COLOGNE GRADUATE SCHOOL
in Management, Economics and Social Sciences

Vol 3 | No 14

CGS WORKING PAPER

Forecasting extreme electricity spot prices

Volodymyr Korniichuk (Cologne Graduate School, University of Cologne)

Cologne Graduate School
in Management, Economics
and Social Sciences
Albertus-Magnus-Platz
50923 Köln
www.cgs.uni-koeln.de

University of Cologne

Forecasting extreme electricity spot prices

Volodymyr Korniichuk[a]

[a] *Cologne Graduate School, University of Cologne*

Abstract

We propose a model for forecasting extreme electricity prices in real time (high frequency) settings. The unique feature of our model is its ability to forecast electricity price exceedances over very high thresholds, where only a few (if any) observations are available. The model can also be applied for simulating times of occurrence and magnitudes of the extreme prices. We employ a copula with a changing dependence parameter for capturing serial dependence in the extreme prices and the censored GPD for modelling their marginal distributions. For modelling times of the extreme price occurrences we propose an approach based on a negative binomial distribution. The model is applied to electricity spot prices from Australia's national electricity market.

Keywords: electricity spot prices, copula, GPD, negative binomial distribution.

1. Introduction

A typical feature of electricity spot prices is their disposition towards sudden extreme jumps. This phenomenon stems from the lack of practical ways to store electricity and is attributed to an inelastic demand for electricity and very high marginal production costs in case of unforeseen shortfalls in electricity supply or unexpected rises in the demand. Although lasting for rather short time intervals, the magnitudes of those jumps may take extreme proportions hundred times exceeding the average electricity prices. This type of price behaviour presents an important topic for the risk management research and is of great relevance for electricity market participants, for example, retailers, who buy electricity at market prices but redistribute it at fixed prices to consumers. Estimating the probabilities of electricity prices to exceed some high thresholds is of paramount importance for the retailers, because even a few hours of extreme prices on the market may cause significant losses in their portfolios.

The problem of modelling extreme electricity prices was considered in many papers, e.g., Eichler, Grothe, Manner, and Tuerk (2012), Christensen, Hurn, and Lindsay (2012), which concentrate on modelling times of extreme price occurrences in Australia's national electricity market, and Klüppelberg, Meyer-Brandis, and Schmidt (2010) for the electricity spot price model applied to daily data of the EEX Phelix Base electricity price index. Considering the recent developments in modelling extreme electricity prices, there is still a lack of an approach for a combined modelling of times of occurrence and magnitudes of extreme electricity prices in real time (high frequency) settings. To fill that gap, we develop in this paper a model for a complete description of extreme electricity spot prices. The model consists of two components (sub-models): one for modelling the magnitudes of extreme electricity prices and the other for modelling times of extreme electricity price occurrences. Once being estimated, our model can be applied (without re-estimation) for forecasting the price exceedances over any sufficiently high threshold. This unique feature is provided by a special construction of the model in which price exceedances over a comparatively small threshold may trigger the exceedances over much larger levels.

Common distributions used in the literature for modelling electricity prices are Gaussian, exponential, and generalized beta (Geman and Roncoroni (2010), Becker, Hurn, and Pavlov (2007)). Since those distributions cannot account for heavy tails of the magnitudes of extreme electricity

Email address: `korniichuk@wiso.uni-koeln.de` (Volodymyr Korniichuk)

December 27, 2012

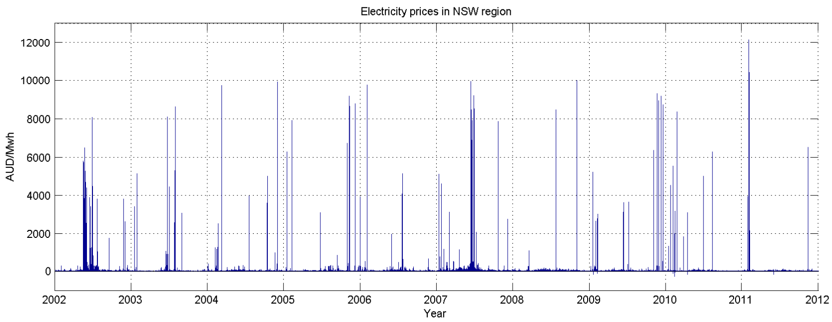

Figure 1: Electricity prices in NSW region of Australia's electricity market over the period Jan 1, 2002–Dec 31, 2011.

spot prices, we suggest, first, to use a generalized Pareto distribution (GPD) for capturing the heavy tails and, second, to employ a copula (survival Clayton) with a changing dependence parameter for capturing the serial dependence between the magnitudes. We account also for possible ceilings in the electricity prices by applying the censored GPD approach.

For modelling times of extreme electricity price occurrences, we propose a duration model based on a negative binomial distribution with a time-varying parameter. That model can capture the main features of time intervals between the extreme price occurrences, namely, the high variability, the strong persistence, and the discreteness. We compare the performance of the proposed model to the performance of other suitable approaches, like the autoregressive conditional duration model (Engle and Russell, 1998) and the Hawkes process (Hawkes, 1971).

The model of this paper is developed on and applied to the dataset of half-hourly electricity spot prices from the four regions of Australia's electricity market: New South Wales (NSW), Queensland (QLD), South Australia (SA), and Victoria (VIC). The dataset consists of 175296 observations, embracing the period over January 1, 2002–December 31, 2011.

The rest of the paper is organized as follows. In Section 2 we define a price spike, a building block of this paper, and provide then a short data-analysis of the prices. In Sections 3 and 4 we present our models for, respectively, magnitudes and times of the spikes. Section 5 combines those models into one model for forecasting extreme electricity prices. Section 6 concludes.

2. Defining a price spike

In intra-day electricity spot prices, one frequently observes a feature which is common for most electricity markets, namely sudden extreme prices. In Australia's electricity market, for example, the magnitude of some prices 300 times exceeds the sample average, see Figure 1 for electricity prices from NSW region and Table 1 for descriptive statistics of the half-hourly prices from the four regions of Australia's electricity market embracing the period over January 1, 2002–December 31, 2011. Modelling and forecasting those extreme electricity prices is the aim of this paper.

A building block of our model is a (price) spike, under which we understand a situation when the electricity price exceeds a certain high threshold. We use the spikes to develop two separate models: one (in Section 3) for the magnitudes of the spikes and the other (in Section 4) for the times of spike occurrences. In Section 5, we combine those two models into one for a complete description of extreme electricity prices. The final model can provide probabilities of the prices to exceed not only the threshold of the spikes, but any other sufficiently high level. All those models are developed on the dataset from Australia's electricity market.

Treating a spike as a situation when the price exceeds a certain high threshold, there are different approaches in the literature how to set that threshold. In Christensen, Hurn, and Lindsay (2009), Eichler, Grothe, Manner, and Tuerk (2012), the choice of the threshold is argued either by needs of the market, e.g., 300AUD/MWh is the strike price of heavily-traded cap products in Australia's electricity market, or simply by convenience, e.g, setting the threshold at the 95% quantile of the prices. In both those cases the threshold is fixed and hence does not incorporate the electricity

Table 1: Descriptive statistics for half-hourly electricity spot prices (AUD/MWh) from the four regions of Australia's electricity market in the period over January 1, 2002–December 31, 2011.

	NSW	QLD	SA	VIC
mean	39.8	36.1	43.8	35.1
median	25.1	22.7	28.1	25.2
st. dev.	224.3	189.8	283.7	158.3
skewness	31.6	31.9	31.4	44.7
kurtosis	1138.5	1191.3	1065.1	2349.7
number of observations	175296	175296	175296	175296

Notes: subscripts on the column headings indicate four regions of Australia's electricity market: New South Wales (NSW), Queensland (QLD), South Australia (SA), and Victoria (VIC).

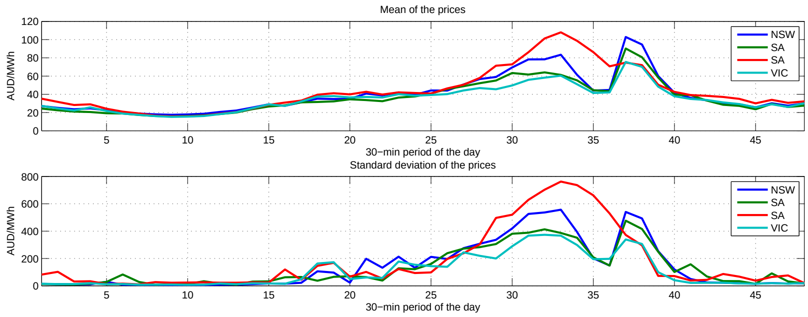

Figure 2: Mean and standard deviation of the electricity prices pooled by 30-min period of the day.

prices' diurnal structure, which is explicitly manifested in the prices' changing mean and variation during the day, see Figure 2.

This diurnal structure has a strong impact on the retailers' expectations of the prices. For instance, a retailer operating on Australia's electricity market at 6am may expect an average price for electricity of approximately 20AUD/MWh, but at 12am the retailer's expectation are completely different: the average price is doubled and the standard deviation is at least tripled. Due to these varying expectations, the price level of 155AUD/MWh at 12am can be regarded as extreme because it exceeds the 99% quantile of the expected prices at 12am, but at 6am the price should exceed the level of only 50AUD/MWh in order to be considered as extreme in the same sense. Those comparatively small extreme prices carry information about the state of the market (indicating, for example, a rise in the demand for electricity or a shortfall in the supply) and they should therefore be accounted for in forecasting electricity prices to exceed some higher (e.g. > 300AUD/MWh) thresholds.

Considering the diurnal distribution of the prices as a representation of the retailers' price expectations, we suggest to define a spike as a situation when the price exceeds a certain high quantile of those expectations. For this reason we set the diurnal threshold – the threshold which consists of 48 values corresponding to the 97% quantile of the prices happened at each of 48 half-hour periods of the day. The choice of the 97% quantile is motivated by the intention to consider high prices, on one hand, and to have enough data for statistical inferences, on the other hand. The spikes defined with help of the diurnal threshold will be used in Section 5 for construction of the model that can forecast the prices to exceed not only the diurnal threshold, but any other sufficiently high levels.

Figure 3 and 4 plot, respectively, the diurnal threshold and monthly proportions of the spikes provided by that diurnal threshold in the four regions of Australia's electricity market in the period over January 1, 2002–December 31, 2011. Note that an atypically high proportion of the spikes in the year 2007 is unrepresentative for the whole dataset and can severely distort the modelling of times of spike occurrences. We will address that issue in Section 4.

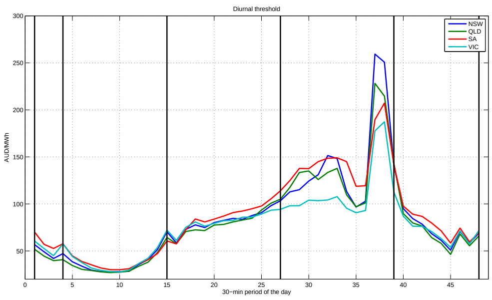

Figure 3: Diurnal threshold. Notes: solid vertical lines illustrate parts of the day where ξ of the GPD can be assumed to be the same, details in Section 3.1.1.

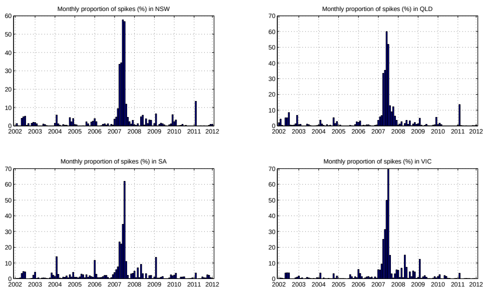

Figure 4: Monthly proportions of the spikes. Notes: the period of atypically high proportion of spikes in 2007 will be removed in modelling times of the spike occurrences.

3. Modelling magnitudes of the spikes

We understand a spike magnitude as the excess of the price level over the corresponding value of the diurnal threshold at times when spikes occur. Throughout the paper, $Y_1, Y_2, \ldots, Y_N$ will denote N consecutive (random) spike magnitudes. In Section 3.1 we develop a model for capturing the main features of the spike magnitudes. Section 3.2 considers a censored estimation procedure to account for the ceiling in the electricity prices. In Section 3.3 we report on the estimation results of fitting the model to magnitudes of the spikes occurred in the four regions of Australia's electricity market in the period over January 1, 2002–December 31, 2011.

3.1. Description of the model

3.1.1. Modelling long tails in magnitudes of the spikes

Magnitudes of extreme electricity prices are often modelled with Gaussian, exponential, or generalized beta distributions, see, for example, Geman and Roncoroni (2010), Becker, Hurn, and Pavlov (2007). Since those methods may significantly underestimate a spike risk in the high-frequency electricity spot prices (by failing to account for their heavy tails), we suggest to use a generalized Pareto distribution (GPD) for modelling magnitudes of the spikes. The distribution function of a GPD is defined as follows

$$G(x;\xi,\beta) = \begin{cases} 1-(1+\xi x/\beta)^{-1/\xi}, & \xi \neq 0, \\ 1-\exp(-x\beta), & \xi = 0, \end{cases} \tag{1}$$

where $\beta > 0$, $x \geq 0$ for $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ for $\xi < 0$. We shall refer to ξ and β as, respectively, the shape and scale parameters. Note that the GPD distributed random variable X is a heavy-tailed and it holds $\mathrm{E}\left(X^k\right) = \infty$ for $k \geq 1/\xi$.

The GPD possesses several properties which may be very beneficial in modelling magnitudes of the spikes. First, it was shown in Pickands (1975) that for distributions belonging to the domain of attraction of an extreme value distribution, i.e., for heavy-tailed data, the GPD is a limiting distribution for excesses over a suitably high threshold. Since the electricity prices are heavy-tailed, the GPD is a natural choice to model the magnitudes of the spikes. Second, the GPD is characterized by threshold stability property stating that if excesses over some threshold u_1 can be modelled by the GPD with the shape parameter ξ and the scale parameter β_{u_1}, then excesses over the higher threshold u_2 can be modelled by the GPD with the same shape parameter ξ and the scale parameter β_{u_2} defined as $\beta_{u_2} = \beta_{u_1} + \xi(u_2 - u_1)$. Using the GPD for the spike magnitudes may provide better estimates of the tail of the spikes and protect against arbitrariness involved in the choice of the diurnal threshold. See Davison and Smith (1990) for a detailed record on using a GPD to model exceedances over high thresholds. Modelling extreme electricity prices with a GPD can also be found in Klüppelberg, Meyer-Brandis, and Schmidt (2010). Note that the use

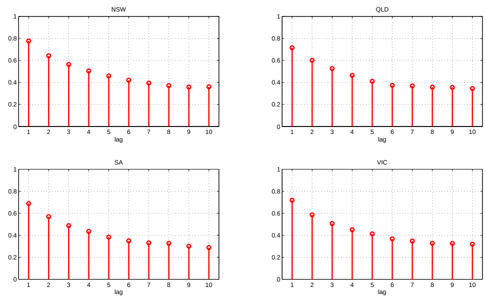

Figure 5: Spearman's rank correlation between the lagged spike magnitudes.

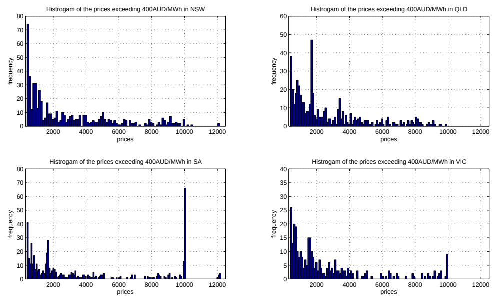

Figure 6: Histogram of the electricity prices exceeding 400AUD/MWh.

of Paretian distributions (a GPD's special case) to model commodity prices was first suggested in Mandelbrot (1963).

Applying the GPD for description of the spike magnitudes is not straightforward, because the diurnal structure of the prices implies that the spikes across the day have different characteristics and hence they cannot be modelled by the same GPD. We suggest to model the spike magnitudes belonging to each of 48 half-hour periods of the day by a separate GPD. In order to reduce the number of the shape parameters to be estimated, which, in turn, simplifies our model and leads to more accurate estimates, we distinguish parts of the day when the shape parameters of the prices can be assumed to be the same (but not the scale parameters). The possible division of the day can be as follows: 12am–2am; 2.30am–7.30am; 8am–13.30pm; 14pm–7pm; 7.30pm–11.30pm. Solid vertical lines on Figure 3 illustrate that division. Further in the text, we will denote by $m(i)$ a function that identifies to which part of the day (where the shape parameters are assumed equal) the i-th observation belongs, and by $n(i)$ a function that identifies to which out of 48 half-hour periods of the day the i-th observation belongs. The corresponding parameters of the GPD will be denoted by $\xi_{m(i)}$ and $\beta_{n(i)}$.

3.1.2. Modelling dependence in magnitudes of the spikes

In addition to the distributional choice for spike magnitudes $Y_1, Y_2, \ldots, Y_N$, there is a need of modelling a dependence between them. Figure 5 plots the estimated autocorrelation of the spike magnitudes calculated as Spearman's rank correlation between k-lagged samples $Y_1, Y_2, \ldots, Y_{N-k}$ and $Y_{k+1}, Y_{k+2}, \ldots, Y_N$. Although those correlations cannot be directly interpreted, because time intervals between the spike occurrences ranges from 30 minutes to 5 months, they still clearly indicate a strong positive dependence between the spike magnitudes. In addition to the strong dependence, extreme electricity prices display a peculiar clustering behaviour around the level of 10000AUD/MWh, see Figure 6, which is attributed to price ceiling on the market. We will address modelling this price ceiling in Section 3.2.

A conventional way of handling dependences between non-identically distributed random variables is a copula. In continuous case, a copula is a multivariate distribution function with uniformly on $[0, 1]$ distributed marginal distributions. A detailed introduction to copulas can be found in Nelsen (2006), for an excellent review of copula based models for econometric time series see Patton (2012). To capture dependence between two consecutive spikes Y_{i-1} and Y_i we suggest to model the conditional distribution of Y_i given $Y_{i-1} = y_{i-1}$ as a conditional distribution of two random variables with copula $C(u_1, u_2)$, namely,

$$\mathrm{P}\left(Y_i \leq y \mid \mathbb{H}_{t_{i-1}}, t_i\right) = \frac{\frac{\partial}{\partial u_2} C\left(F_{Y_i}(y), F_{Y_{i-1}}(y_{i-1})\right)}{f_{Y_{i-1}}(y_{i-1})}, \tag{2}$$

where $\mathbb{H}_{t_{i-1}}$ is a history of the first $(i-1)$ spikes including their magnitudes $(y_1, \ldots, y_{i-1})$ and times of occurrences $(t_1, \ldots, t_{i-1})$; $\frac{\partial}{\partial u_2} C$ denotes a derivative of copula $C(u_1, u_2)$ by the second component; F_{Y_i} is an unconditional distribution function of Y_i (which is assumed to be the GPD), and f_{Y_i} is the corresponding density. Note that $F_{Y_i}(Y_i)$ follows the uniform on $[0,1]$ distribution.

Considering the specification for $C(u_1,u_2)$, we prefer a dependence structure which is both simple, to provide explicit simulation formulas, and flexible, to capture a changing dependence between the spike magnitudes. We suggest to use the survival Clayton copula, which is defined as follows

$$C(u_1,u_2) = u_1 + u_2 + \left((1-u_1)^{-\theta} + (1-u_2)^{-\theta} - 1\right)^{-1/\theta} - 1, \qquad 0 < \theta < \infty. \tag{3}$$

In the limit this copula approaches the independence copula as $\theta \to 0$ and the two-dimensional comonotonicity copula as $\theta \to \infty$. Beneficially to modelling clustering of large spikes, copula (3) implies asymptotically dependent tails with a coefficient of upper tail dependence $\lambda_u = 2^{-1/\theta}$. Furthermore, the survival Clayton copula is the upper threshold limit of the Galambos copula providing in practice an accurate approximation for Gumbel and t extreme value copulas, see McNeil, Frey, and Embrechts (2005).

Applying copula for modelling dependence between two consecutive spikes Y_{i-1} and Y_i, it is reasonable to assume that the more time has elapsed between the spikes the less dependent they are. To capture that idea, we suggest to model the dependence parameter θ of copula (3) as $\theta_i = \gamma_0\ D_i^{-\gamma_1}$, $\gamma_0 > 0$, $\gamma_1 \geq 0$, where D_i denotes a time interval between occurrence of two consecutive spikes Y_{i-1} and Y_i, i.e, $D_i = t_i - t_{i-1}$. This specification of θ_i implies a constant (not time-varying) level of dependence between the spikes that are separated by the same time interval.

With copula (3) and the GPD as an unconditional distribution of the spike magnitudes, the conditional distribution of Y_i in (2) takes the form

$$\mathrm{P}\left(Y_i \leq y \mid \mathbb{H}_{t_{i-1}}, t_i\right) = 1 - \left(1 + \frac{g_i(y)^{\theta_i/\xi_{m(i)}} - 1}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}\right)^{-(1/\theta_i+1)}, \tag{4}$$

with $\mathrm{P}\left(Y_1 \leq y \mid \mathbb{H}_0, t_1\right) = 1 - g_1(y)^{-1/\xi_{m(1)}}$, where again $m(i)$ and $n(i)$ denotes a function that identifies, respectively, to which part of the day where the shape parameters are assumed equal and to which out of 48 half-hour periods of the day the i-th observation belongs; $\xi_{m(i)}$ and $\beta_{n(i)}$ denote parameters of the GPD used for modelling magnitudes of the i-th spike; $g_i(y) = 1 + \xi_{m(i)} \frac{y}{\beta_{n(i)}}$. Note that when the time interval between occurrence of consecutive spikes is large, θ_i approaches zero vanishing any dependence between spikes Y_{i-1} and Y_i yielding the conditional distribution function of Y_i as

$$\mathrm{P}\left(Y_i \leq y \mid \mathbb{H}_{t_{i-1}}, t_i\right) = 1 - g_i(y)^{-1/\xi_{m(i)}},$$

which is the distribution function of the GPD.

3.1.3. Estimation

For fully parametric copula-based models the most efficient estimation method is maximum likelihood. Assuming that the conditional distributions of the spikes are independent, the likelihood of N realization $y_1, y_2, \ldots, y_N$ of spike magnitudes from model (4) takes the form

$$\mathcal{L} = \prod_{i=1}^{N} \left. \frac{\partial \mathrm{P}\left(Y_i \leq y \mid \mathbb{H}_{t_{i-1}}, t_i\right)}{\partial y} \right|_{y=y_i}, \tag{5}$$

where

$$\left. \frac{\partial \mathrm{P}\left(Y_i \leq y \mid \mathbb{H}_{t_{i-1}}, t_i\right)}{\partial y} \right|_{y=y_i} = \frac{(\theta_i+1)}{\beta_{n(i)}} \left(1 + \frac{g_i(y_i)^{\theta_i/\xi_{m(i)}} - 1}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}\right)^{-(2+1/\theta_i)} \frac{g_i(y_i)^{\theta_i/\xi_{m(i)}-1}}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}, \tag{6}$$

with $\left. \frac{\partial \mathrm{P}(Y_1 \leq y \mid \mathbb{H}_0, t_1)}{\partial y} \right|_{y=y_1} = \frac{1}{\beta_{n(1)}} g_1(y_1)^{-1/\xi_{m(1)}-1}$.

Computing standard errors of the estimated parameters we will consider standard errors based on an inverse Hessian of the likelihood function and simulation-based standard errors, which are computed on parameter estimates of the model fitted on many simulated samples from the originally estimated model. Provided the correct model specification, the latter approach yields correct finite-sample standard errors when a number of simulation is sufficient, see Patton (2012).

3.1.4. Simulation and Goodness-of-fit

Applying the above model in practice, where true distributions are unknown, it is vital to conduct a goodness-of-fit test and a simulation study to check the fit of the estimated model. Our approach to the goodness-of-fit test is based on the probability integral transformation (Diebold, Gunther, and Tay, 1998) of the sample of spike magnitudes $y_1, y_2, \dots, y_N$ with the estimated conditional density forecast (4). Recalling that for a continuous random variable X with distribution function F, $F(X)$ is uniformly distributed on the unit interval, we obtain from (4) that residuals defined as

$$u_i = 1 - \left(1 + \frac{g_i(y_i)^{\theta_i/\xi_{m(i)}} - 1}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}\right)^{-(1/\theta_i+1)}, \qquad i = 2, \dots, N, \tag{7}$$

with $u_1 = 1 - g_1(y_1)^{-1/\xi_{m(1)}}$, are supposed to be N independent realizations from the uniform on $[0,1]$ distribution, if the estimated model is correct (suitable). Since the goodness-of-fit looks for evidence that the model is misspecified, the test of the estimated model can be limited to checking a hypothesis that the residuals are independent realizations from the standard uniform distribution.

From (7), it is also immediate to construct the simulation procedure. It follows that

$$\left(1 + \frac{g_i(Y_i)^{\theta_i/\xi_{m(i)}} - 1}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}\right)^{-(1/\theta_i+1)} \sim \text{ Unif[0,1]}, \qquad i = 2, \dots, N, \tag{8}$$

with $g_1(Y_1)^{-1/\xi_{m(1)}} \sim$ Unif[0,1], where $Y_1, Y_2, \dots, Y_N$ denote N consecutive (random) spike magnitudes. Provided the knowledge of time intervals between the spikes (to calculate θ_i), one can obtain a simulated spike magnitude by expressing Y_i from the above equation for each realization of Unif[0,1]. By adding to the simulated magnitude the corresponding value of the diurnal threshold one obtains a simulated (extreme) electricity price.

3.2. Accounting for the price ceiling in magnitudes of the spikes

Due to legal regulations of Australia's electricity market, the prices are capped at a maximum of 12500AUD/MWh. This ceiling was 5000AUD/MWh until April 1, 2002 and 10000AUD/MWh in the period from April 1, 2002 to July 1, 2010. The number of prices in the sample which have approximately reached the ceilings, we call those prices censored, are the following: 3 in NSW, 0 in QLD, 66 in SA, and 9 in VIC. Although there are only a few censored prices, they still may have a strong effect on estimating tails of the spike magnitudes.

Constructing likelihood function $\mathcal{L}^C$ which accounts for the ceiling in the electricity prices, we distinguish four types of contribution $\mathcal{L}_i^C(y_i)$ of observation $Y_i = y_i$ to that likelihood function. In case spike $Y_{i-1} = y_{i-1}$ is censored and $Y_i = y_i$ is not censored, the contribution of $Y_i = y_i$ to the likelihood is as follows

$$\mathcal{L}_i^C(y_i) = \left.\frac{\partial \mathrm{P}(Y_i \leq y \mid \mathbb{H}_{t_{i-1}}, t_i, Y_{i-1} \geq y_{i-1})}{\partial y}\right|_{y=y_i} =$$
$$\frac{1}{\beta_{n(i)}}\left(1 + \frac{g_i(y_i)^{\theta_i/\xi_{m(i)}} - 1}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}\right)^{-(1+1/\theta_i)} \frac{g_i(y_i)^{\theta_i/\xi_{m(i)}-1}}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}. \tag{9}$$

As $Y_{i-1} = y_{i-1}$ is not censored and $Y_i = y_i$ is censored then

$$\mathcal{L}_i^C(y_i) = \mathrm{P}(Y_i \geq y_i \mid \mathbb{H}_{t_{i-1}}, t_i, Y_{i-1} = y_{i-1}) = \left(1 + \frac{g_i(y_i)^{\theta_i/\xi_{m(i)}} - 1}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}\right)^{-(1+1/\theta_i)}. \tag{10}$$

If both $Y_{i-1} = y_{i-1}$ and $Y_i = y_i$ are censored then

$$\mathcal{L}_i^C(y_i) = \mathrm{P}(Y_i \geq y_i \mid \mathbb{H}_{t_{i-1}}, t_i, Y_{i-1} \geq y_{i-1}) = \left(1 + \frac{g_i(y_i)^{\theta_i/\xi_{m(i)}} - 1}{g_{i-1}(y_{i-1})^{\theta_i/\xi_{m(i-1)}}}\right)^{-1/\theta_i} \frac{1}{g_{i-1}(y_{i-1})^{1/\xi_{m(i-1)}}}. \tag{11}$$

In the case of $Y_{i-1} = y_{i-1}$ and $Y_i = y_i$ being not censored, the contribution of $Y_i = y_i$ is as in (6).

The final likelihood function $\mathcal{L}^C$ is a product of contributions $\mathcal{L}_i^C(y_i)$ across all spikes, where $\mathcal{L}_i^C(y_i)$ takes one form of (6),(9)-(11).

3.3. Estimation results

In this section we report on the estimation results of fitting the model of Section 3.1 to the magnitudes of the spikes occurred in the four regions of Australia's electricity market over the period January 1, 2002–December 31, 2011. Note that the model estimation is not adversely affected by the atypically high proportion of spikes in the year 2007, see Figure (4), because the conditional distribution (4) of the spike magnitudes depends only on the information of the previous spike and time of the current spike occurrence, not on the whole history of the spikes.

Estimating the model, we define spikes that occurred by April 1, 2002 as censored if the associated prices are higher than 4995AUD. Spikes happened in the period over April 1, 2002–July 1, 2010 are set as censored if the associated prices exceed the level of 9995AUD, and the spikes occurred after July 1, 2010 are considered to be censored if the associated prices reach 12495AUD. Note that, since the prices tend to cluster up to the ceiling, we allow for 5AUD deviation in identifying the censored data. The numbers of observations in the sample which were set as censored are the following: 3 in NSW, 0 in QLD, 66 in SA, and 9 in VIC. The total number of the spikes in the samples is as follows: 5241 in NSW, 5279 in QLD, 5271 in SA, and 5278 in VIC.

Table 2: Parameter estimates of the model for spike magnitudes. An inverse Hessian of the likelihood function is used to obtain the standard errors reported in parentheses right to the estimates.

	uncensored	censored	uncensored	censored
	NSW		QLD	
$\hat{\xi}_1$	0.4822 (0.2351)	0.4855 (0.0514)	0.5709 (0.0529)	0.5709 (0.0529)
$\hat{\xi}_2$	0.5125 (0.6881)	0.5161 (0.0351)	0.5298 (0.0326)	0.5298 (0.0326)
$\hat{\xi}_3$	1.1917 (0.2677)	1.1995 (0.0571)	1.2812 (0.0549)	1.2812 (0.0549)
$\hat{\xi}_4$	1.7956 (1.2866)	1.8161 (0.0877)	1.8200 (0.0765)	1.8200 (0.0765)
$\hat{\xi}_5$	0.8911 (0.1201)	0.8972 (0.0581)	1.2042 (0.0593)	1.2042 (0.0593)
$\hat{\gamma}_0$	2.8289 (2.0404)	2.8487 (0.1064)	2.7253 (0.0937)	2.7253 (0.0937)
$\hat{\gamma}_1$	0.3677 (0.8009)	0.3636 (0.0383)	1.4367 (0.1742)	1.4367 (0.1742)
	SA		VIC	
$\hat{\xi}_1$	1.0049 (0.0498)	1.1463 (0.0764)	0.4693 (0.0340)	0.4841 (0.0731)
$\hat{\xi}_2$	0.7680 (0.0357)	0.8913 (0.0386)	0.8732 (0.0530)	0.9076 (0.1445)
$\hat{\xi}_3$	1.1501 (0.0395)	1.3036 (0.0433)	1.2580 (0.0543)	1.3031 (0.2219)
$\hat{\xi}_4$	1.7370 (0.0482)	2.3066 (0.0770)	1.6176 (0.0689)	1.6668 (0.3598)
$\hat{\xi}_5$	1.1263 (0.0404)	1.2982 (0.0420)	1.6290 (0.0875)	1.7265 (0.1936)
$\hat{\xi}_6$	–	–	0.5561 (0.0392)	0.5735 (0.0505)
$\hat{\gamma}_0$	2.6284 (0.0511)	3.0280 (0.0605)	2.4701 (0.0562)	2.5457 (0.0552)
$\hat{\gamma}_1$	1.5777 (0.0641)	1.4117 (0.0825)	0.6722 (0.1216)	0.6400 (0.0847)

Notes: This table presents estimates of the shape (ξ) and dependence (γ_0, γ_1) parameters of the model for spike magnitudes. For estimation of ξ in NSW, QLD, and SA, five parts of the day were distinguished: 12am–2am; 2.30am–7.30am; 8am–13.30pm; 14pm–7pm; 7.30pm–11.30pm. For VIC region the following partition was used: 12am–8am; 8.30am–12pm; 12.30pm–14.30pm; 15pm–5.30pm; 6pm–8pm, 8.30pm–11.30pm.

Table 2 presents parameter estimates of model (4) obtained by maximizing the ceiling adjusted likelihood $\mathcal{L}^C$ (column "censored") and the unadjusted likelihood function $\mathcal{L}$ in (5) (column "uncensored"). For estimation of the model we distinguish five parts of the day: 12am–2am; 2.30am–7.30am; 8am–13.30pm; 14pm–7pm; 7.30pm–11.30pm, – and model magnitudes of the spikes within the part with the GPD which has the same shape parameter ξ but different scale parameters β corresponding to every half-hour period of the day. Note that improving fit of the model for spike magnitudes in VIC region, we use another partition of the day where the shape parameters are assumed equal. The partition is as follows: 12am–8am; 8.30am–12pm; 12.30pm–14.30pm; 15pm–5.30pm; 6pm–8pm, 8.30pm–11.30pm. In the table, we report only the estimates of the shape parameter (ξ) of the GPD and of the dependence parameters (γ_0, γ_1). To save space, estimates of the scale parameter are not displayed in the table, but they are available upon request.

Accounting for the price censoring has insignificant effects on the estimates of the shape parameter for NSW spikes, which have a few censored observations, but on the estimates for VIC and SA spikes that effect is strong leading to a significant upward adjustment of the uncensored estimates. Without that upward adjustment of the shape parameter estimates, the risk of extreme prices would be underestimated.

Provided by the asymptotic distributional properties of the maximum likelihood estimators, see, e.g., Greene (2003), calculation of standard errors for the parameter estimates in Table 2 is based on the inverse Hessian of the likelihood function. Since the use of asymptotic arguments in finite-size samples may yield inaccurate results, we conduct a further analysis of the estimators based on the ceiling adjusted maximum likelihood $\mathcal{L}^C$. We simulated 500 paths of (censored) spike magnitudes with the parameter values of Table 2 (column "censored") and estimated the ceiling adjusted model on every simulated path. The mean, the standard deviation, the mean relative bias, and the mean squared error of the estimated parameters are summarized in Table 3. Non-surprisingly, the estimators show a large variation and tend to be upward biased. This large variation of the estimates is a typical feature of all extreme value statistics, especially of those based on threshold data, see Klüppelberg, Meyer-Brandis, and Schmidt (2010).

Table 3: Estimated mean, standard deviation (std), mean relative bias (MRB), and mean squared error (MSE) of estimated parameters for the ceiling adjusted model from 500 simulated paths.

	true value	mean	std	MRB	MSE	true value	mean	std	MRB	MSE
			NSW					QLD		
$\hat{\xi}_1$	0.4855	0.64	0.09	0.31	0.03	0.5709	0.67	0.09	0.17	0.02
$\hat{\xi}_2$	0.5161	0.68	0.08	0.32	0.03	0.5298	0.62	0.07	0.17	0.01
$\hat{\xi}_3$	1.1995	1.58	0.19	0.32	0.18	1.2812	1.51	0.15	0.18	0.08
$\hat{\xi}_4$	1.8161	2.48	0.28	0.37	0.53	1.8200	2.25	0.23	0.24	0.24
$\hat{\xi}_5$	0.8972	1.18	0.14	0.32	0.10	1.2042	1.41	0.14	0.17	0.06
$\hat{\gamma}_0$	2.8487	3.79	0.42	0.33	1.06	2.7253	3.23	0.27	0.19	0.33
$\hat{\gamma}_1$	0.3636	0.34	0.03	-0.07	0.00	1.4367	1.36	0.15	-0.05	0.03
			SA					VIC		
$\hat{\xi}_1$	1.1463	1.36	0.13	0.19	0.06	0.4841	0.55	0.07	0.13	0.01
$\hat{\xi}_2$	0.8913	1.05	0.10	0.18	0.04	0.9076	1.03	0.12	0.13	0.03
$\hat{\xi}_3$	1.3036	1.56	0.14	0.19	0.08	1.3031	1.50	0.18	0.15	0.07
$\hat{\xi}_4$	2.3066	2.92	0.26	0.27	0.44	1.6668	1.95	0.22	0.17	0.13
$\hat{\xi}_5$	1.2982	1.55	0.15	0.20	0.09	1.7265	1.99	0.21	0.16	0.12
$\hat{\xi}_6$	–	–	–	–	–	0.5735	0.64	0.08	0.12	0.01
$\hat{\gamma}_0$	3.0280	3.64	0.28	0.20	0.46	2.5457	2.90	0.27	0.14	0.20
$\hat{\gamma}_1$	1.4117	1.35	0.13	-0.04	0.02	0.6400	0.62	0.05	-0.04	0.00

Notes: This table presents characteristics for estimates of the shape (ξ) and dependence (γ_0, γ_1) parameters of the ceiling adjusted model for the spike magnitudes estimated on 500 simulations from that model with parameter values of Table 2 column "censored".

As a goodness-of-fit test of the estimated ceiling adjusted model, Figure 7 plots the autocorrelation of the residuals $(\hat{u}_1, \ldots, \hat{u}_N)$ computed according to (7). The estimated autocorrelations

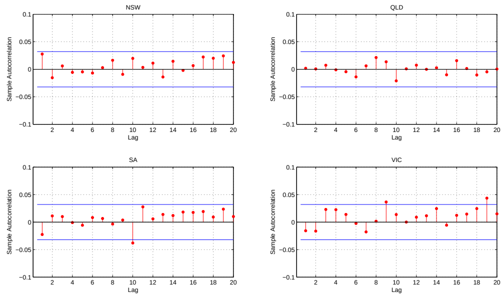

Figure 7: Autocorrelation of the residuals. Solid vertical lines show 99% confidence intervals.

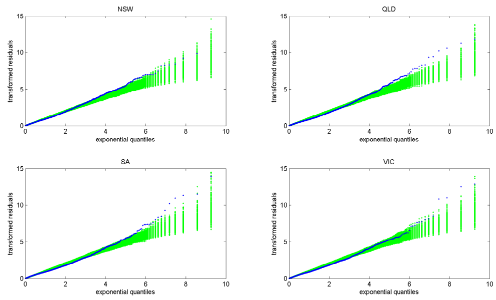

Figure 8: QQ-plot of the transformed residuals. Green points show expected deviations of the residuals.

lie mainly within the confidence bounds indicating no evidence against an assumption of zero autocorrelations. This suggestion is supported by the Ljung-Box test (15 lags) which failed to reject the null of no autocorrelation with p-values 61.82% for NSW, 91.44% for QLD, 16.58% for SA, and 7.39% for VIC. The hypothesis of no autocorrelation was also supported by investigating the squares of the residuals (p-values: 59.48% for NSW, 84.97% for QLD, 12.89% for SA, 28.06% for VIC). The absence of significant autocorrelation in the estimated residual indicates the ability of our model to capture the serial dependence between the spike magnitudes.

Considering the distributional properties of $(\hat{u}_1, \ldots, \hat{u}_N)$, Figure 8 illustrates the plot of the quantiles of transformed residuals $(-\log(\hat{u}_1), \ldots, -\log(\hat{u}_N))$ versus the corresponding quantiles of the standard exponential distribution. We have also added to the figure the QQ-plots of 100 realizations of the standard exponential random variable (in green color), to illustrate what type of deviations one can expect. In case of a good fit of the estimated model, the transformed residuals are supposed to be standard exponentially distributed, implying the uniform on $[0, 1]$ distribution of the estimated residuals. After an inspection of the plot, it becomes apparent that the transformed residuals seem indeed be consistent with the standard exponential distribution, although comparatively few of them deviate from the expected boundaries. Those few deviations may be attributed, among others, to an inevitable estimation error of the model (55 estimated parameters), the price ceiling, and a peculiar clustering behaviour of the prices (especially in QLD) around the level of 1700AUD/MWh, see Figure 6.

For a further analysis of the estimated ceiling adjusted model, we investigate its properties in a small simulation study. Using the original time intervals between the spike occurrences to compute θ_i, we simulated 500 samples of the spike magnitudes (of the same length as the original ones) and added to them the corresponding values of the diurnal threshold. The obtained values can be considered as simulated extreme electricity prices in the absence of any ceilings. To compare those prices with the original (censored) ones, we truncated the simulated values at the level equal to the price ceilings of the corresponding original spikes, i.e, at the level of 12500AUD/MWh, 10000AUD/MWh, or 5000AUD/MWh depending on the time of the original spike occurrences. The results, documented in Table 4, clearly indicate that the simulated prices acceptably reproduce (in range of one standard deviation) the first two moments of the original extreme prices and autocorrelation of the original spike magnitudes.

In light of the estimation results presented in this section, it seems that our model provides a reasonable description of the spike magnitudes by capturing their heavy-tails, strong positive dependence, and intra-day variability.

4. Modelling durations between spike occurrences

In this section, we concentrate on modelling times of the spike occurrences. Inspecting Figure 4, it becomes apparent that there was a systematic shock in Australia's electricity market at the beginning of 2007 causing monthly proportions of the spikes to reach the level of 60% in all the

Table 4: Descriptive statistics of the actual and simulated prices (500 simulations).

	actual	simulated	actual	simulated
	NSW		QLD	
mean	412.4	390.9 (55.97)	364.5	386.1 (39.76)
std	1237.9	1306.3 (174.9)	1037.9	1317.0 (128.2)
autocorr(1)	0.876	0.866 (0.004)	0.851	0.854 (0.004)
	SA		VIC	
mean	474.5	452.3 (51.95)	263.3	232.6 (30.04)
std	1573.1	1488.6 (143.6)	878.3	893.6 (136.9)
autocorr(1)	0.812	0.799 (0.006)	0.853	0.827 (0.006)

Notes: Standard deviations of the characteristics for simulated prices are reported in parentheses. Row "autocorr(1)" denotes the Spearman's rank correlation between 1-lagged simulated spike magnitudes.

regions. Since explaining the occurrence of those systematic shocks is not the aim of statistical models, we omit that period in modelling times of the spike occurrences covering only the period over January 1, 2008–December 31, 2010 for estimation of the model, leaving the spikes happened over January 1, 2011–December 31, 2011 for the out-of-sample evaluation. In Section 4.1 we define spike durations and indicate their main features. A comparison of some existing approaches for modelling times of spike occurrences is provided in Section 4.2. Section 4.3 introduces a new model for spike durations. Estimation results are provided in Section 4.4.

4.1. Spike durations

Under a spike duration, or simply duration, we understand a time interval between occurrences of two consecutive spike. In Australia's electricity market, the smallest duration constitutes 30 minutes, – we shall refer to that duration as a unit duration and assign a value of one to it. Note that the unit duration denotes the smallest time interval between occurrences of two consecutive spikes. Time intervals of 60 minutes correspond to durations of two, intervals of 90 minutes correspond to durations of three, and so on. Throughout the paper $D_1, D_2, \dots, D_N$ will denote a sample of N consecutive (random) spike durations.

A major challenge of modelling spike durations in Australia's electricity market lies in their large variation and high proportion of unit durations (at least 62%), see Table 5, indicating a strong persistence of the spike occurrences and a distinctive integer character of the durations. There are many models in the literature which may capture those distinctive features of the spike durations. In the next section, we compare the performance of some of those models.

Table 5: Descriptive statistics for the spikes durations.

	NSW	QLD	SA	VIC
mean	64.72	95.57	52.46	43.76
std	284.09	438.83	226.78	213.59
proportion of unit durations	0.68	0.62	0.69	0.67
number of observations	760	539	969	1168

Notes: spike durations are measured in units of 30 minutes.

4.2. Models for the spike durations

A well-known model for durations is the autoregressive conditional duration (ACD) suggested by Engle and Russell (1998), see Bauwens and Hautsch (2009) for an overview of extensions and applications of this model. Another suitable approach for duration (actual time) modelling is the Hawkes process (Hawkes, 1971). Applications of the Hawkes process to modelling financial time series can be found in Embrechts, Liniger, and Lin (2011), Aït-Sahalia, Cacho-Diaz, and Laeven (2011), Chavez-Demoulin, Davison, and McNeil (2005), Bowsher (2007).

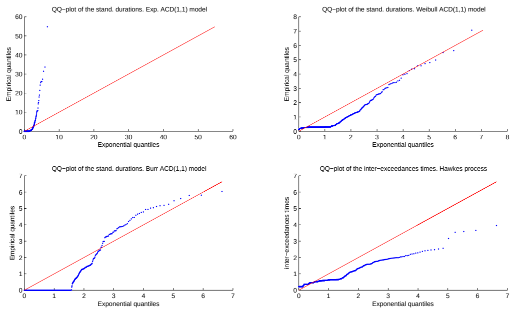

Figure 9: QQ-plot of the standardized durations (transformed by the theoretically implied distribution to the standard exponential) of the estimated ACD models and the residual inter-arrivals times of the estimated Hawkes process. The models were estimated on NSW spike durations occurred in the period over January 1, 2008–December 31, 2010.

To demonstrate the performance of the those approaches, we estimated four models on NSW spike durations from the period over January 1, 2008–December 31, 2010: Exponential ACD(1,1), Weibull ACD(1,1), Burr ACD(1,1) (Grammig and Maurer, 2000), and the univariate Hawkes process with an exponential response function. As a measure of goodness-of-fit of the estimated models, Figure 9 illustrates a plot of empirical quantiles of the standardized durations (transformed by theoretically implied distribution into standard exponential) of the estimated ACD models and the residual inter-arrivals times, see Embrechts, Liniger, and Lin (2011) for a definition, of the estimated Hawkes process versus corresponding quantiles of the standard exponential distribution.

For a reasonable fit of the models one expects the standardized durations and the residual inter-arrivals times to follow the standard exponential distribution. The QQ-plots indicate a strong deviation from the standard exponential distribution suggesting that the estimated ACD models and the Hawkes process are inappropriate for describing the spike durations (these estimation results are similar to SA, QLD, and VIC regions). A possible reason for a poor performance of the ACD model is its implied linearity of impact of past durations on the expected value of the future ones. This linear structure of the conditional expectation may be insensitive to capture both the large dispersion and the strong clustering behaviour of the spike durations. Furthermore, none of the models can accommodate the prominent integer character of the spike durations, which clearly can be observed as a sharp bend in the QQ-plots.

4.3. Negative binomial duration model

For description of the spikes durations we need a model which can reproduce their large variation and strong clustering pattern, and, finally, be of a discrete nature as the spikes durations are. A possible candidate which can meet those requirements is a model based on a negative binomial distribution. This distribution can be regarded as a gamma mixture of Poisson distributions, implying that it always has a higher ratio of variance to mean than a corresponding Poisson distribution. This feature is beneficial for modelling a large variation of the durations. A recent application of the negative binomial model for time series can be found in Davis and Wu (2009).

A random variable X whose distribution is negative binomial with parameters $r > 0$ and $p \in (0,1)$ has the mean $\mu = \frac{r(1-p)}{p}$, the variance $\sigma^2 = \frac{r(1-p)}{p^2}$, and the probability mass function

$$f_{NB}(k; r, p) := P(X = k) = \frac{\Gamma(r+k)}{\Gamma(k+1)\,\Gamma(r)} p^r (1-p)^k, \qquad k = 0, 1, 2, \dots, \tag{12}$$

where $\Gamma(\cdot)$ is the gamma function: $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$. Note for a positive integer k, the gamma function is related to the factorial by $\Gamma(k) = (k-1)!$.

4.3.1. Model description

For modelling durations we propose the following approach. Let $D_1, D_2, \ldots, D_N$, be a series of (spike) durations. We assume that the conditional distribution of D_i depends only on $D_1, \ldots, D_{i-1}$ and it can be expressed in the following way

$$\mathrm{P}\left(D_i = k \mid \mathbb{H}_{t_{i-1}}\right) = f_{NB}(k-1; r, p_i), \qquad k = 1, 2, \ldots \tag{13}$$

where $f_{NB}(\cdot; r, p)$ is a probability mass function of the negative binomial distribution, see (12), and p_i is a function of $(D_1, \ldots, D_{i-2}, D_{i-1})$. Recall that $\mathbb{H}_{t_{i-1}}$ is a history of the first $(i-1)$ spikes including their magnitudes $(y_1, \ldots, y_{i-1})$ and times of occurrences $(t_1, \ldots, t_{i-1})$. To account for the strong persistence of the spike occurrences, we suggest the following parametrization for p_i:

$$p_i = \omega + \alpha^{D_{i-1}} p_{i-1}, \qquad \alpha \in (0, 1). \tag{14}$$

This parametrization comes from a simple $AR(1)$ process and is intuitive as it adjust the dependence between two consecutive spikes to a time span between their occurrences: when D_{i-1} is small then D_i is almost identically distributed (provided α is close to 1); in case D_{i-1} is large, $\alpha^{D_{i-1}} \to 0$ implying that p_{i-1} has a small contribution to p_i.

4.3.2. Estimation

Estimation of model (13) is easily performed by the maximum likelihood method. With conditional distribution (13) and probability mass function (12), the log-likelihood function of durations $D_1, \ldots, D_N$ takes the form

$$\mathcal{L}(D_1, \ldots, D_N; r, \omega, \alpha) = \sum_{i=1}^{N} \left(\log \Gamma(r + D_i - 1) - \log \Gamma(r) + r \log p_i + (D_i - 1) \log(1 - p_i)\right). \tag{15}$$

Maximizing the likelihood we set a condition that a sample mean of durations $D_1, \ldots, D_N$ should be equal to the expected value of their conditional distributions implied by the model, namely, $\frac{1}{N}\sum_{i=1}^{N} D_i = \frac{1}{N}\sum_{i=1}^{N}\left(1 + r\frac{1-p_i}{p_i}\right)$. This condition is quite useful as it allows to express $r = \frac{\sum_{i=1}^{N}(D_i - 1)}{\sum_{i=1}^{N}\frac{1-p_i}{p_i}}$ and by this to reduce the number of parameters to be estimated from three parameters to two: ω ans α. For the overall stability of the estimated model, it is necessary that the condition $p_i \in (0, 1)$ (implied by definition (12)) holds. Expressing p_i from (14) as

$$p_i = \omega\left(1 + \alpha^{D_{i-1}} + \alpha^{D_{i-1}+D_{i-2}} + \alpha^{D_{i-1}+D_{i-2}+D_{i-3}} + \ldots\right),$$

it is easy to see that p_i achieves its minimum of ω as $D_{i-1} \to \infty$ and its maximum of $\frac{\omega}{1-\alpha}$ as $D_{i-k} = 1$, for all $k = 1, \ldots, (i-1)$. Since p_i should lie within the unit interval, the following two conditions on the parameter estimates should hold: $\omega > 0$ and $\frac{\omega}{1-\alpha} < 1$.

4.3.3. Simulation and Goodness-of-fit

In this section we consider a simulation procedure and a goodness-of-fit test for the proposed duration model. By considering the inverse of the negative binomial distribution function, it is straightforward to simulate duration D_i that follows model (13) in the following way

$$D_i \sim 1 + \min\left[k : U_i \leq F_{NB}(k; r, p_i)\right], \tag{16}$$

where $U_i \sim \mathrm{Unif}[0, 1]$ and $F_{NB}(\cdot; r, p)$ is a distribution function of the negative binomial distribution with parameters r and p. Note that $\min\left[k : u \leq F_{NB}(k; r, p)\right]$ denotes an inverse of $F_{NB}(\cdot; r, p)$ in point u. To construct a sample of simulated durations, function p_{i+1} should be updated according to (14) after each realization of D_i, and then the realization of D_{i+1} can be found from (16).

Our approach to the goodness-of-fit test is based on the adaptation of the probability integral transformation, discussed in the continuous case in Diebold, Gunther, and Tay (1998), to the discrete case. Considering (16) in the way that D_i (sample duration) and p_i are known and U_i is unknown one can can reproduce D_i from U_i, but vice versa it is not true, because the distribution function of D_i is discrete. The only information on series $U_1, U_2, \ldots, U_N$ which can be extracted from the sample data $D_1, D_2, \ldots, D_N$ is that U_i satisfies

$$U_i \sim \text{Unif}\left[F_{NB}\left(D_i - 2; r, p_i\right), F_{NB}\left(D_i - 1; r, p_i\right)\right], \qquad i = 1, 2, \ldots, N. \tag{17}$$

We shall refer to U_i defined above as a generator of D_i.

In case $D_1, D_2, \ldots, D_N$ really follow model (13), generators $U_1, U_2, \ldots, U_N$ should constitute N realizations from the uniform on $[0, 1]$ distribution. In practice, therefore, the goodness-of-fit test of the negative binomial duration model can be checked by testing the null hypothesis that the sample of generators $U_1, U_2, \ldots, U_N$ estimated according to (17) for a given sample of durations follow the uniform on $[0, 1]$ distribution. The goodness-of-fit can be checked either graphically using QQ-plots, or formally using the Kolmogorov-Smirnov and Anderson-Darling tests. Note that since for a fixed sample of durations $D_1, D_2, \ldots, D_N$ the sample of estimated generators is random, the testing of the null hypothesis should be conducted sufficient many times and then an non-rejection rates of the null hypothesis should be analysed.

4.4. Estimation results

In this section we estimate the model of Section 4.3.1 on the spike durations from the four regions of Australia's national electricity market covering the period over January 1, 2008–December 31, 2010. The parameters estimates, with the 99% confidence intervals in parentheses, are reported in Table 6. The confidence intervals are computed by using the profile log-likelihood function, because simulations and practical experience suggest these intervals provide better results than those derived by using the numerical Hessian matrix, see, e.g., Coles (2001). Note that the parameter estimates

Table 6: Parameter estimates of the negative binomial duration model estimated on the spike durations.

	NSW	QLD	SA	VIC
$\hat{\omega} * 10^4$	4.83 [3.67, 7.58]	4.01 [3.02, 6.02]	2.10 [1.53, 2.84]	2.13 [1.57, 2.92]
$\hat{\alpha} * 10$	6.93 [1.65, 8.79]	7.63 [3.85, 8.80]	9.91 [9.81, 9.95]	9.94 [9.88, 9.96]
$\hat{r}$	0.0541	0.0667	0.0605	0.0687

Notes: The profile log-likelihood function is used to compute the 99% confidence intervals reported in squared parentheses right to the estimates.

meet the necessary conditions of the overall stability of the model, namely, $\hat{\omega} > 0$ and $\frac{\hat{\omega}}{1-\hat{\alpha}} < 1$.

In order to check the goodness-of-fit of the estimated model, we employ the procedure of Section 4.3.3 and test the hypotheses that the estimated generators, first, follow the uniform distribution on $[0, 1]$ and, second, exhibit no autocorrelation. Those hypotheses were tested with, respectively, the Kolmogorov-Smirnov and Ljung-Box (10 lags) tests, which were conducted on 1000 different realisations of the estimated generators. Table 7 reports the non-rejection rates of the conducted tests with a significance level of 1%.

Table 7: Goodness-of-fit test: non-rejection rates (in %) of the Kolmogorov-Smirnov and Ljung-Box (10 lags) tests with a significance level of 1% conducted on 1000 random samples of the estimated generators.

	NSW	QLD	SA	VIC
Kolmogorov-Smirnov	99.7	99.3	99.6	99.5
Ljung-Box(10)	75.8	92.9	40.4	26.4

When the estimated generators really follow the uniform on $[0, 1]$ distribution, then the non-rejection rate of the Kolomogorov-Smirnov test with 1% significance level would approximately be 99%, which exactly corresponds to the rates in the above table. The results of the Ljung-Box test are less convincing but still in a high proportion of cases the generators can be assumed to have no autocorrelation. In order to get a graphical presentation of the goodness-of-fit and to compare it with the fit of models in Section 4.2, we transform a typical sample of estimated generators $\left(\hat{U}_1, \ldots, \hat{U}_N\right)$

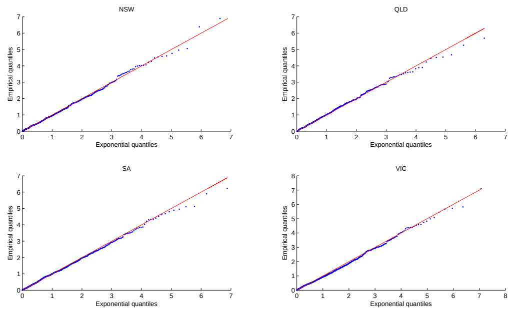

Figure 10: QQ-plot of a typical sample of the estimated transformed generators. Compare this figure with Figure 9.

(which are supposed to be uniformly on $[0, 1]$ distributed) into $\left(-\log \hat{U}_1, \ldots, -\log \hat{U}_N\right)$ (which are hence supposed to have the standard exponential distribution) and plot its quantiles versus quantiles of the standard exponential distribution, see Figure 10. Comparing this QQ-plot with that of Figure 9, one can observe a clear improvement in the fit of the estimated model to the spike durations.

Verifying accuracy of the estimated duration models, we have simulated 500 samples of durations (of the same length as the original ones) and compared their characteristics to those of the original spike durations. Simulation results are summarized in Table 8. The characteristics of the simulated

Table 8: Descriptive statistics of the actual and simulated durations (500 simulations).

	actual	simulated	actual	simulated
	NSW		QLD	
mean	64.72	65.24 (10.90)	95.57	97.64 (19.75)
std	284.09	291.57 (65.11)	438.83	400.01 (103.43)
proportion of unit durations	0.683	0.685 (0.018)	0.622	0.621 (0.021)
	SA		VIC	
mean	52.46	53.92 (14.20)	43.76	49.97 (13.06)
std	226.78	301.92 (116.15)	213.59	275.56 (115.77)
proportion of unit durations	0.689	0.680 (0.019)	0.666	0.656 (0.018)

Notes: The length of the simulated samples corresponds to the length of the original durations, see Table 5.

data are very close (in the range of one standard deviation) to those of the initial sample indicating the ability of our model to produce realistic simulations of spike durations. The major reason for some inconsistencies in the simulated data lies in the large variation of durations implied by the model, for example, with parameters estimates of the VIC region, the implied standard deviation of the duration varies from 7.6 to 1230.2 (depending whether p_i approaches respectively its maximum or minimum value). On the other hand, exactly that feature in combination with the dynamic structure of the model is necessary in reproducing the high variation of the spike durations.

5. Forecasting extreme electricity prices

For good risk management at electricity markets it is essential to accurately forecast extreme electricity prices in order to prevent unexpected losses. In this section we combine the results from Section 3, modelling magnitudes of the spikes, and Section 4, modelling the spike durations, into one model for forecasting extreme electricity prices.

5.1. Forecasting approach

The estimated in Section 4.4 duration model readily lends itself for estimating the probability of price spikes conditional on their past. The probability that a spike occurs at time t conditional

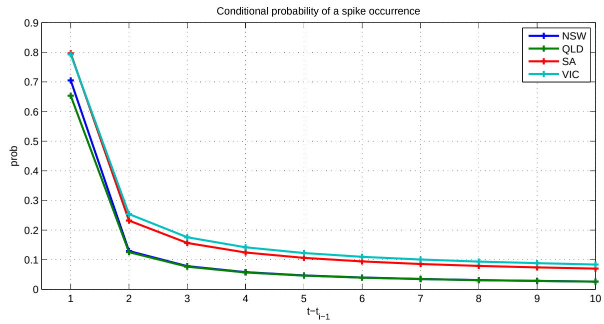

Figure 11: The conditional probability of a spike occurrence on the four regions of Australia's electricity market. The probability was estimated according to (18) with parameters values from Table 6. p_i was set on its max achievable value: $p_i = 0.0016$ for NSW; $p_i = 0.0017$ for QLD; $p_i = 0.0232$ for SA; $p_i = 0.0335$ for VIC.

the last spike with duration D_{i-1} happened at time t_{i-1} can be expressed as

$$\mathrm{P}(\text{spike occurs at time } t \mid \mathbb{H}_{t-1}) = \frac{\mathrm{P}(D_i = t - t_{i-1} \mid \mathbb{H}_{t-1})}{\mathrm{P}(D_i > t - t_{i-1} - 1 \mid \mathbb{H}_{t-1})},$$

where D_i follows model (13), $\mathbb{H}_{t-1}$ is an information set consisting of times $(t_1, \ldots, t_{i-1})$ and magnitudes $(y_1, \ldots, y_{i-1})$ of the spikes up to time $(t-1)$. In terms of model (13) the above probability takes the form

$$\mathrm{P}(\text{spike occurs at time } t \mid \mathbb{H}_{t-1}) = \frac{f_{NB}\left(t - t_{i-1} - 1; r, p_i\right)}{1 - F_{NB}\left(t - t_{i-1} - 2; r, p_i\right)}, \tag{18}$$

where $F_{NB}(\cdot; r, p)$ is a distribution function of the negative binomial distribution with parameters r and p; $f_{NB}(\cdot; r, p)$ is the corresponding probability mass function. Figure 11 illustrates the above conditional probability calculated with the parameters estimates of Table 6. Note that for the calculation of the conditional probabilities on the plot, we set parameter p_i on its maximum achievable value in order to get the maximum achievable conditional probabilities of spike occurrences.

Equation (18) provides a conditional probability that a spike occurs, i.e., that electricity price exceeds the diurnal threshold defined in Section 2. Surely, the electricity market participants may be interested in probabilities that the prices exceed other thresholds: a common example is a price level of 300AUD/MWh which is the strike price of heavily-traded cap products in Australia's electricity market. Designing an approach to estimate those probabilities, one can informally express the probability of the price to exceed some threshold y (provided it is higher than the diurnal threshold) in the following way

$$\begin{aligned}&\mathrm{P}(\text{price at time } t \text{ exceeds } y \ \mid \mathbb{H}_{t-1}) = \\ &\quad \mathrm{P}(\text{spike occurs at time } t \mid \mathbb{H}_{t-1})\mathrm{P}(\text{price at time } t \text{ exceeds } y \ \mid \text{spike occurs at time } t, \mathbb{H}_{t-1}).\end{aligned}$$

Combining the model of Section 3 for the spike magnitudes (see Eq. (4)), and the model of Section 4.3 for the spike durations (see Eq. (18)), the above probability takes the form

$$\begin{aligned}&\mathrm{P}(\text{price at time } t \text{ exceeds } y \mid \mathbb{H}_{t-1}) = \\ &\qquad \frac{f_{NB}\left(t - t_{i-1} - 1; r, p_i\right)}{1 - F_{NB}\left(t - t_{i-1} - 2; r, p_i\right)} \left[1 + \frac{g_t(y - th_{m(t)})^{\theta_t/\xi_{m(t)}} - 1}{g_{t_{i-1}}(y_{i-1})^{\theta_t/\xi_{m(t_{i-1})}}}\right]^{-(1/\theta_t + 1)}\end{aligned} \tag{19}$$

where t_{i-1} is the time of the last up to time $(t-1)$ spike occurrence, y_{i-1} is the magnitude of that spike; $g_t(y) = 1 + \xi_{m(t)} \frac{y}{\beta_{n(t)}}$, where $\xi_{m(t)}$ and $\beta_{n(t)}$ denote parameters of the GPD used for modelling the magnitude of the t-th observation; $\theta_t = \gamma_0(t - t_{i-1})^{-\gamma_1}$, and $th_{m(t)}$ denotes the value of diurnal threshold corresponding to the t-th observation; finally, $m(t)$ and $n(t)$ denote a function that identifies, respectively, to which part of the day where the shape parameters are assumed equal or to which out of 48 half-hour periods of the day the t-th observation belongs. For a further explanation of the parameters see Section 3.1 and Section 4.3.1.

Model (19) for forecasting the probability of extreme price occurrences provides two beneficial features. First, although the model is estimated on the price exceedances over the diurnal threshold (this approach yields sufficient number of observation for the estimation), the model can provide probabilities of the prices to exceed any higher levels without a need for re-estimation of the model. Second, (19) suggests a mechanism how spikes, defined as price exceedances over the comparatively small diurnal threshold, may trigger the occurrence of price exceedances over much higher thresholds than the diurnal one. This relationship is provided in two channels: first, spike occurrence triggers the occurrence of further spikes through (14); second, magnitude y_{i-1} of the last spike impacts the conditional distribution of the magnitude of the next spike through (4).

5.2. Out-of-sample forecasting performance

As it was noted in Section 4, the period over January 1, 2011–December 31, 2011 of the data of electricity prices was left for the out-of-sample forecasting evaluation. Note that this evaluation requires computation of (19), which, in turn, requires the estimates of the model for spike magnitudes reported in Table 2. Although those estimates were obtained analysing the whole sample of spikes, excluding the out-of-sample period in modelling of the magnitudes does not significantly affect the out-of-sample forecasting results of this section.

In order to analyse the forecasting performance of the model presented in this paper, we adopt the procedure suggested in Eichler, Grothe, Manner, and Tuerk (2012). In that study, the out-of-sample performance of seven different models was compared based on their ability to make 1-step ahead forecasts of electricity prices to exceed 300AUD/MWh (and 100AUD/MWh) analysing the same dataset as we use for this out-of-sample evaluation. According to that study, a sample of forecasted price exceedances over 300AUD/MWh was constructed using the true history of the process in the way that for each period when the estimated probability exceeds the value of 0.5 then a price exceedances was forecasted. The forecasting performance of the models was compared based on the correct detection rate (CDR), the ratio between correctly detected and the observed spikes, and the false detection rate (FDR), the ratio between falsely detected and the total number of detected spikes.

Using our model for forecasting exceedances of the electricity prices over 300AUD/MWh, we changed slightly the procedure of deciding whether an exceedance was forecasted. Since our duration model is based on a discrete distribution, it cannot provide probabilities filling the whole interval of $[0, 1]$, contrary to the models in Eichler, Grothe, Manner, and Tuerk (2012). To analyse accurately the performance of our model, we adjust the probability level, exceeding which we decide whether the price exceedance occurs, from 0.5 to one half of the maximum spike probability that our model can provide (see probabilities at $t - t_{i-1} = 1$ on Figure 11). For example, with parameter estimates for VIC region, that probability threshold is approximately equal to 0.4.

Table 9 provides the out-of-sample performance of the models in forecasting the electricity prices to exceed the level of 300AUD/MWh. We compare the performance of our model to the best models analysed in Eichler, Grothe, Manner, and Tuerk (2012). The best models were chosen (for each region) based on two criteria: the model with the best (i.e., maximum) CDR and the model with the best (i.e., minimum) FDR. Corresponding columns that refer to those models are denoted as "Best CDR" and "Best FDR". Note that those models are not the same for each of the regions.

An ideal model for spike forecasting provides CDR= 100% and FDR= 0%. In practice, however, there is often a trade-off between the high CDR and the low FDR. From Table 9 it is apparent that the performance of our model is always somewhere in the middle compared to performance of the other models analysed in the table: our model provides either a higher CDR or a smaller FDR. The only exception constitutes the performance of our model for the VIC region.

Table 9 provides only a limited assessment of the forecasting performance of our model, because it is suited to estimate probabilities of the prices to exceed any sufficiently high level, not just the level for which the model was estimated. As a demonstration of that feature, we estimated

Table 9: Out-of-sample performance of the models in forecasting electricity prices exceeding 300AUD/MWh.

	Our model	Best CDR	Best FDR	Our model	Best CDR	Best FDR
	NSW			QLD		
exceedances	38	38	38	37	37	37
detections	58	77	38	43	30	30
CDR	84.2	94.7	76.3	59.5	54.1	54.1
FDR	44.8	53.6	23.7	48.8	33.3	33.3
	SA			VIC		
exceedances	29	29	29	11	11	11
detections	25	29	12	10	10	10
CDR	48.3	55.2	34.5	54.6	63.6	63.6
FDR	44.0	44.8	16.7	40.0	30.0	30.0

Notes: rows "exceedances" and "detections" denote respectively the number of the actual and forecasted prices exceeding the level of 300AUD/MWh. Columns with headings "Best CDR" and "Best FDR" refer to the models with respectively maximum CDR and minimum FDR analysed in Eichler, Grothe, Manner, and Tuerk (2012). CDR and FDR are reported in %.

1-step ahead probabilities of the electricity prices to exceed different price levels: 500AUD/MWh, 1000AUD/MWh, 2000AUD/MWh, and 5000AUD/MWh, – and, applying the same procedure as used for construction of Table 9, we evaluated the out-of-sample forecasting performance of our model. Table 10 provides the evaluation results. Unexpectedly, but the forecasting performance of the model for higher price threshold was only slightly decreased compared to the results in Table 9. Moreover, for some regions the duration model showed even better results, for example, for SA region, eight from nine spikes over 5000AUD/MWh were correctly forecasted (in the sense that the probability exceeds some level).

This ability to forecast the electricity price exceedances over high thresholds is a unique and valuable feature of our model. Other approaches for modelling extreme electricity prices can experience estimation problems because very few data may be available fitting the model to the prices that exceed very high thresholds. For example, in Australia's electricity market, in the period over January 1, 2002–December 31, 2010, there were only a few out of 157728 observations when the electricity prices exceed the level of 5000AUD/MWh: 99 in NSW, 72 in QLD, 135 in SA, and 45 in VIC.

Table 10: Out-of-sample performance of our model in forecasting electricity prices exceeding 500AUD/MWh, 1000AUD/MWh, 2000AUD/MWh, and 5000AUD/MWh levels.

	NSW	QLD	SA	VIC	NSW	QLD	SA	VIC
	500AUD/MWh				1000AUD/MWh			
exceedances	30	28	24	8	30	23	22	8
detections	34	29	23	8	28	24	22	7
CDR	70.0	50.0	54.2	62.5	63.3	56.5	50.0	50.0
FDR	38.2	51.7	43.5	37.5	32.1	45.8	50.0	42.9
	2000AUD/MWh				5000AUD/MWh			
exceedances	22	19	19	5	13	8	9	3
detections	22	17	13	5	11	5	9	2
CDR	63.6	63.2	42.1	40.0	61.5	62.5	88.9	33.3
FDR	36.4	29.4	38.5	60.0	27.3	0	11.1	50.0

Notes: rows "exceedances" and "detections" denote respectively the number of the actual and forecasted price exceedances. CDR and FDR are in %.

6. Conclusions

This study presents a model for forecasting extreme electricity prices in real-time (high frequency) settings. The model consists of two components (sub-models) which deal separately with times of occurrence and magnitudes of extreme electricity prices. We employ a copula with a changing dependence parameter for capturing serial dependence in the magnitudes of extreme electricity prices and the censored GPD distribution for modelling their heavy tails. For modelling times of the extreme price occurrences, we propose an approach based on the negative binomial distribution. For both of the sub-models, the simulation procedure and the goodness-of-fit test are presented.

The model is applied to half-hourly electricity prices from the four regions of Australia's national electricity market embracing the period over January 1, 2002–December 31, 2011. The simulation studies and the goodness-of-fit tests indicate an ability of our model in capturing the main characteristics of extreme electricity prices. In particular, our approach to times of the extreme price occurrences outperforms the ACD models and the Hawkes process. The out-of-sample evaluation also indicates a convincing performance of our model in forecasting the prices to exceed very high thresholds.

A promising direction for a future research is to consider a multivariate approach for modelling extreme electricity spot prices. That suggestion is motivated by the fact that in interconnected regional markets, spikes in one region tend to trigger the occurrence of spikes in the other regions. Multivariate approaches can capture those interdependences and describe the contagion effects of extreme electricity prices.

References

Aït-Sahalia, Y., J. Cacho-Diaz, and R. J. Laeven (2011): "Modeling Financial Contagion Using Mutually Exciting Jump Processes," *NBER Working Paper No. w15850.*

Bauwens, L., and N. Hautsch (2009): "Modelling financial high frequency data with point processes," in *Handbook of Financial Time Series*, ed. by T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch. Springer Verlag.

Becker, R., S. Hurn, and V. Pavlov (2007): "Modeling Spike in Electricity Prices," *The Economic Record*, 83(263), 371–382.

Bowsher, C. G. (2007): "Modelling security market events in continuous time: Intensity based, multivariate point process models," *Journal of Econometrics*, 141, 876–912.

Chavez-Demoulin, V., A. C. Davison, and A. J. McNeil (2005): "Estimating Value-at-Risk: A point process approach," *Quantitative Finance*, 5(2), 227–234.

Christensen, T. M., S. Hurn, and K. A. Lindsay (2009): "It never rains but it pours: modelling the persistence of spikes in electricity prices," *The Energy Journal*, 30, 25–48.

——— (2012): "Forecasting spikes in electricity prices," *International Journal of Forecasting*, 28(2), 400–411.

Coles, S. G. (2001): *An Introduction to Statistical Modeling of Extreme Values*. Springer.

Davis, R. A., and R. Wu (2009): "A negative binomial model for time series of counts," *Biometrika*, 96(3), 735–749.

Davison, A. C., and R. L. Smith (1990): "Models for Exceedances over High Thresholds," *Journal of the Royal Statistical Society B*, 52, 393–442.

Diebold, F. X., T. A. Gunther, and A. S. Tay (1998): "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39(4), 863–883.

EICHLER, M., O. GROTHE, H. MANNER, AND D. TUERK (2012): “Modeling spike occurrences in electricity spot prices for forecasting,” *METEOR Research Memoranda*.

EMBRECHTS, P., T. LINIGER, AND L. LIN (2011): “Multivariate Hawkes Processes: an Application to Financial Data,” *Journal of Applied Probability*, 48(A), 367–378.

ENGLE, R. F., AND J. R. RUSSELL (1998): “Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data,” *Econometrica*, 66(5), 1127–62.

GEMAN, H., AND A. RONCORONI (2010): “Understanding the fine structure of electricity prices,” *Journal of Business*, 79, 1225–1261.

GRAMMIG, J., AND K.-O. MAURER (2000): “Non-monotonic hazard functions and the autoregressive conditional duration model,” *Econometrics Journal*, 3, 16–38.

GREENE, W. H. (2003): *Econometric Analysis*. Prentice Hall.

HAWKES, A. G. (1971): “Point Spectra of Some Mutually Exciting Point Processes,” *Journal of the Royal Statistical Society B*, 33(3), 438–443.

KLÜPPELBERG, C., T. MEYER-BRANDIS, AND A. SCHMIDT (2010): “Electricity spot price modelling with a view towards extreme spike risk,” *Quantitative Finance*, 10:9, 963–974.

MANDELBROT, B. B. (1963): “The variation of certain speculative prices,” *The Journal of Business of the University of Chicago*, 36, 394–419.

MCNEIL, A. J., R. FREY, AND P. EMBRECHTS (2005): *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press.

NELSEN, R. B. (2006): *An Introduction to Copulas*. Springer.

PATTON, A. J. (2012): “A review of copula models for econometric time series,” *Journal of Multivariate Analysis*, 110, 4–18.

PICKANDS, J. (1975): “Statistical inference using extreme order statistics,” *The Annals of Statistics*, 3, 119–131.