

Modelling the Impact of Wind Power Production on Electricity Prices by Regime-Switching Lévy Semistationary Processes

Almut E.D. Veraart

Abstract This paper studies the impact of wind power production on electricity prices in the European energy market. We propose a new modelling framework based on so-called regime-switching Lévy semistationary processes to account for forward-looking information consisting of predicted wind power generation. We show that our new regime-switching model, where the regime switch depends on the so-called wind penetration index, can describe recent electricity price data well.

1 Introduction

Renewable sources of energy are of increasing importance in modern energy markets. For instance, the European Union has set the target of increasing the share of energy from renewable sources by 2020 to 20%. In the German and Austrian energy market, which will be the focus of this paper, the most important source of renewable energy is wind, followed by biogas and solar. Since many renewable sources are highly dependent on weather conditions, they tend to increase the volatility of the corresponding energy prices. It is hence urgent and important to find reliable models which can describe electricity prices in these changing market conditions which can be used for risk assessment and management in energy markets.

The recent literature has presented a variety of both discrete-time and continuous-time series models which promise to describe the stylised facts of energy markets, see e.g. [3, 17] for reviews. However, reliable models which incorporate information on renewable sources have only recently emerged and have currently been restricted to discrete-time models, see e.g. [9–12, 18].

This paper contributes to the continuous-time literature by introducing for the first time a modelling framework which takes the forward-looking information available to market participants through wind production forecasts into account when

A.E.D. Veraart (✉)
Department of Mathematics, Imperial College London,
180 Queen's Gate, SW7 2AZ London, UK
e-mail: a.veraart@imperial.ac.uk

modelling electricity day-ahead prices. In doing so, it extends recent work by [1] who proposed to model electricity spot prices by so-called Lévy semistationary processes. Their model consists of a reduced form approach, which only considers electricity prices directly and does not take any price information from other fuels or commodities into account. Note that other types of forward-looking information, such as capacity constraints, have been incorporated in models for electricity prices by [8]. Also, [2] have developed a framework for incorporating forward-looking information in electricity or weather markets through an enlargement of filtrations approach.

With the increasing power generation through wind farms, we have observed that electricity prices at the European Energy Exchange (EEX) started to become negative, which happened for the first time in October 2008, and partially even exhibited rather extreme negative price spikes, see e.g. [16] and the references therein. Also, various articles have argued that increasing wind power production seems to decrease the overall price level, but tends to increase the observed volatility in the market, see e.g. [12]. These are important findings, which need to be incorporated into a modelling framework, one of which will be presented in this paper.

The outline for the remaining part of this article is as follows. In Sect. 2 we give a detailed description of the data from the EEX which will be used in our empirical analysis and we carry out an exploratory data analysis to motivate the new model we are going to introduce in this paper. Section 3 contains the main contribution, where we introduce the new class of regime-switching Lévy semistationary processes and show how they can be calibrated to our empirical data. Finally, Sect. 4 concludes.

2 Exploratory Data Analysis

This section presents the results of an exploratory data analysis of electricity price and wind production data from the European Energy Exchange, which motivates the new modelling framework which we will introduce in Sect. 3.

2.1 Description of the Data

Our empirical data analysis focuses on electricity prices and wind data for a time period from 01.01.2011 to 31.07.2014, i.e. consisting of 1308 days.

We consider three sets of data: electricity prices, their corresponding volumes (sometimes called loads) and wind production data. More precisely, daily EEX Phelix baseload and EEX Phelix baseload volume data (for Germany and Austria) were downloaded from Datastream and the EPEX spot website. In addition, we downloaded the *forecasted wind production data* for the four German Transmission System Operator (TSOs) (50Hz Transmission, Amprion, Tennet TSO, EnBW Transportnetze (Transnet)) and one Austrian TSO (Verbund (APG)). These data

Table 1 Summary statistics of the EEX Phelix baseload from 01.01.2011 to 31.07.2014

Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
-56.87	33.87	43.10	41.96	50.69	98.98

have been aggregated to obtain *daily* forecasts for the wind production for each of the five TSOs.

2.2 EEX Phelix Baseload Prices

First of all, we want to explore the specific properties of the time series of the electricity price data. We focus on the day-ahead electricity prices determined by a daily auction at 12:00 pm, 7 days a week all year (including statutory holidays). The underlying quantity to be traded is the electricity for delivery the following day in 24h intervals. The prices are bounded (currently between [-500, 3000] EUR/MWh). The EEX Phelix baseload is obtained as the daily averages of the 24h day-ahead prices for Germany and Austria.

From the summary statistics in Table 1, we notice that the times series does not feature any truly extremes spikes, which have occurred in older data sets from the EEX market. In addition, we notice that there are negative prices even in the baseload prices which consists of the average of the 24h prices, see also Fig. 1. The plot of the autocorrelation function reveals a clear weekly pattern which is one of the well-known stylised facts of such data.

When we study the distributional properties of the price data, see Fig. 2, we clearly observe that the empirical distribution is not well described by a Gaussian distribution, but appears to be asymmetric and features heavier tails, the latter is particularly pronounced when we focus on the left tail of the distribution.

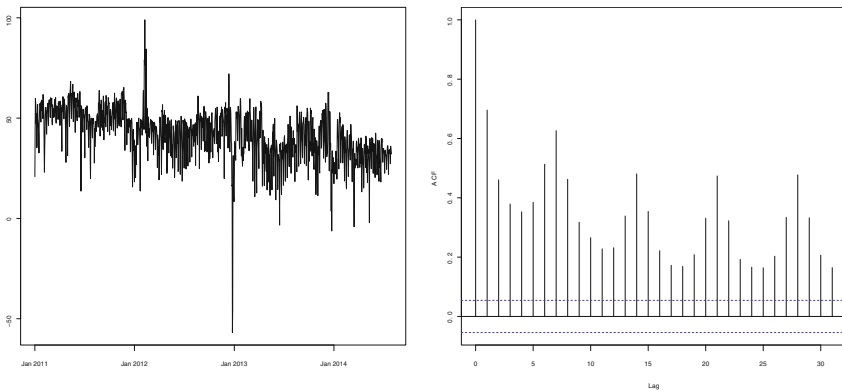


Fig. 1 Time series plot of the baseload prices (in EUR/MWh) and autocorrelation plot

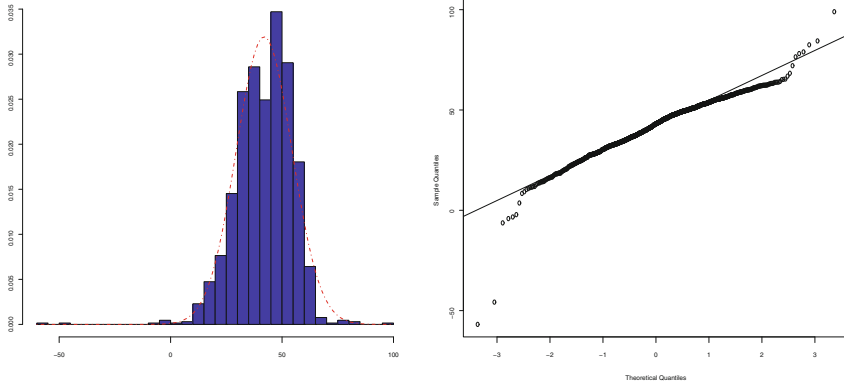


Fig. 2 Distributional properties of the prices: The standardised histogram of the empirical distribution and estimated kernel density function of the prices are depicted in the first plot. The second plot compares the empirical distribution to a Gaussian distribution via a quantile-quantile plot

2.3 Predicted Wind Energy Feed-In

Next we investigate the data from the five Transmission System Operator (TSOs) in Germany and Austria. Note that we are studying the *one-day ahead predicted wind feed-in* since we assume that this is the quantity which impacts the one-day ahead electricity prices determined in the daily auction.

Note that for each TSO, the data is available in 15 min intervals, where the unit of measurement is Megawatt (MW). In order to get the hourly forecasts, we aggregated the data as follows. Let $V_{t,q(i)}^{(q)}$ denote the 15 min wind power forecast for quarter i within hour t . We then obtain the hourly forecasts (recorded in Megawatt hours (MWh)), denoted by $V_{t(j)}^{(h)}$, where $t(j)$ denotes the j th hour on the t th day, from

$$V_{t(j)}^{(h)}[\text{MW} \cdot \text{h}] = \sum_{i=1}^4 \frac{1}{4} [\text{h}] V_{t(j),q(i)}^{(q)}[\text{MW}].$$

Further, we obtain the daily forecasts, denoted by $V_t^{(d)}$ from

$$V_t^{(d)}[\text{MW} \cdot \text{h}] = \sum_{j=1}^{24} V_{t(j)}^{(h)}[\text{MW} \cdot \text{h}].$$

The data are summarised by TSO in Table 2.

In the following study, we focus on the aggregated time series which accumulates the predicted wind feed-in of all five TSOs, see Table 3 for the summary statistics and Fig. 3 for the time series plot. We clearly observe a yearly cycle in the wind data showing that the predicted wind feed-in is always highest during the winter time period.

Table 2 Forecasted wind feed-in in Gigawatt hours (GWh) (rounded)

	50 Hz	Amprion	APG	Tennet	Transnet	Total
2011	19490	6687	18	18484	416	45095
2012	20203	7253	25	20464	246	48191
2013	19129	7742	31	21259	425	48585
2014 (Until July)	11648	4799	21	13412	399	30279

Table 3 Summary statistics of the forecasted aggregated wind production data (in GWh) from 01.01.2011 to 31.07.2014

Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
12510	53870	98940	131600	176000	572600

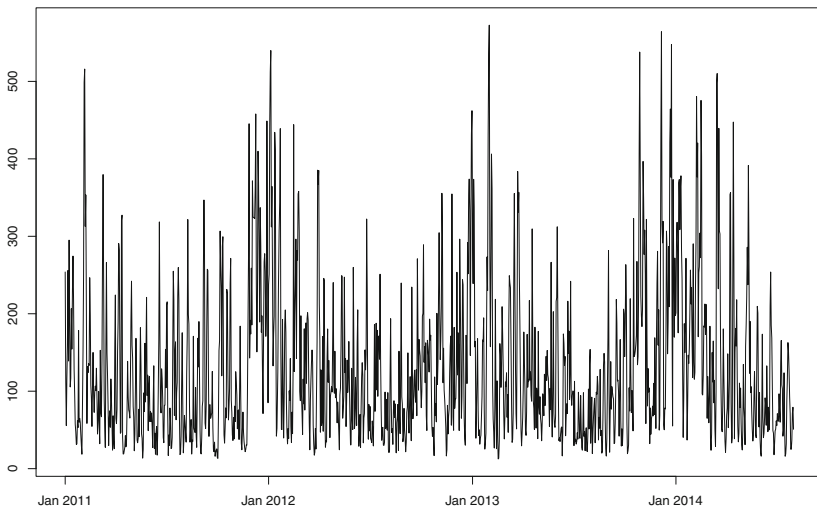


Fig. 3 Time series of the forecasted aggregated wind production data from 01.01.2011 to 31.07.2014 reported in GWh

2.4 Wind Penetration Index

Jónsson et al. [10] pointed out that there is a non-linear and time dependent relationship between wind power forecasts and spot prices. Moreover, they found that “it is in fact the ratio between the forecasted wind power generation and the forecasted load that has the strongest association with the spot prices”, see [10, p. 314]. Hence, in the following analysis, we will not use the predicted wind feed-in data directly, but rather focus on the so-called *wind penetration index*, which describes the percentage of the wind feed-in compared to the total energy production.

Table 4 Summary statistics of the wind penetration index from 01.01.2011 to 31.07.2014

Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
0.02104	0.08796	0.15480	0.19050	0.26450	0.67540

In order to compute this index, we follow the approach outlined in [10]: They argued that according to recent work by [15], state-of-the-art forecasting models lead to load forecasts where the predicted load equals the actual load plus an error term, i.e. let AL_t denote the actual load on day t and let L_t denote the corresponding predicted load. Then

$$AL_t = L_t + \epsilon_t, \quad \text{where } \epsilon_t \sim N(0, \sigma^2).$$

Typically, the standard deviation σ is chosen as 2 % of the average load for the period considered.

Following this methodology, we downloaded the actual load data from Datastream and the EPEX website and computed the predicted load data by adding Gaussian perturbations to the actual load data. Clearly, this is not exactly the same as working with the predicted load information from each TSO directly, see [10] for a discussion, but the practical impact of this approximation has been found to be marginal.

We can now define the so-called *wind penetration index* on day t as

$$WP_t := \frac{V_t^{(d)}}{L_t}.$$

This is in fact the prediction of the wind penetration on day t , which is available on day $t - 1$ and can hence be considered as forward-looking information. We provide the summary information of the wind penetration index for our sample in Table 4.

Moreover, a time series plot and the corresponding histogram of the wind penetration index is depicted in Fig. 4. We observe that the time series plot of the wind penetration index resembles the one for the original wind data—including a yearly seasonal pattern. Also, the wind penetration index is overall rather low, which is indicated by the histogram and the quantile information contained in Table 4. This is not surprising since the conventional fuels still account for the majority of the electricity production in the European energy market.

2.5 The Relation Between Prices and Wind Data

Finally, we carry out an exploratory study of the relation between the electricity prices and the wind penetration index. In Fig. 5, we plot the electricity prices versus the wind penetration index to check whether we can spot any association between the

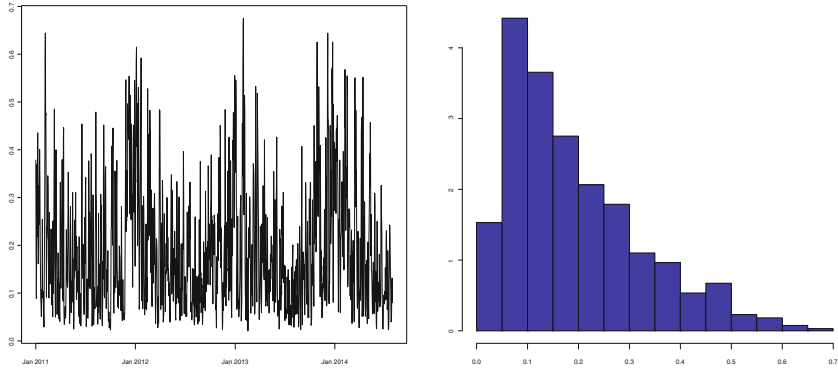


Fig. 4 Time series plot and histogram of the wind penetration index

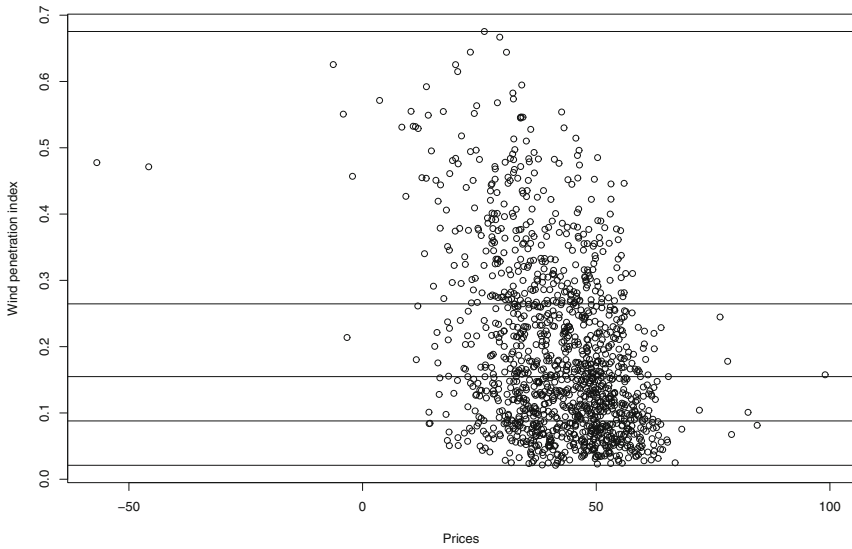


Fig. 5 The electricity prices are plotted versus the wind penetration index. The five horizontal lines correspond to the minimum, the 25%, 50%, 75% quantiles and the maximum of the wind penetration index, respectively

two variables. We observe that the two lowest electricity prices are associated with a rather high wind penetration index. Also, for a very high wind penetration index, the prices seem to be below their mean value. This is in line with earlier studies which found that a high wind production typically results in lower electricity prices. However, we need to keep in mind that by comparing the wind and the electricity prices, we can only obtain a partial picture, since clearly other fuels, such as coal, gas and nuclear, play a key role in determining the corresponding electricity price and are for the purpose of this study excluded from the analysis.

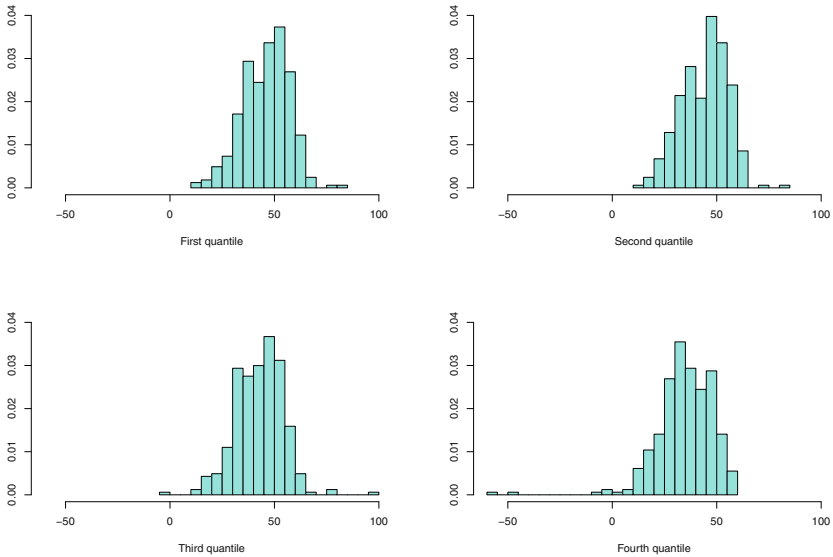


Fig. 6 Distribution of the electricity prices for different quartiles of the wind penetration index. E.g. the first plot corresponds to a low wind penetration index (the first quartile) and the last one to a rather high wind penetration index (the fourth quartile)

We also compare the distribution of the electricity prices associated with different quartiles of the wind penetration index. That is, we have divided our price data into four groups corresponding to the 1st, 2nd, 3rd and 4th quartile of the wind penetration index. When comparing the corresponding marginal distributions, we observe again that smaller price data are associated with a higher wind penetration index, see Fig. 6.

The finding from this exploratory study motivates the new modelling framework which we are going to introduce in the next section.

3 Model Building

Recent work by [1] suggests that the class of so-called *Lévy semistationary* (LSS) processes is very suitable for modelling electricity day-ahead prices. In their work, the class of LSS processes was used in a truly reduced form modelling set-up, meaning that the (deseasonalised) electricity prices were modelled directly by LSS processes and no other external variables were included in the analysis.

Here we will go one step further and explore the possibility of including *forward-looking information* in form of the wind penetration index into a new modelling framework which is based on LSS processes.

In a first step, we are going to review the basic traits of LSS processes and then we will discuss how an LSS-based model can be extended to account for forward-looking information.

An LSS process $Y = \{Y(t)\}_{t \in \mathbb{R}}$ on \mathbb{R} without drift is defined as

$$Y(t) = \int_{-\infty}^t g(t-s)\sigma(s-)dL(s), \tag{1}$$

where L denotes a two-sided Lévy process, $g : \mathbb{R} \rightarrow \mathbb{R}$ denotes a deterministic weight function satisfying $g(s) = 0$ whenever $s < 0$ and σ denotes a càdlàg, adapted stochastic volatility process, which is assumed to be independent of L . In order to ensure the existence of the stochastic integral, we need suitable integrability conditions on the kernel function g , see [1] for details.

Note that the name Lévy semistationary process indicates that the process Y is stationary as soon as the stochastic volatility process σ is a stationary process. The reason for choosing a stationary process for modelling deseasonalised electricity prices is that commodity prices typically exhibit strong mean reversion. A stationary process can in fact mimic such a behaviour since it ensures that the process cannot move away from its long term mean indefinitely, but will need to return to it since otherwise the stationarity would not be preserved.

Many well-known stochastic processes belong to the LSS class, including volatility modulated Ornstein-Uhlenbeck processes, continuous-time autoregressive moving average (CARMA) processes and fractionally integrated CARMA processes.

It is important to note that LSS processes are in general not semimartingales, which has been discussed in detail in [1]. However, in this paper, we will in fact be staying within the traditional semimartingale framework since we will be working with volatility modulated CARMA processes as the main building blocks for our new model. To this end, let us briefly recall their definition.

Suppose we have nonnegative integers $p > q$ and we wish to define a CARMA (p, q) process. We introduce the autoregressive (AR) and moving average (MA) polynomials:

$$P^{AR(p)}(z) = z^p + a_1z^{p-1} + \dots + a_p,$$

$$P^{MA(q)}(z) = b_0 + b_1z + \dots + b_{p-1}z^{p-1},$$

where $b_q = 1$ and $b_j = 0$ for $q < j < p$. Moreover, we assume that the polynomials have no common roots and then write formally

$$P^{AR(p)}(D)Y(t) = P^{MA(q)}(D)DL(t),$$

where $D = \frac{d}{dt}$. We can make sense of the “derivative” of the Lévy process through a state space representation, where we write

$$Y(t) = \mathbf{b}^\top \mathbf{V}(t), \quad \text{for } d\mathbf{V}(t) = \mathbf{A}\mathbf{V}(t)dt + \mathbf{e}dL(t), \quad \text{where}$$

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_p & -a_{p-1} & \cdots & \cdots & -a_1 \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{pmatrix}.$$

Assuming that all eigenvalues of \mathbf{A} have negative real parts, we know that

$$\mathbf{V}(t) = \int_{-\infty}^t e^{\mathbf{A}(t-s)} \mathbf{e} dL(s)$$

is the (strictly) stationary solution of the stochastic differential equation above, see [5]. That is, in our LSS specification we can choose $g(x) = \mathbf{b}^\top e^{\mathbf{A}x} \mathbf{e}$ and $\sigma \equiv 1$ to obtain a CARMA(p, q) process. As soon as a stochastic volatility process is added, we would call the corresponding LSS process a *volatility modulated CARMA process*.

3.1 Deseasonalising the Data

We argued before that stationary processes can easily accommodate key stylised facts of commodity prices. However, at the same time, we cannot ignore that strong seasonal effects are typically present in such markets and need to be accounted for. We proceed by introducing an *arithmetic* model for the electricity day-ahead price, denoted by $S = (S(t))_{t \geq 0}$, where

$$S(t) = \Lambda(t) + Y(t).$$

Here, Λ denotes a deterministic seasonality and trend function and Y denotes a stochastic process. In the original framework proposed by [1], the process Y was chosen to be an LSS process. In the following, however, we will introduce a modification of that modelling framework.

The seasonality and trend function is chosen to be

$$\Lambda(t) = c_0 + c_1 t + c_2 \cos\left(\frac{\tau_1 + 2\pi t}{365}\right) + c_3 \cos\left(\frac{\tau_2 + 2\pi t}{7}\right),$$

which accounts for a linear trend and weekly and yearly seasonal cycles. We used a robust least squares estimation procedure to estimate the parameters (by truncating the spikes in the estimation procedure) and obtained the following estimates and standard errors, see Table 5, all of which were highly significant. The estimated parameters in the seasonality function confirm the existence of both weekly and yearly seasonality as well as the presence of a negative trend.

Table 5 Estimated parameters in the seasonality and trend function Δ

	c_0	c_1	c_2	c_3	τ_1	τ_2
Estimate	53.47	-0.01724	2.377	-6.815	349.25	33.59
Standard error	0.4693	0.0006203	0.3356	0.3302	50.46	0.3394

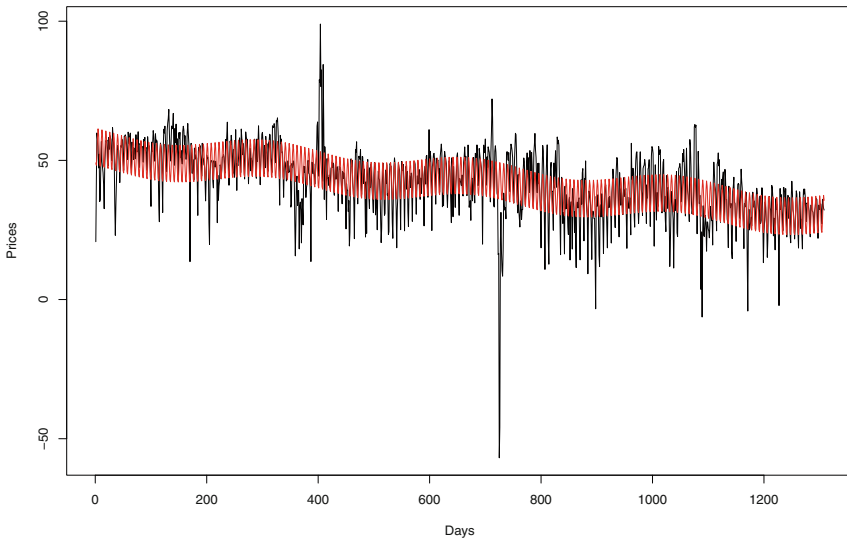


Fig. 7 Electricity day-ahead data and fitted seasonality and trend function

Figure 7 shows a graph of the original data with the fitted seasonality function super-imposed. Also, Fig. 8 depicts the deseasonalised price data and their autocorrelation function.

It should be noted that a variety of alternative procedures could be followed to deal with the problem of seasonality in the electricity prices. Here we are dealing with a rather simple deterministic parametric function to mimic the trend and the seasonal cycles. Using e.g. weekly and yearly dummy variables could refine this approach even further, but would result in a less parsimonious model. A more interesting alternative to the approach pursued here is to acknowledge the fact that the seasonality cannot only be determined by historical data, but also through other market data available to market participants. Example, it has been observed that gas and coal prices, given that they are important fuels used to produce electricity, play a key role in determining trend and also seasonal cycles of electricity prices. This suggests that e.g. forward curve data for gas and/or coal could be used to model the trend, see e.g. [8] for research along those lines. Also [14] give a detailed account on robust estimation procedures of the long-term seasonal component of electricity prices.

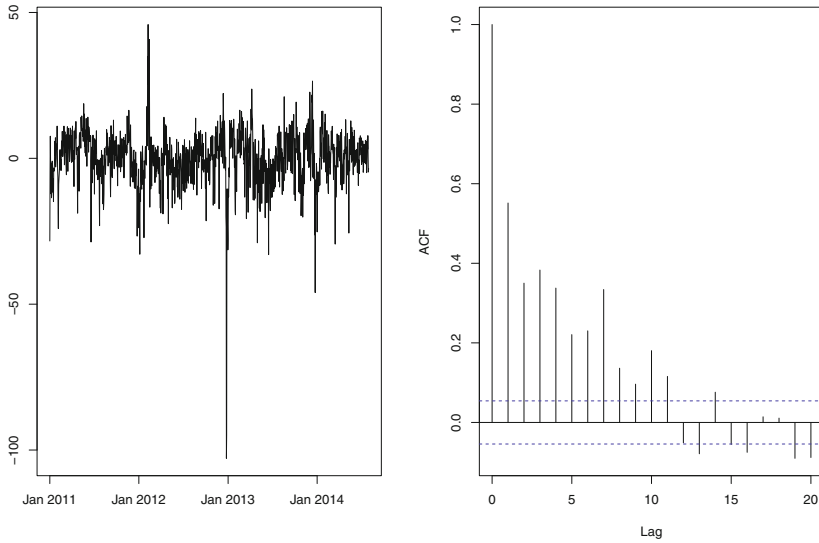


Fig. 8 Time series and autocorrelation plot of the deseasonalised electricity day-ahead data

Table 6 Estimated parameters of the associated ARMA(2,1) process

	AR1	AR2	MA
Estimate	1.1480	-0.2324	-0.6962
Standard error	0.0597	0.0447	0.0501

Here AR1 and AR2 corresponds to the first and second autoregressive parameter, respectively, and MA corresponds to the moving average parameter

3.2 Fitting a CARMA Process

After the seasonality has been removed, we need to find a suitable model for the stochastic process Y . Following the success of the CARMA processes in describing electricity prices, we choose a Lévy semistationary process where the kernel function is given by a kernel associated with a CARMA(p,q) process. More specifically we choose $p = 2$ and $q = 1$. Note that when choosing the order of the CARMA(p,q) process, we need to consider pairs (p,q) such that $p > q$ so that the CARMA process is well defined. We choose a CARMA(2,1) process due to reasons of analytical tractability and increased flexibility compared to a simple Ornstein-Uhlenbeck model. In our goodness-of-fit study, we indeed find that such a model choice is suitable here.

Note that [6] have discussed in detail how a discretely sampled CARMA process can be represented as a weak ARMA process. Using this representation we have first estimated the corresponding ARMA(2,1) parameters by a quasi-maximum likelihood method. The corresponding parameters estimates and standard errors are provided in Table 6.

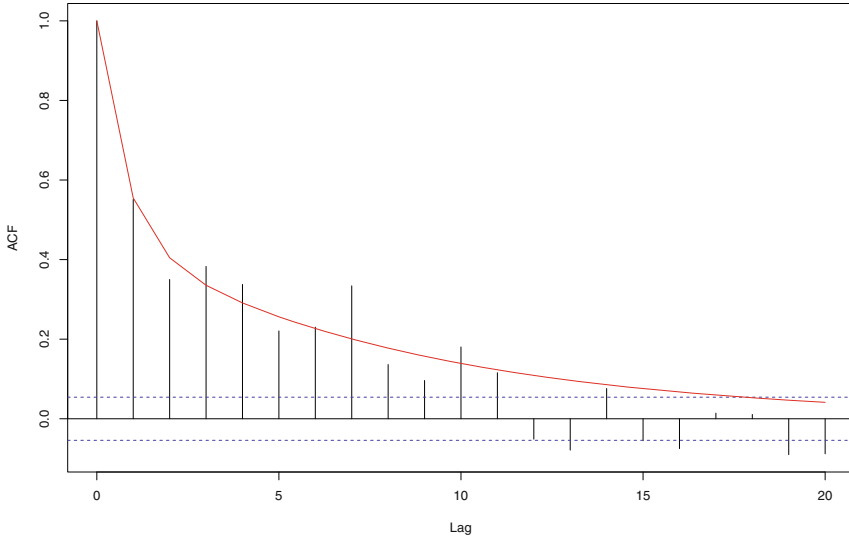


Fig. 9 Empirical (*bars*) and estimated (*solid line*) autocorrelation function of the estimated CARMA(2,1) process

Following the procedure outlined in [6] we can then recover the corresponding continuous-time parameters. In our case, we have $a_1 = 1.459, a_2 = 0.162, b_0 = 0.383$. Note that one can easily verify that the estimates satisfy the condition that the eigenvalues of \mathbf{A} have negative real parts, which implies a stationary model. We compare the empirical and estimated autocorrelation function in Fig. 9, where we observe a good fit.

Under the assumption that the CARMA process is driven by a subordinator, [6] have shown how the corresponding increments of the driving Lévy process can be recovered from a discretely observed CARMA process, see also [7] for the multivariate case. Here we have implemented their algorithm for the case of a CARMA(2,1) process and have recovered the driving process. Note that the original algorithm was designed for driving Lévy processes, but can in fact be adapted to the case of volatility modulated Lévy processes as well, the case which will be relevant in the next section.

Let us briefly recap our estimation procedure until now: We started off with a spot price model $S(t) = \Lambda(t) + Y(t)$, where we have estimated the seasonality function Λ and have removed it from the data. In the next step, we have assumed that Y is an LSS process of the form

$$Y(t) = \int_{-\infty}^t g(t - s) dM(s),$$

where g corresponds to the kernel function associated with a CARMA(2,1) process and initially M was assumed to be a Lévy process, which can be recovered from the observations of the CARMA process. We will now leave this traditional framework behind and will introduce a new regime-switching model based on LSS processes.

3.3 The New Model Based on a Regime-Switching LSS Process

The predicted wind penetration index can be viewed as forward-looking information since the information is available before the prices for the next day are determined in the auction market. Hence it is reasonable to try to incorporate this information in the model.

Previous studies have included such information in discrete-time models such as e.g. ARMAX-GARCHX models, see [12], where the wind is treated as an exogenous variable. However, we are interested in a continuous-time modelling framework. E.g. one could consider CARMA-X models or regime-switching models. Here we will follow the latter approach which is motivated by the work by [8] who incorporated forward-looking capacity constraints into a jump-diffusion model for electricity prices.

We introduce an exogenous regime-switching variable based on the forward-looking variable given by the predicted wind penetration index ρ , where

$$\rho(s) = \begin{cases} 1, & \text{if the predicted wind penetration at time } s \text{ is "high"} \\ 0, & \text{if the predicted wind penetration at time } s \text{ is "low"}. \end{cases} \quad (2)$$

The new spot price model is then given by $S(t) = \Lambda(t) + Y(t)$, where

$$Y(t) = \int_{-\infty}^t g(t-s)dM(s), \quad t \geq 0.$$

Here

$$dM(s) = \rho(s)dM^{(1)}(s) + (1 - \rho(s))dM^{(2)}(s),$$

where

$$dM^{(i)}(s) = a^{(i)}(s)ds + \sigma^{(i)}(s-)dL^{(i)}(s),$$

for independent Lévy processes $L^{(1)}$ and $L^{(2)}$. Also, $a^{(i)}$ denote suitable drift and $\sigma^{(i)}$ stochastic volatility processes, for $i \in \{1, 2\}$.

The key question which remains to be addressed is how exactly the regime-switching variable ρ should be chosen, given that the expression in (2) appears

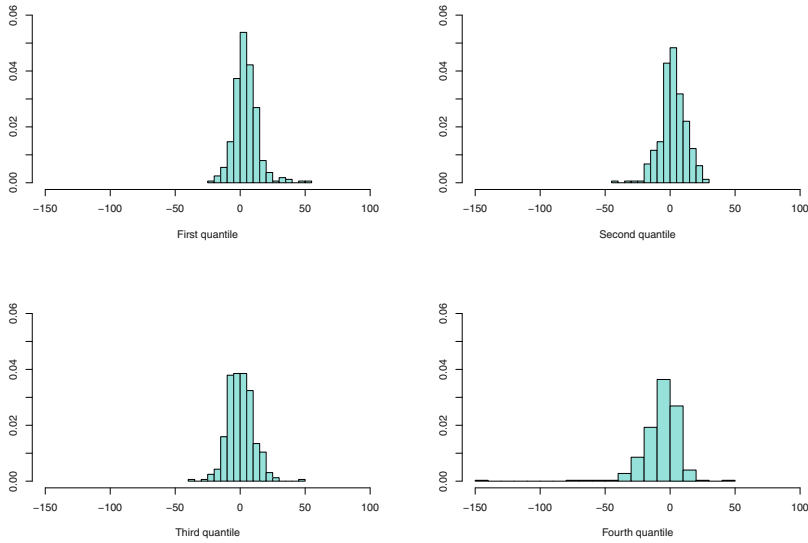


Fig. 10 Standardised histograms of the increments of M for different levels (associated with the four quartiles) of the wind penetration index

rather casual. In order to answer this question, we study the empirical properties of the recovered driving process M . More precisely, we investigate how the marginal distribution of the driving process of the CARMA process changes in relation to different levels of the wind penetration index. We split the sample of the recovered increments of M into four parts corresponding to the four quartiles of the wind penetration index, which are given in Table 4.

Their empirical distributions are described in form of standardised histograms, which describe the empirical probability density functions, in Fig. 10. Similarly to the finding in our exploratory data analysis, we observe that also the distribution of the increments of the driving process M changes quite remarkably for different levels of the wind penetration index. In particular, we observe that rather extreme negative increments are associated with a relatively high wind penetration index.

One can imagine a variety of rather sophisticated methods for choosing the cut-off point for our regime-switching variable. Here we are interested in a rather simple rule, which at the same time allows for a reasonable amount of observations in the high regime so that inference is still feasible and does not just rely on a very small number of observations. Hence, we choose the cut-off point to be 26.4% as a hard threshold, meaning that all increments of M associated with a wind penetration index in the fourth quartile belong to the high regime.

3.4 Model for M Based on the Generalised Hyperbolic Distribution

Motivated by the empirical study in [1], we will fit the class of *generalised hyperbolic* (GH) distributions to the increments of M in the two regimes.

Our notation for the GH distribution follows the one used in [13]. See also [4] for more details on the implementation of the corresponding estimation procedures in R available through the `ghyp` package.

Let us denote by $d, k \in \mathbb{N}$ some constants and let \mathbf{X} denote a k -dimensional random vector. Recall that we say that the law of \mathbf{X} is given by the *multivariate generalised hyperbolic* (GH) distribution if

$$\mathbf{X} \stackrel{\text{law}}{=} \boldsymbol{\mu} + W\boldsymbol{\gamma} + \sqrt{W}\mathbf{C}\mathbf{Z},$$

where $\mathbf{Z} \sim N(\mathbf{0}, I_k)$, $\mathbf{C} \in \mathbb{R}^{d \times k}$, $\boldsymbol{\mu}, \boldsymbol{\gamma} \in \mathbb{R}^d$. Here $W \geq 0$ denotes a one-dimensional random variable, independent of \mathbf{Z} and with Generalised Inverse Gaussian (GIG) distribution, i.e. $W \sim GIG(\lambda, \chi, \psi)$. The density of the GIG distribution with parameters (λ, χ, ψ) has the following functional form:

$$f_{GIG}(x) = \left(\frac{\psi}{\chi}\right)^{\frac{\lambda}{2}} \frac{x^{\lambda-1}}{2K_{\lambda}(\sqrt{\chi\psi})} \exp\left(-\frac{1}{2}\left(\frac{\chi}{x} + \psi x\right)\right),$$

where K_{λ} denotes the modified Bessel function of the third kind, and the parameters have to satisfy one of the following three restrictions

$$\chi > 0, \psi \geq 0, \lambda < 0, \quad \text{or} \quad \chi > 0, \psi > 0, \lambda = 0, \quad \text{or} \quad \chi \geq 0, \psi > 0, \lambda > 0.$$

The parameter $\boldsymbol{\mu}$ is called the location parameter, $\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}'$ is the dispersion matrix and $\boldsymbol{\gamma}$ is the symmetry or skewness parameter. The three (scalar) parameters λ, χ, ψ of the GIG distribution determine the shape of the GH distribution. The parametrisation described above is referred to as the so-called $(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ -parametrisation of the GH distribution. Since this parametrisation causes an identifiability problem when one tries to estimate the parameters, we will work with the so-called $(\lambda, \bar{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ -parametrisation in our empirical study. One can show that the $(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ -parametrisation can be obtained from the $(\lambda, \bar{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ -parametrisation by setting

$$\psi = \bar{\alpha} \frac{K_{\lambda+1}(\bar{\alpha})}{K_{\lambda}(\bar{\alpha})}, \quad \chi = \frac{\bar{\alpha}^2}{\psi} = \bar{\alpha} \frac{K_{\lambda}(\bar{\alpha})}{K_{\lambda+1}(\bar{\alpha})},$$

and $\lambda, \boldsymbol{\Sigma}, \boldsymbol{\gamma}$ remain the same, cf. [4].

We estimated 11 distributions within the GH class—consisting of the asymmetric and symmetric versions of the GH, hyperbolic, Student’s t, variance gamma, normal inverse Gaussian distribution and the Gaussian distribution and compared them

Table 7 Parameter estimates for the Student’s t-distribution in the low regime (symmetric case) and high regime (asymmetric case)

	$\hat{\nu}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\gamma}$
Low regime	6.70	2.26	9.77	0
High regime	4.76	0.57	11.31	-7.68

Note that the parameter $\bar{\alpha} = 0$ in the case of the Student’s t-distribution

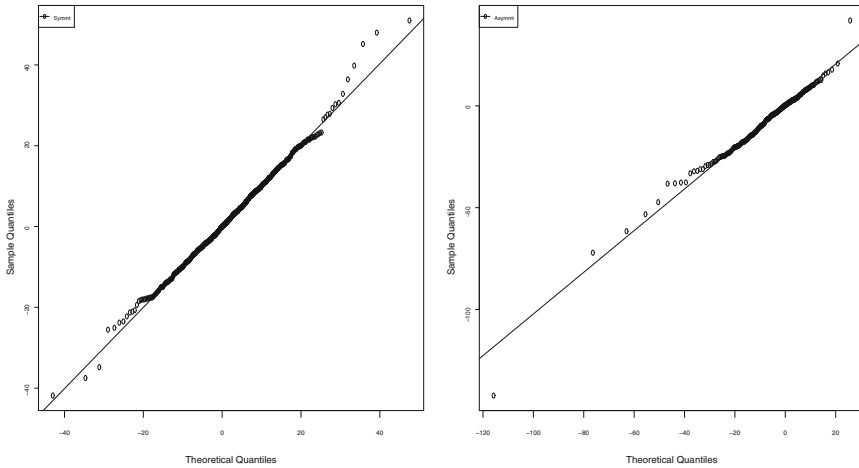


Fig. 11 Diagnostic quantile-quantile plots: The first picture compares the empirical quantiles of the data in the *low* regime with the estimated symmetric Student’s t-quantiles, and the second picture compares the empirical quantiles of the data in the *high* regime with the estimated asymmetric Student’s t-quantiles

according to the Akaike information criterion. We found that the best model for the low regime is given by the symmetric Student’s t-distribution and for the high regime by the asymmetric Student’s t-distribution, see Table 7 for the corresponding parameter estimates.

When comparing the parameter estimates for the low and the high regime provided in Table 7, we observe that the skewness, fatness of the tails and the volatility increases for the high regime and that the mean parameter decreases compared with the low regime. This is in line with previous findings in the literature, which suggest that the price level typically decreases with increasing wind energy production and that the volatility and the risk for negative spikes (represented through negative skewness and fatter tails) typically increases.

Also, we have provided quantile-quantile plots to assess the goodness-of-fit of the Student’s t-distribution in Fig. 11, which overall look reasonable.

The estimation results suggest that a good model for the driving process M in the regime-switching LSS specification is given by

$$dM(s) = \rho(s)dM^{(1)}(s) + (1 - \rho(s))dM^{(2)}(s),$$

where

$$dM^{(1)}(s) = \left(\mu^{(1)} + \gamma \left(\sigma^{(1)}(s) \right)^2 \right) ds + \sigma^{(1)}(s)dW^{(1)}(s),$$

$$dM^{(2)}(s) = \mu^{(2)} ds + \sigma^{(2)}(s)dW^{(2)}(s),$$

for independent Brownian motions $W^{(1)}$ and $W^{(2)}$. Here the stochastic volatility processes $\sigma^{(i)}$ are chosen as Ornstein-Uhlenbeck processes with inverse Gamma marginal distribution, since a mean-variance mixture with the inverse Gamma distribution results in the Student's t-distribution.

Note that the reason for choosing volatility modulated Brownian motions rather than Lévy processes with Student's t-distribution is that we found a significant short term (2 lags) autocorrelation in the increments of the recovered process M suggesting that a stochastic volatility model is more suitable than a pure jump model. This finding reveals that stochastic volatility is a key feature in energy markets, but it typically only exhibits short memory. Stochastic volatility is naturally incorporated into the LSS framework making it a convincing modelling tool for energy markets.

4 Conclusion

This paper has presented an extension of the modelling framework based on Lévy semistationary (LSS) processes introduced by [1]. Since forward-looking information in terms of weather forecasts is available to market participants, the corresponding predictions for the day-ahead wind production can be derived and used when determining day-ahead electricity spot prices. We incorporated this information through the so-called predicted wind penetration index in a regime-switching model based on LSS processes. We have observed that the flexibility offered through the regime switching component allows to model electricity prices in a more refined way than it was possible in the original (reduced form) LSS modelling framework. In particular, we have found that a relatively high wind penetration index leads to a lower mean level, higher skewness, fatter tails and increased volatility in the distribution of the electricity prices. This confirms earlier findings in the literature and for the first time links them to a flexible continuous-time stochastic modelling framework.

Given the increasing importance of renewable sources of energy, it will be interesting to extend the current investigation to include a wider range of renewables, including solar and biogas which up to now do not play as big of a role as wind power generation in the European energy market.

Another area for future research would be to develop a stochastic model for the wind penetration index, which could either result in a regime-switching model with a stochastic switching parameter or in a joint model for electricity prices and the

wind generation index. A preliminary analysis along those lines has revealed that a reasonable model for ρ needs to take both yearly seasonality and clusters into account. This could be seen as a first step for constructing models which can be used for mid-term forecasts of electricity prices influenced by renewables and would help to find modelling and inference tools for reliable risk management in energy markets.

Acknowledgments A.E.D. Veraart acknowledges financial support by a Marie Curie FP7 Integration Grant within the 7th European Union Framework Programme.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Barndorff-Nielsen, O.E., Benth, F.E., Veraart, A.E.D.: Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. *Bernoulli* **19**(3), 803–845 (2013)
2. Benth, F.E., Meyer-Brandis, T.: The information premium for non-storable commodities. *J. Energy Markets* **2**(3), 111–140 (2009)
3. Benth, F.E., Šaltytė Benth, J., Koekebakker, S.: Stochastic modelling of electricity and related markets. In: *Advanced Series on Statistical Science and Applied Probability*, vol. 11. World Scientific (2008)
4. Breymann, W., Lüthi, D.: *ghyp: a package on generalized hyperbolic distributions. Manual for the R package ghyp* (2010)
5. Brockwell, P.: Representations of continuous-time ARMA processes. *J. Appl. Probab.* **41**(A), 375–382 (2004)
6. Brockwell, P.J., Davis, R.A., Yang, Y.: Estimation for non-negative Lévy-driven CARMA processes. *J. Bus. Econ. Statist.* **29**(2), 250–259 (2011)
7. Brockwell, P.J., Schlemm, E.: Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations. *J. Multivar. Anal.* **115**, 217–251 (2013)
8. Cartea, Á., Figueroa, M.G., Geman, H.: Modelling electricity prices with forward looking capacity constraints. *Appl. Math. Finance* **16**(1–2), 103–122 (2009)
9. Forrest, S., MacGill, I.: Assessing the impact of wind generation on wholesale prices and generator dispatch in the Australian national electricity market. *Energy Policy* **59**, 120–132 (2013)
10. Jónsson, T., Pinson, P., Madsen, H.: On the market impact of wind energy forecasts. *Energy Econ.* **32**(2), 313–320 (2010)
11. Jónsson, T., Pinson, P., Nielsen, H., Madsen, H., Nielsen, T.: Forecasting electricity spot prices accounting for wind power predictions. *Sustain. Energy, IEEE Trans.* **4**(1), 210–218 (2013). Jan
12. Ketterer, J.C.: The impact of wind power generation on the electricity price in Germany. *Energy Econ.* **44**, 270–280 (2014)
13. McNeil, A.J., Frey, R., Embrechts, P.: *Quantitative risk management: concepts, techniques and tools. Princeton Series in Finance.* Princeton University Press, Princeton, NJ (2005)
14. Nowotarski, J., Tomczyk, J., Weron, R.: Robust estimation and forecasting of the long-term seasonal component of electricity spot prices. *Energy Econ.* **39**, 13–27 (2013)
15. Taylor, J.W., McSharry, P.E.: Short-term load forecasting methods: an evaluation based on European data. *IEEE Trans. Power Syst.* **22**(4), 2213–2219 (2007)

16. Veraart, A.E.D., Veraart, L.A.M.: Modelling electricity day-ahead prices by multivariate Lévy semi-stationary processes. In: Benth, F.E., Kholodnyi, V., Laurence, P. (eds.) *Wolfgang Pauli Proceedings*, pp. 157–188. Springer (2014)
17. Weron. R.: *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley, Chichester (2006). With CD-ROM with MFE Toolbox
18. Würzburg, K., Labandeira, X., Linares, P.: Renewable generation and electricity prices: taking stock and new evidence for Germany and Austria. *Energy Econ.* **40**, S159–S171 (2013). Supplement Issue: Fifth Atlantic Workshop in Energy and Environmental Economics