table 10). This can be explained by the start-up of industrial companies' machinery.

	Seasonality and drift parameter						
Weekly seasonality	3.1297	7.9410	8.2671	6.4057	2.3192	−9.2020	−18.8607
Drift	$0.0127t + 229.0927$						
Yearly seasonality	$-4.6969 \cdot \cos\left(\frac{2\pi(t-129.0719)}{365}\right)$						

Table 1: Seasonality and drift parameters (peakload process)

[6]We use the term jump when we describe the properties of the compound Poisson processes although we use the term spike when we describe the events.

4.2 Mean-Reversion Rate

After removing the seasonality from the sample time series, we analyze the stochastic part. We start with determining the mean-reversion rates α_X and α_Y. We construct Z with

$$Z(t) = X(t) + Y(t) \tag{14}$$

and therefore

$$dZ(t) = dX(t) + dY(t). \tag{15}$$

Using (6) as well as (7) and the discrete version of (15) it follows

$$\begin{aligned}
\Delta Z(t) &= \Delta X(t) + \Delta Y(t) && (16)\\
&= -\alpha_X X(t)\Delta t + \sigma \Delta B(t) - \alpha_Y Y(t)\Delta t + \Delta I(t)\\
&= -(\alpha_X X(t) + \alpha_Y Y(t))\Delta t + \sigma \Delta B(t) + \Delta I(t) && (17)\\
&= -\left(\alpha_X \tfrac{X(t)}{X(t)+Y(t)} + \alpha_Y \tfrac{Y(t)}{X(t)+Y(t)}\right)(X(t)+Y(t))\,\Delta t + \sigma \Delta B(t) + \Delta I(t)\\
&= -\alpha_Z Z(t)\Delta t + (\sigma \Delta B(t) + \Delta I(t)). && (18)
\end{aligned}$$

Since the time series consists of daily data, a discretization of the interval length $\Delta t = 1$ is reasonable. Using linear regression of log-prices against the log-returns as in (19), we find a value for α_Z.

$$\Delta Z(t) = \gamma Z(t)\Delta t + \varepsilon(t) \tag{19}$$

where γ is the regression coefficient and ε describes the daily changes not caused by the mean-reversion effect. Using $\alpha_Z = (-1) \times \gamma$, it is possible to estimate the mean-reversion rate.

By removing the mean-reversion effect from the process Z we get the process ε. From this process, we determine a vector ζ of spikes using the tolerance parameter of 2.3 times the standard deviation. The mean-reversion rate α_Z and the spike vector ζ are then used to determine the mean-reversion rates α_X and α_Y.

For the calculation of X and Y, the stochastic part needs to be divided between the processes σdB and dI, first. As the final values of σdB and dI change because

of the mean-reversion, we name the processes initially σdB_0 and dI_0. Generally, we set as a first step $\sigma dB_0 = dZ$ and $dI_0 = 0$. However, at the spike points the value of σdB_0 is set to 0 and the values of dI_0 at these points equal to the values of dZ. Afterward, the processes X and Y are calculated recursively. We assume as starting values $X(0) = Z(0)$ and $Y(0) = 0$. The stochastic influences of σdB and dI have to be adjusted for the mean-reversion effect. This adjustment is necessary to offset the movement caused by the mean-reversion by the means of the stochastic components, to achieve the same movement intensity as in the process dZ . Generally, the mean-reversion effect is associated with the "normal" part σdB and is deducted from σdB_0 . The only exception are movements, where an association with σdB would lead to a new spike. In this case it is associated with spike process dI. The adaption formulas for σdB and dI are

$$\sigma dB(t) = \sigma dB_0(t) - (-\alpha_X X(t) - \alpha_Y Y(t)) \tag{20}$$

and in the case that it would lead to a new spike

$$dI(t) = dI_0(t) - (-\alpha_X X(t) - \alpha_Y Y(t))\,. \tag{21}$$

Here, $\sigma dB_0(t)$ and $dI_0(t)$ on the right hand side of equations (20) and (21) represent in each case the first assignment and $\sigma dB(t)$and $dI(t)$ on the left hand side of the equations (20) and (21) represent the final values. With these values the iterative update formulas for X (22) and Y (23) are

$$\begin{aligned} X(t+1) &= X(t) + dX(t) \\ &= (1-\alpha_X)X(t) + \sigma dB(t) \end{aligned} \tag{22}$$

$$\begin{aligned} Y(t+1) &= Y(t) + dY(t) \\ &= (1-\alpha_Y)Y(t) + dI(t) \end{aligned} \tag{23}$$

When the recursive filtering procedure is finished the processes X and Y can have different spikes than we originally assumed for the calculation. Therefore, we determine the spikes in $\sigma dB + dI$ again with a second algorithm. This algorithm

classifies all price movements that are larger than 2.3 times the standard deviation of the process σdB as spikes, but in contrast to the first algorithm it does not perform any recursions. Instead the spikes are calculated solely on the specified volatility. This enables us to take stochastic volatility into account. Although not using a recursive approach might result in a very low number of price moment not to be identified as spikes, we use this approach because otherwise the run-time of the algorithm becomes extraordinary long and the accuracy does not increase significantly. The required stochastic volatility processes are determined by the maximum likelihood estimation from the process dB .

With the obtained processes X and Y, the mean-reversion factors α_X and α_Y can be determined in analogy to 19. These parameters will then replace the initial values for the mean-reversion rates and the calculation of the processes X and Y starts again until the mean-reversion rates converge, too. In the case of alternating mean-reversion factors those factors are averaged and the calculation of the processes X and Y is done again. The procedure gives us as results the mean-reversion factors α_X and α_Y, the processes X and Y, the spike vector with the entries of the spikes, and the stochastic parts σdB and dI of the processes X and Y. The mean-reversion rates for the peakload process and the different volatility models are shown in table 2.

Volatility Model	Constant	GARCH	EGARCH
α_X	0.1064	0.1073	0.1069
α_Y	0.4802	0.4779	0.4744

Table 2: Mean-reversion parameters with different volatility models (peakload process)

4.3 Volatility and Correlation analysis

We analyze the distribution of the stochastic part σdB and dI. Figure 5 shows the division between the two processes σdB and dI for the three different volatility models. By analyzing the individual stochastic processes, we determine the stochastic processes B and I for each price processes. The combination of the six sub-processes results in the overall model with the random changes.

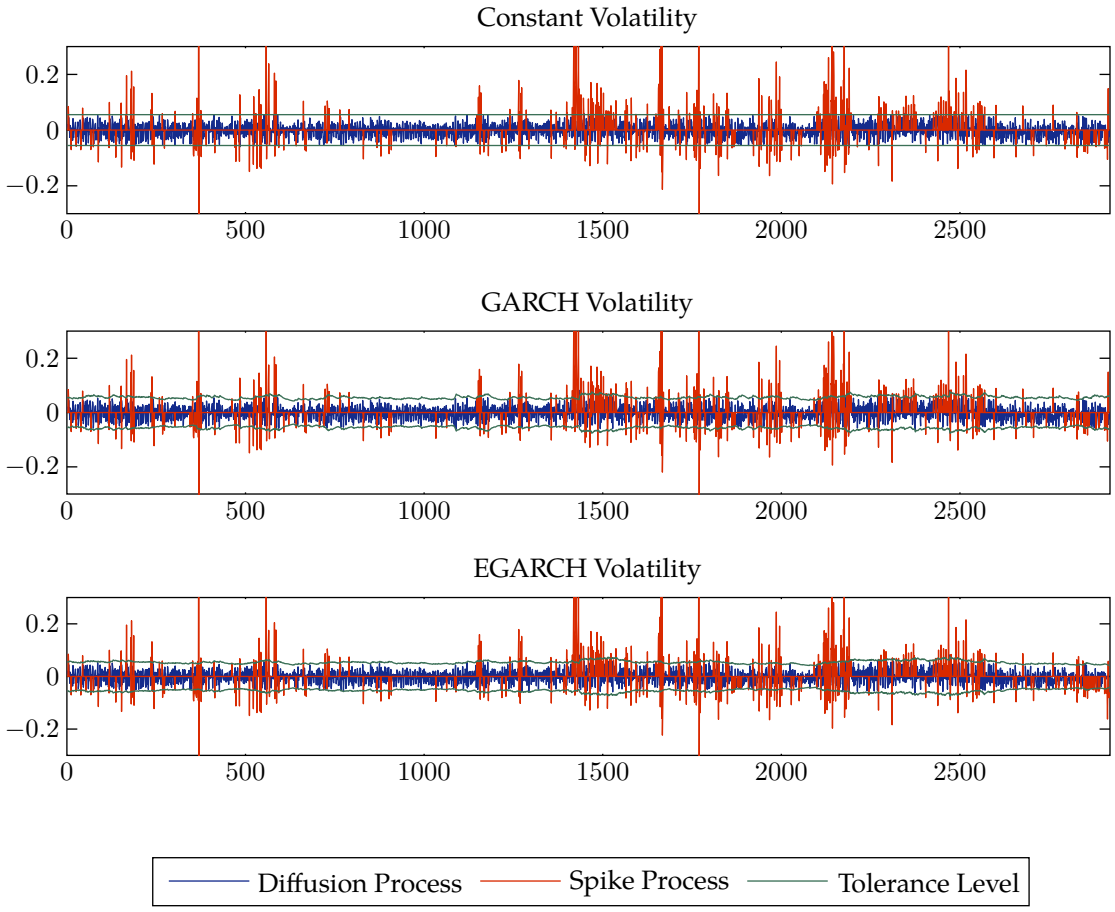

Figure 5: Spike detection for the three different volatility models

The standard assumption in most models, including the model of Benth et al. (2003), is to assume independence between the spike processes of different time series. However, correlation between spikes in different electricity spot prices is a common phenomenon. For example, price spikes caused by a shut down of a large power plant will most likely have an impact on the whole day, i.e. the peakload, off-peak 1 and off-peak 2 price processes. Even spikes in other price series, e.g. gas, can lead to spikes in the electricity price series. Consequently, the assumption of independence for the spike processes over different time series has to be rejected and a suitable electricity price model has to consider this dependency.

4.3.1 Diffusion process

When analyzing the distribution of σdB, we assume that for the stochastic volatility models GARCH and EGARCH a normalized process B, with zero mean and standard deviation of 1, exists. This allows us to use a constant parame-

ter σ, while the volatility of the normalized process B is stochastic. The parameters for the stochastic volatility are estimated from the normalized process dB by maximum-likelihood. Table 3 shows the distribution parameters and the GARCH-parameters for the peakload price process. The results indicate a positive leverage effect.

Volatility Model	Constant	GARCH	EGARCH
	Distribution Parameter		
μ	0	0	0
σ	0.0241	0.0245	0.0236
	GARCH Parameter		
κ	1	0.0304	0.0003
ARCH	0	0.0402	0.0530
GARCH	0	0.9297	0.9673
Leverage	0	0	0.0349

Table 3: Volatility parameters (peakload process)

After analyzing the individual price processes, we analyze the correlation structure in order to obtain the final parameters for the simulation. Tables 4, 5 and 6 show the correlation structure between the three electricity price processes σdB using different types of volatility. It is apparent that the correlation is the highest with EGARCH volatility and the lowest with GARCH volatility.

	Peakload	Off-Peak 1	Off-Peak 2
Peakload	1	0.3029	0.3207
Off-Peak 1	0.3029	1	0.2146
Off-Peak 2	0.3207	0.2146	1

Table 4: Correlation matrix Σ between the processes using constant volatility

	Peakload	Off-Peak 1	Off-Peak 2
Peakload	1	0.2950	0.3314
Off-Peak 1	0.2950	1	0.1949
Off-Peak 2	0.3314	0.1949	1

Table 5: Correlation matrix Σ between the processes using GARCH volatility

	Peakload	Off-Peak 1	Off-Peak 2
Peakload	1	0.3219	0.3537
Off-Peak 1	0.3219	1	0.2296
Off-Peak 2	0.3537	0.2296	1

Table 6: Correlation matrix Σ between the processes using EGARCH volatility

4.3.2 Spike process

As already mentioned dependencies between the spikes are very likely. Two types of dependencies are possible (i) among the jump intensity, i.e. the point of time, when a spike occurs, and (ii) the jump height, i.e. the size of the spike. We test for both dependencies. First, we analyze the jump intensity where a dependency is likely, because the events that have an effect on the price of electricity at a time of day may be longer term in nature and, thus, affect the subsequent contracts.

In order to test for the dependency, we identify all days on which one of the processes had a positive spike or a negative spike. Both for the positive and the negative spikes, there are 7 possible combinations[7] of spikes in the electricity price processes. First, there is the possibility that all three processes have a spike at the same day. Second, there is the possibility that two of the three processes may spike with variations which two processes spike. Third, there is the possibility that only one process has a spike. The derived likelihood for all 7 possibilities are presented in table 7.

The summation of the probabilities for a process leads to the jump probabilities for each process. For example, the resulting probabilities for a combination of a positive spike of the peakload process present a positive spike in one of other processes is 56.67% $(= 5.15\% + 8.43\% + 13.58\% + 29.51\%)$. Multiplying this by the probability that there is, indeed, a positive spike in the process gives the jump intensity λ^p for the peakload process of $8.29\%(= 56.67\% \times 14.63\%)$.

Second, we analyze the dependency among the jump heights. The analysis does not provide significant values and, therefore, we assume that no dependency ex-

[7] A total of $2^3 = 8$ spike combinations, less the combination of no spikes results in 7 possible combinations.

Volatility model			Constant		GARCH		EGARCH	
Combinations			Jump probablities					
Peak	OP 1	OP 2	positive	negative	positive	negative	positive	negative
x	x	x	5.15%	6.40%	3.50%	7.09%	4.20%	7.49%
x	x	0	8.43%	7.82%	5.75%	7.09%	4.94%	8.45%
x	0	x	13.58%	7.82%	16.50%	9.29%	15.80%	9.90%
x	0	0	29.51%	20.38%	32.00%	22.49%	34.32%	27.78%
0	x	x	3.28%	4.27%	1.75%	3.18%	2.72%	2.66%
0	x	0	20.54%	32.94%	18.25%	27.87%	16.79%	21.26%
0	0	x	19.20%	20.38%	22.25%	22.98%	21.23%	22.46%
Jump intensity			14.63%	14.46%	13.70%	14.01%	13.87%	14.18%

Table 7: Probablities of the jump combinations with a GARCH volatility diffusion process

ists between the jump heights. The jump distributions' parameters for the peakload process are shown in table 8.

Volatility Model	Constant	GARCH	EGARCH
μ^+	-2.3606	-2.3569	-2.4091
σ^+	0.6262	0.6415	0.6507
λ^+	0.0829	0.0791	0.0832
μ^-	-2.5387	-2.5614	-2.6168
σ^-	0.4276	0.4496	0.4442
λ^-	0.0613	0.0644	0.0757
# Jumps	421 (242 | 179)	419 (231 | 188)	464 (243 | 221)

Table 8: Parameters of the jump distributions (peakload process)

4.4 Results

To test the model we perform a Monte Carlo simulation for each price series (peakload, off-peak 1,off-peak 2) and each volatility model (constant volatility, GARCH, EGARCH) with $K = 100000$ simulation paths for the period between January 1^{st} and June 30^{th} 2010 ($T = 181$ days). The results are compared to the realized electricity price for that period. Figure 6 shows the realized peakload prices and a simulated price path (for off-peak 1 and 2, refer to figures 14 and 15). The model is able to capture the mean-reversion of the spikes and reverts to its seasonality function in the long run, as expected from (1) with (6) and (7).

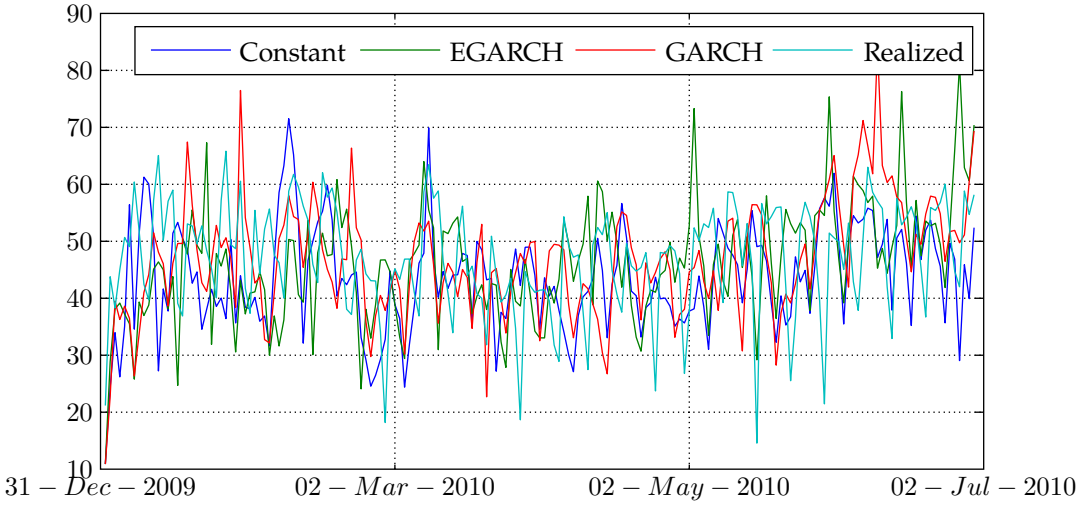

Figure 6: Spot and Simulated Spot Prices (peakload process)

We use the root mean squared error as goodness-of-fit criteria. It is defined as

$$SE_{t,k,v,a} = \left(S^{sim}_{t,k,v,a} - S^{rel}_{t,a}\right)^2 \qquad \forall 1 \leq k \leq K, 1 \leq t \leq T \tag{24}$$

$$RMSE_{i,k,a} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(SE_{t,k,v,a} - \overline{SE}_{k,v,a}\right)} \qquad \forall 1 \leq k \leq K \tag{25}$$

where $SE_{t,n,v,a}$ denotes squared error and $S^{sim}_{t,n,v,a}$ the simulated electricity price at time t for the k simulation path for asset a and volatility model v. $S^{rel}_{t,a}$ denotes the realized electricity price at time t.

Table 9 shows the mean, standard deviation, minimum, and maximum of the root mean squared error resulting from simulating the peakload, off-peak 1 and off-peak 2 price processes, using the three different volatility models. Figure 7 shows the cumulative distribution of the root mean squared error for all three volatility models for the peakload price process. As expected by looking at figure 6, figure 7 reveals that the stochastic volatility performs better than constant volatility. It is apparent that the EGARCH model is better able to capture the behavior of the spot price. A possible explanation is that the EGARCH allows asymmetry in the effects on the volatility.

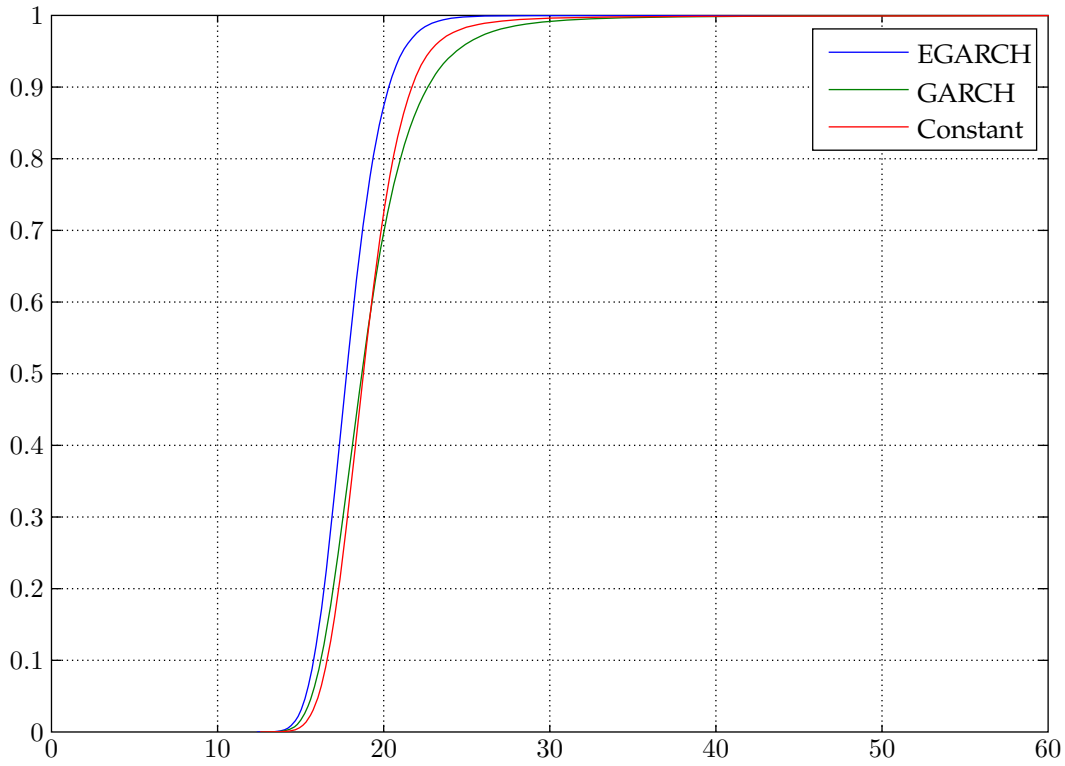

Figure 7: RMSE using different volatility models (peakload process)

Volatility model	Constant	GARCH	EGARCH
Mean	19.0970	19.2160	17.9402
Std Dev	0.2083	0.1778	0.1034
Min	12.5539	12.6129	12.3645
Max	829.4487	372.3992	116.4280

Table 9: Root mean squared error (peakload process)

5 Conclusion

In this paper, we propose a self-contained model for electricity spot prices that is able to capture (i) mean-reversion (with different speeds of mean-reversion for "normal" and "extreme" periods), (ii) spikes, (iii) stochastic volatility, and (iv) correlations among the spike processes in different time series. In this instance, we extend former models in three major aspects: First, we provide, based on Benth et al. (2003), a framework that allows us to incorporate a stochastic volatility in the diffusion process. Second, we use a self-contained estimation procedure for all necessary distribution and process parameters based only on the realized historic spot prices. Third, the model is able to consider correlations in spikes of

different time series, e.g. peakload and off-peak electricity prices. For the calibration and the testing of the performance of the model, we use electricity spot price data from the German electricity market, i.e. the EEX.

Regarding the model with stochastic volatility, we conclude that, although it has a higher complexity resulting in a more difficult extraction of parameters from empirical data, it is worth the effort. Especially in the case of the peakload price process the EGARCH model, which is the most complex model in our analysis, outperforms both the GARCH volatility and the constant volatility model. Furthermore, our analysis reveals that modeling the correlation among the spike process is a crucial step to capture the specific nature of the electricity spot price. In our empirical validation, we show that the price paths simulated with this model have the same characteristics as realized electricity spot price data.

However, future research is necessary to further increase our understanding of electricity spot price behavior. For example, models with an own stochastic process for the volatility might be promising. To sum up, we proposed a mode for the electricity spot price that explicitly considers all distinct features of this commodity. Furthermore, it can easily be estimated only based on the time series, a fact that makes it interesting for practical applications, e.g. option pricing.

References

Benth, F. E., Ekeland, L., Hauge, R., Nielsen, B., 2003. A note on arbitrage-free pricing of forward contracts in energy markets. Applied Mathematical Finance 10 (4), 325–336.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81 (3), 637–654.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Botterud, A., Kristiansen, T., Ilic, M. D., 2010. The relationship between spot and futures prices in the nord pool electricity market. Energy Economics 32(5), 967–978.

Bowden, N., Payne, J. E., 2008. Short term forecasting of electricity prices for miso hubs: Evidence from arima-egarch models. Energy Economics 30, 3186–3197.

Cartea, Á., Figueroa, M. G., 2005. Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. Applied Mathematical Finance 22 (4), 313–335.

Chan, K. F., Gray, P., 2006. Using extreme valaue theory to measure value-at-risk for daily electricity spot prices. International Journal of Forecasting 22, 283–300.

Deng, S., 2001. Stochastic Models of Energy Commodity Prices and Their Applications: Mean-reversion with Jumps and Spikes. Journal of Regulatory Economics 19 (3), 239–270.

Escribano, Á., Pena, J. I., Villaplana, P., 2002. Modeling electricity prices: International evidence. Working Paper Economic Series 08 02-27, 1–32.

Furio, D., Meneu, V., 2010. Expectations and forward risk premium in the Spanish deregulated power market. Energy Policy 38, 784–793.

Geman, H., Roncoroni, A., 2006. Understanding the structure of electricity prices. Journal of Business 79 (3), 1225–1261.

Hambly, B., Howison, S., Kluge, T., 2009. Modelling spikes and pricing swing options in electricity markets. Quantitative Finance 9 (8), 937–949.

Higgs, H., Worthington, A. C., 2008. Modelling Spot Prices in Deregulated Wholesale Electricity Markets: A Selected Empirical Review. SSRN eLibrary.

Kluge, T., 2006. Pricing swing options and other electricity derivatives. Ph.D. thesis, University of Oxford.

Knight, J., Satchell, S. E., 2001. Return Distributions in Finance, first edition Edition. Butterworth-Heinemannl.

Lucia, J. J., Schwarz., E. S., 2002. Electricity prices and power derivatives: Evidence from the nordic power exchange. Review of Derivatives Research 5 (1), 5–50.

Merton, R. C., 1973. Theory of rational option pricing. Bell Journal of Economics and Management Science 4 (1), 141–183.

Merton, R. C., 1976. Option pricing when underlying stock returns are discontinous. Journal of Financial Economics 3 (1), 125–144.

Merton, R. C., 2001. Continous-Time Finance. Blackwell.

Nelson, D. B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347–370.

Pindyck, R. S., Rubinfeld, D. L., 1998. Econometric Models and Economic Forecasts, fourth edition Edition. McGraw-Hill.

Redl, C., Haas, R., Huber, C., Böhm, B., 2009. Price formation in electricity forward markets and the relevance of systematic forecast errors. Energy Economics 31, 356–364.

Samuelson, P. A., 1965. Proof that properly anticipated prices fluctuate randomly. Industrial Management Review 6 (1), 41–49.

Schwarz, E. S., 1997. The stochastic behaviour of commodity prices: Implications for valuation and hedging. Journal of Finance 52 (52), 923–973.

Shrevel, S. E., 2004. Stochastic calculus for Finance II, first edition Edition. Spinger Finance.

Weron, R., 2006. Modeling and forecasting electricity loads and prices - A Statistical Approach. Wiley Finance.

Wilkens, S., Wimschulte, J., 2007. The pricing of electricity futures: Evidence from the European Energy Exchange. Journal of Futures Markets 27(4), 387–410.

A Off-peak1 and off-peak2 processes results

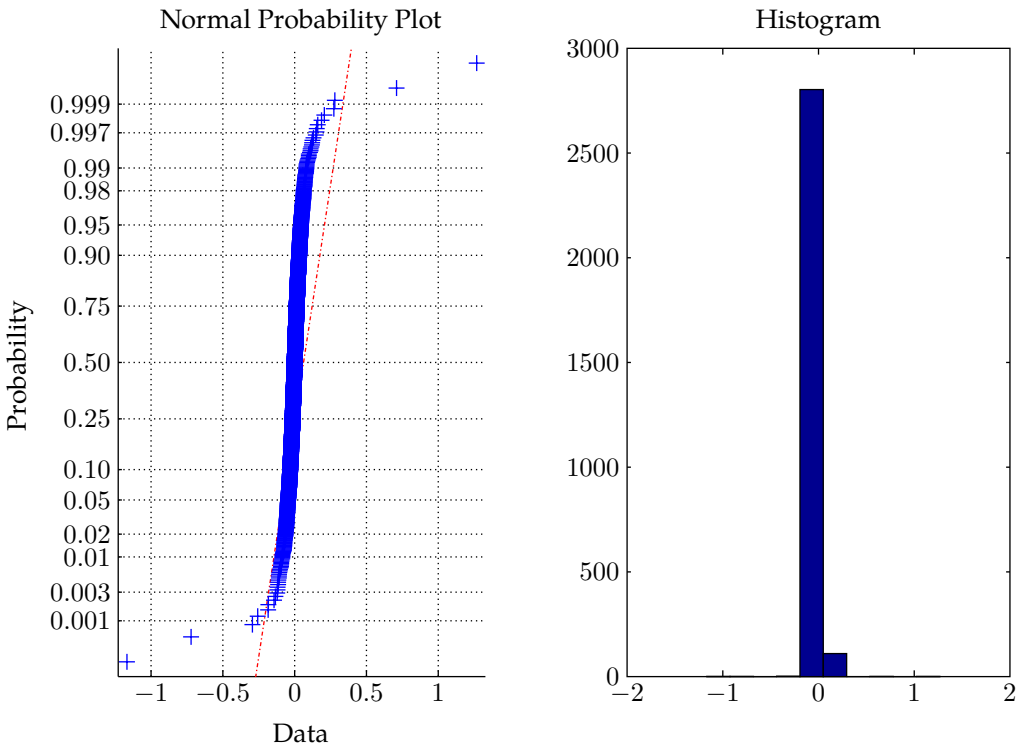

Figure 8: Normal probability plot and histogram of the unseperated off-peak 1 process

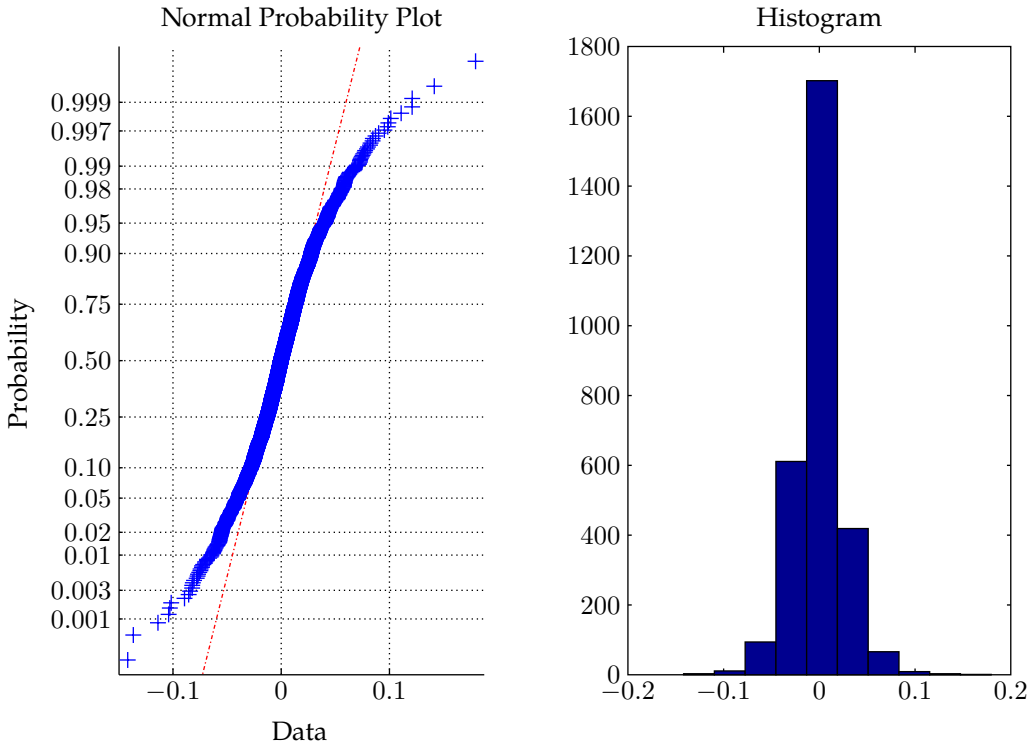

Figure 9: Normal probability plot and histogram of the unseperated off-peak 2 process

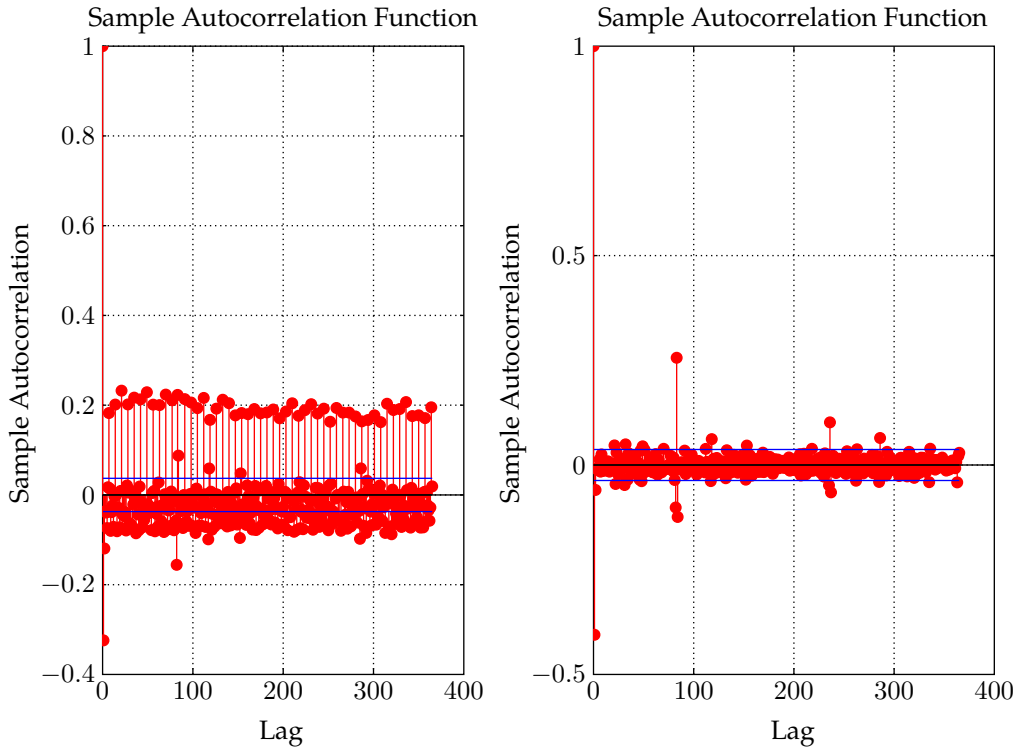

Figure 10: ACF plots of the log price and the log price without seasonality (off-peak 1 process)

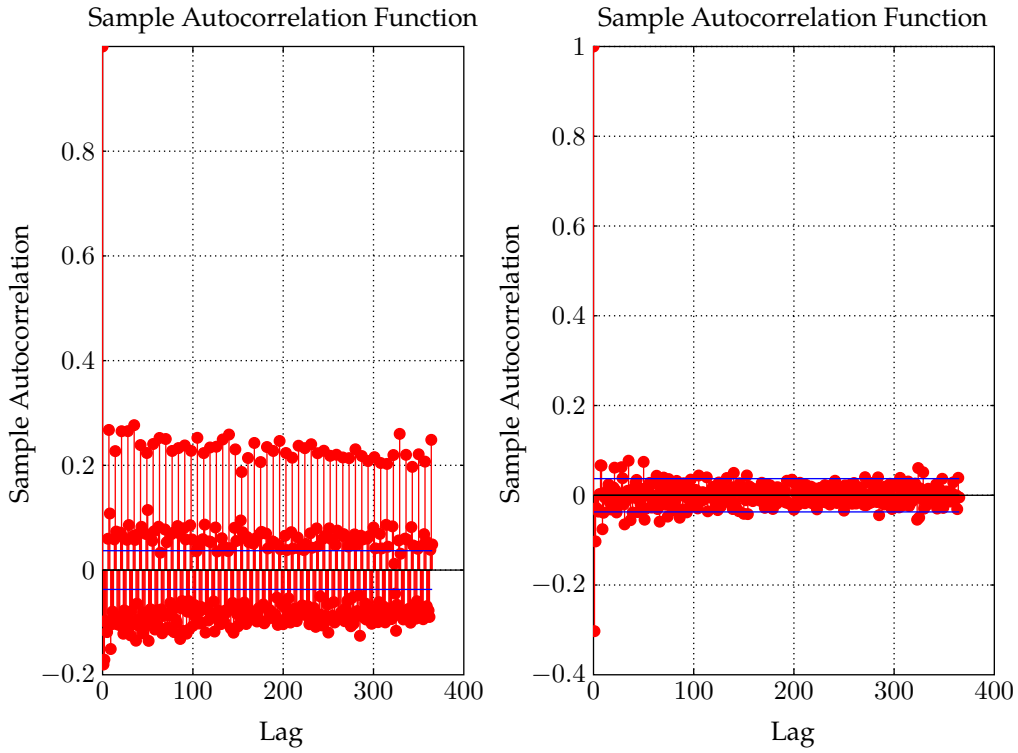

Figure 11: ACF plots of the log price and the log price without seasonality (off-peak2 process)

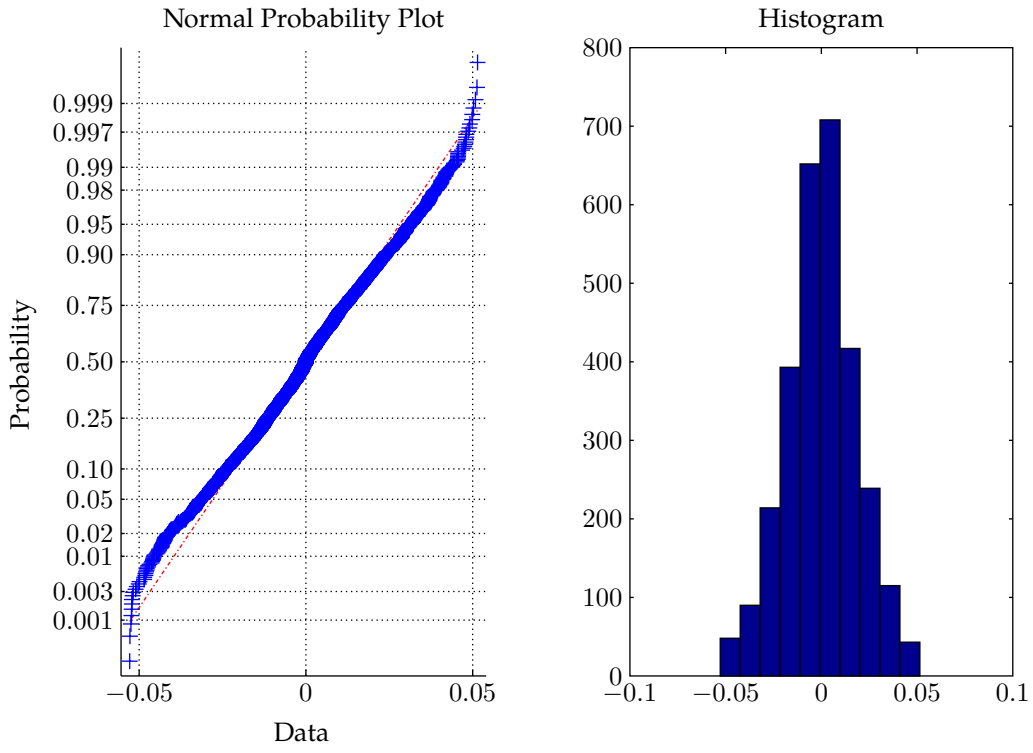

Figure 12: Normal probability plot and histogram of the diffusion part (off-peak1 process)

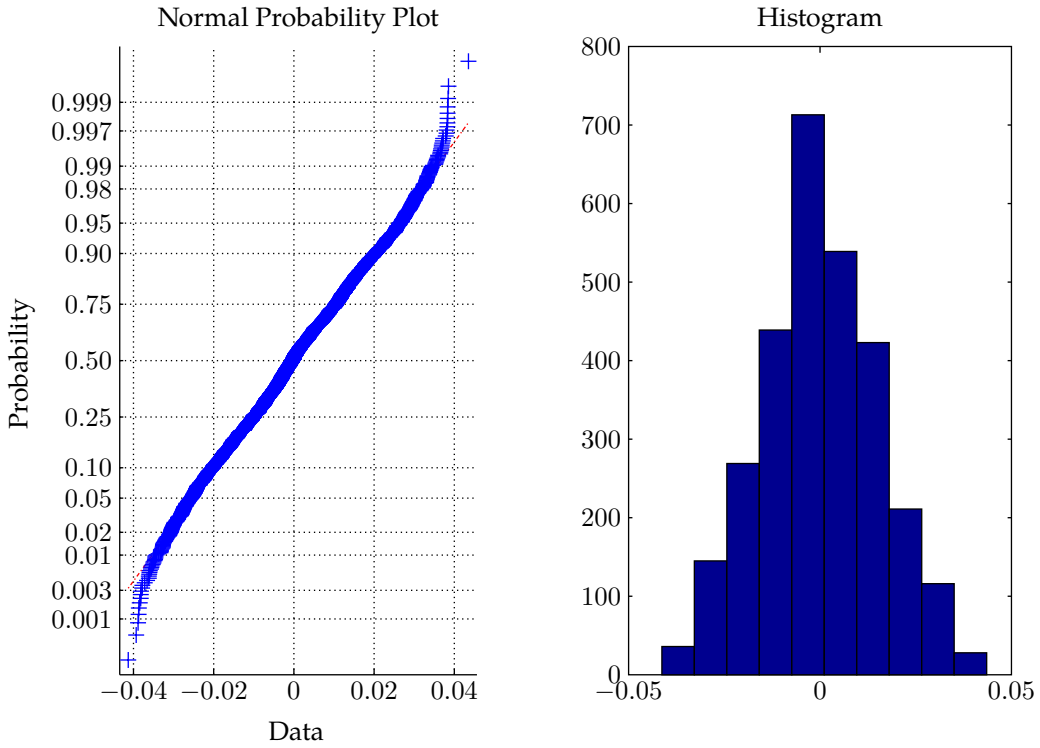

Figure 13: Normal probability plot and histogram of the diffusion part (off-peak2 process)

	Seasonality and drift parameter						
	Off-peak 1 price process						
Weekly seasonality	−1.2857	2.7180	3.1261	3.6671	2.9411	−1.3487	−9.8179
Drift	$0.0077t + 215.6862$						
Yearly seasonality	$2.2565 \cdot \cos\left(\frac{2\pi(t-357.0434)}{365}\right)$						
	Off-peak 2 price process						
Weekly seasonality	1.7547	2.5923	2.7773	2.1662	−0.0345	−4.8357	−4.4202
Drift	$0.0108t + 222.0488$						
Yearly seasonality	$-2.3967 \cdot \cos\left(\frac{2\pi(t-144.5787)}{365}\right)$						

Table 10: Seasonality and drift parameters (off-peak1 and off-peak2 processes)

Volatility Model	Constant	GARCH	EGARCH
	Off-peak 1 price process		
α_X	0.0816	0.0969	0.0994
α_Y	0.7732	0.7991	0.7998
	Off-peak 2 price process		
α_X	0.0524	0.0556	0.0553
α_Y	0.5197	0.5614	0.5454

Table 11: Mean-reversion parameters (off-peak1 and off-peak2 processes)

Volatility Model	Constant	GARCH	EGARCH
	Off-peak 1 price process		
	Distribution Parameter		
μ	0	0	0
σ	0.0167	0.0187	0.0184
	GARCH Parameter		
κ	1	0.0054	0.00001
ARCH	0	0.0460	0.0735
GARCH	0	0.9489	0.9942
Leverage	0	0	0.0001
	Off-peak 2 price process		
	Distribution Parameter		
μ	0	0	0
σ	0.0152	0.0154	0.0153
	GARCH Parameter		
κ	1	0.0108	-0.0001
ARCH	0	0.0482	0.0575
GARCH	0	0.9411	0.9922
Leverage		0	0.0177

Table 12: Volatility parameters (off-peak1 and off-peak2 processes)

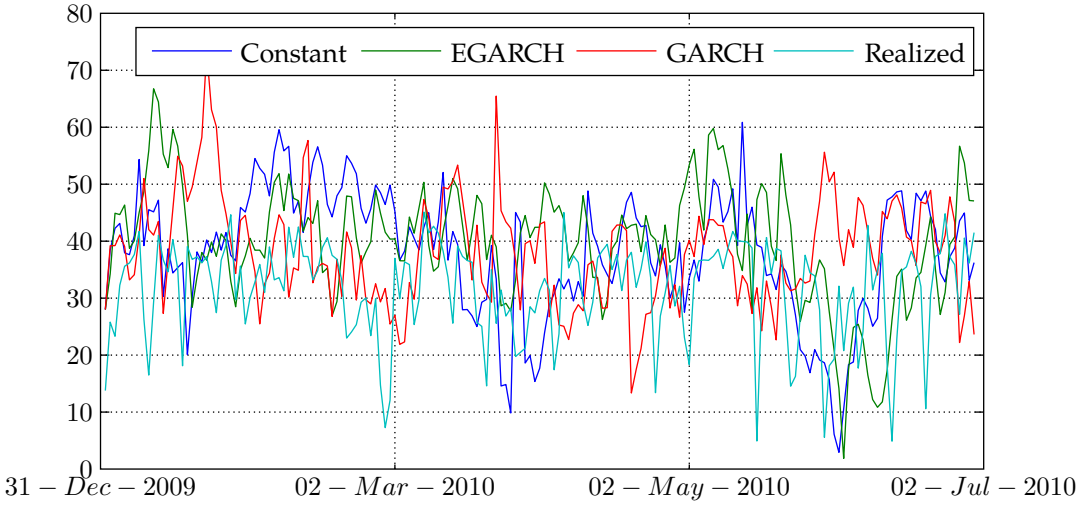

Figure 14: Spot and Simulated Spot Prices (off-peak1 process)

Volatility Model	Constant	GARCH	EGARCH
	Off-peak 1 price process		
μ^+	-3.0579	-2.9780	-3.0349
σ^+	0.7049	0.6462	0.8040
λ^+	0.0552	0.0401	0.0397
μ^-	-3.0361	-3.0280	-2.9377
σ^-	0.8183	0.7786	0.7300
λ^-	0.0743	0.0634	0.0565
# Jumps	378 (161 | 217)	302 (117 | 185)	281 (116 | 165)
	Off-peak2 price process		
μ^+	-3.1348	-3.1458	-3.1276
σ^+	0.5332	0.5107	0.4966
λ^+	0.0603	0.0603	0.0610
μ^-	-3.0916	-3.1425	-3.1482
σ^-	0.3603	0.3990	0.4126
λ^-	0.0562	0.0596	0.0603
# Jumps	340 (176 | 164)	350 (176 | 174)	354 (178 | 176)

Table 13: Parameters of the jump distributions (off-peak1 and off-peak2 processes)

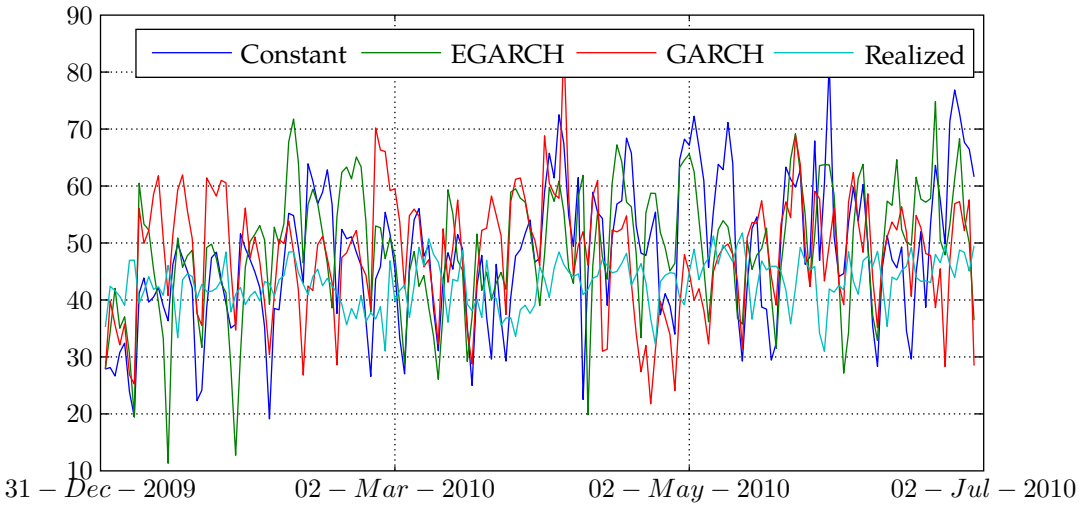

Figure 15: Spot and Simulated Spot Prices (off-peak2 process)

Volatility model	Constant	GARCH	EGARCH
	Off-peak 1 price process		
Mean	30.7696	30.8413	30.7081
Std Dev	0.1590	0.1574	0.1538
Min	14.1359	14.3356	15.0585
Max	57.5053	71.7823	60.1374
	Off-peak 2 price process		
Mean	28.9948	28.8258	28.5571
Std Dev	0.2263	0.1696	0.1693
Min	13.4463	12.6675	13.7845
Max	807.9479	207.5292	241.4611

Table 14: Root mean squared error (off-peak1 and off-peak2 processes)