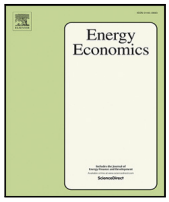
Energy Economics 104 (2021) 105640

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneeco



A multi-factor approach to modelling the impact of wind energy on electricity spot prices

Paulina A. Rowińska [a], Almut E.D. Veraart [a,*], Pierre Gruet [b]

[a] Imperial College London, United Kingdom
[b] EDF Lab Paris-Saclay and FiME, Laboratoire de Finance des Marchés de l'Energie, France

ARTICLE INFO

JEL classification:
C0
C1
C3
C5
Q4

Keywords:
CARMA model
Electricity spot prices
Electricity futures prices
Lévy process
Lévy semistationary process
Wind energy

ABSTRACT

We introduce a four-factor arithmetic model for electricity baseload spot prices in Germany and Austria. The model consists of a deterministic seasonality and trend function, both short- and long-term stochastic components, and exogenous factors such as the daily wind energy production forecasts, the residual demand and the wind penetration index. We describe the short-term stochastic factor by a Lévy semi-stationary (LSS) process, and the long-term component is modelled as a Lévy process with increments belonging to the class of generalised hyperbolic distributions.

We derive the corresponding futures prices and develop an inference methodology for our multi-factor model. The methodology allows to infer the various factors in a step-wise procedure taking empirical spot prices, futures prices and wind energy production and total load data into account.

Our empirical work shows that taking into account the impact of the wind energy generation on the prices improves the goodness of fit. Moreover, we demonstrate that the class of LSS processes can be used for modelling the exogenous variables including wind energy production, residual demand and the wind penetration index.

1. Introduction

One of the main challenges of the 21st century is reinforcing sustainable economic growth in order to tackle climate change. An important part of this task is a more effective use of renewable energy sources, such as the wind power. From the economic point of view, these sources are notorious for being risky to invest in because of their unpredictable influence on the electricity prices. This is due to their high dependence on the weather — and weather forecasts still do not reach the desirable level of accuracy.

In this article, we aim to develop a new model for electricity spot prices which takes the impact of renewable sources of energy, and wind energy in particular, into account. Electricity is considered a commodity with unique characteristics that make the use of standard tools of financial mathematics difficult or even impossible. First, energy storage to date is either limited or expensive, so supply and demand must match at almost all times. Over time the prices tend to a long-term average determined by this balance, so they exhibit a mean reversion. Any disturbances of this equilibrium can result in significant spikes in the electricity spot (day-ahead) market, which leads to a strong and heteroscedastic (time-varying) volatility of electricity spot prices. Furthermore, electricity prices are seasonal. The demand is much higher in winter months (due to the need of heating and longer use of lights) as well as during hot summer months (due to the use of air conditioning). Since prices are elastic functions of demand, very cold or very warm weather usually results in more expensive electricity. The periodic behaviour can also be observed at a smaller, weekly scale, namely the demand is higher in the peak time, i.e. Monday to Friday between 8 am and 8 pm, when people need electricity for their activities at work and home.

In the literature, one can find a variety of electricity price models: some start from a model for the spot prices and then derive the corresponding futures prices (eg. Carmona et al. (2013), Cartea et al. (2009) and Benth et al. (2014)), others model the futures prices directly (eg. Benth and Paraschiv (2018), Barndorff-Nielsen et al. (2015, 2013) and Borovkova and Geman (2006)). For a comprehensive recent survey on the electricity price models, we refer to Deschatre et al. (2021).

Recently, modelling approaches have been developed that try to take into account the increasingly important role of renewable sources of energy. For instance, Ketterer (2014) develops an ARMA-GARCH model for German electricity prices where wind power features as an exogenous factor both in the conditional mean and the conditional variance model. Elberg and Hagspiel (2015) develops a copula

* Corresponding author.
E-mail addresses: p.rowinska15@imperial.ac.uk (P.A. Rowińska), a.veraart@imperial.ac.uk (A.E.D. Veraart), pierre.gruet@edf.fr (P. Gruet).

https://doi.org/10.1016/j.eneco.2021.105640
Received 26 January 2018; Received in revised form 8 July 2021; Accepted 1 October 2021
Available online 16 October 2021

model for the spatial dependence structure of wind power in Germany; Veraart (2016) proposes to model German electricity prices by regime-switching Lévy semi-stationary processes, where the regimes are determined based on the level of the wind penetration index. Also, Deschatre and Veraart (2018) investigate the impact of wind energy production on the spikes in the spot prices.

Multi-factor models have also become popular recently: Benth et al. (2014) propose an arithmetic model for spot prices with three factors: a deterministic seasonality and trend function as well as short- and long-term stochastic parts. Lingohr and Müller (2021) extend this work by introducing both a dependence on residual demand in the model and by using conditionally independent increment processes as driving processes for the short-term price variability.

This article extends the work by Benth et al. (2014) and complements the work by Lingohr and Müller (2021) as follows: We propose to model electricity prices by an arithmetic model consisting of four components: a seasonality and trend term, exogenous variables based on wind energy production forecasts, a short-term stochastic factor and a long-term stochastic factor. We propose to model the short-term stochastic factor by a Lévy semi-stationary (*LSS*) process (rather than a CARMA(2,1) process as in Benth et al. (2014) or a generalised Ornstein–Uhlenbeck process driven by a conditionally independent increment process as in Lingohr and Müller (2021)). Also, rather than using an α-stable driving process as in Benth et al. (2014), we choose processes from the generalised hyperbolic class, which not only capture the features of electricity prices but also are easy to fit thanks to the algorithm implemented by Lüthi and Breymann (2016). We note that while we are using maximum-likelihood and GMM methods for parameter estimation, Müller and Seibert (2019) propose an alternative estimation procedure for CARMA process on Markov chain Monte Carlo methods.

Moreover, we allow for exogenous variables in order to take into account the relationship between electricity spot prices and wind energy production forecasts. We stress that it is the wind energy production forecasts, not the actual values, that impact the price, as they represent the information available to market participants at the time of transaction (Ketterer, 2014). It is crucial that we focus on day-ahead forecasts, because longer forecasting horizons reduce their accuracy and usefulness for making market decisions. We will denote the day-ahead forecasts of wind energy production by WD.

Jónsson et al. (2010) argue that the same level of wind energy production will affect the price in a different way depending on the total demand. In the literature we find two ways of combining these two variables: the wind penetration index, suggested for example by Jónsson et al. (2010), as well as the residual load (or demand) studied by Nicolosi and Fürsch (2009) and others. In our empirical work, we consider both time series and compare the goodness-of-fit of the resulting models.

The remainder of this article is structured as follows. In Section 2, we introduce a four-factor arithmetic model for spot prices, and we derive a formula for futures prices in Section 3. In Section 4 we describe how to fit the model to empirical data and study all model terms in detail. Our empirical results are presented in Section 5, and Section 6 gives an outlook on the use of Lévy stationary processes for modelling the exogenous variables. We summarise our key findings in Section 7 and relegate the proofs of our theoretical results to Appendix. Additional theoretical background material and empirical results are available in the supplementary material (Rowińska et al., 2021). This articles contains material from the PhD thesis (Rowińska, 2020).

2. The multi-factor model

Benth et al. (2014) argue that electricity spot prices can be well described by a three-factor model consisting of (1) a seasonality and trend component, (2) a long-term factor given by a Lévy process and (3) a short-term stationary factor given by a stable CARMA process. We extend their model by replacing the model for the short-term factor by a linear combination of a Lévy semi-stationary process and exogenous variables. In doing so, we find that we can further improve the model fit.

2.1. Model definition

We denote by $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\in\mathbb{R}}, \mathbb{P})$ a filtered probability space satisfying the usual conditions of right-continuity and completeness. Let $S = (S(t))_{t\geq 0}$ denote the electricity spot price process. Suppose that $\mathbf{x} = (\mathbf{x}(t))_{t\geq 0}$ denotes an M-dimensional vector of exogenous (deterministic) variables for $M \in \mathbb{N}$. We denote by $\bar{\mathbf{x}} = (\bar{\mathbf{x}}(t))_{t\geq 0}$ their detrended and deseasonalised counter-parts, see Section 5 in Rowińska et al. (2021) for details. We propose the multi-factor model with exogenous variables given by

$$S(t) = \Lambda(t) + Z(t) + Y(t) + \mathbf{a}^\top \bar{\mathbf{x}}(t), \quad t \geq 0, \tag{2.1}$$

where $\Lambda = (\Lambda(t))_{t\geq 0}$ denotes a deterministic seasonality and trend function, $Z = (Z(t))_{t\geq 0}$ is a Lévy process with zero mean (under $\mathbb{P}$), which describes the non-stationary long-term and low-frequency stochastic dynamics of the spot price. The (stationary) short-term dynamics of the price process are modelled as a sum of a stationary process $Y = (Y(t))_{t\geq 0}$ and a linear combination of exogenous variables, where $\mathbf{a} \in \mathbb{R}^M$ is a parameter vector. We shall assume that Z and Y are independent.

2.2. Modelling the short term factor by a Lévy semi-stationary process

Our main innovation compared to the model described in Benth et al. (2014) is two-fold: first, to model the short-term factor by a Lévy semi-stationary process, see Barndorff-Nielsen et al. (2013), rather than by an α-stable CARMA process (which is in fact a special case of our more general model), and, second, to include exogenous variables. Suppose that $Y = (Y(t))_{t\geq 0}$ is a Lévy semi-stationary (LSS) process without drift denoted by

$$Y(t) = \int_{-\infty}^{t} g(t-s)\sigma(s-)dL(s). \tag{2.2}$$

Here $g : (0,\infty) \to \mathbb{R}$ denotes a deterministic function with $\lim_{t\to\infty} g(t) = 0$. Furthermore, $L = (L(t))_{t\in\mathbb{R}}$ is a two-sided Lévy process with characteristic triplet given by (d, b, l_L) (with a truncation function $h(z) = \mathbb{1}_{\{x:|x|\leq 1\}}(z)$), where d denotes the drift, b the variance of the Gaussian component and l_L the Lévy measure. The volatility process $\sigma = (\sigma(t))_{t\in\mathbb{R}}$ is assumed to be càdlàg, non-negative, strictly stationary and independent of L. Let us define $\phi_t(s) := g(t-s)\sigma(s-)$. According to Barndorff-Nielsen et al. (2013), the process $(\phi_t(s))_{s\leq t}$ is integrable with respect to L if and only if $(\phi_t(s))_{s\leq t}$ is $\{\mathcal{F}(t)\}_{t\in\mathbb{R}}$-predictable and the following three conditions hold almost surely:

(1) $b\int_{-\infty}^{t} (\phi_t(s))^2 ds < \infty$,

(2) $\int_{-\infty}^{t}\int_{-\infty}^{\infty} (1 \wedge |\phi_t(s)z|^2) l_L(dz)ds < \infty$,

(3) $\int_{-\infty}^{t} \left| d\phi_t(s) + \int_{-\infty}^{\infty} (h(z\phi_t(s)) - \phi_t(s)h(z)) l_L(dz) \right| < \infty$.

To ensure the square integrability, we assume that L and σ have finite second moments and replace condition (1) by $\int_{-\infty}^{t} \mathbb{E}_{\mathbb{P}}[\phi_t(s)^2] ds = \int_{-\infty}^{t} g(t-s)^2 \mathbb{E}_{\mathbb{P}}[\sigma(s)^2] ds < \infty$ and $\mathbb{E}_{\mathbb{P}}[(g(t-s)\sigma(s)ds)^2] < \infty$. For the latter condition it is enough to ensure that for some $a \in (0,1)$ $\int_0^\infty g^{2a}(x)dx < \infty$ and $\int_{-\infty}^{t} g^{2(1-a)}(t-s)\mathbb{E}_{\mathbb{P}}[\sigma(s)^2] < \infty$.

Note that Y is stationary if and only if $\sigma(t)$ and the increments of $L(t)$ are jointly stationary. While our empirical studies (Section 5) do not indicate much stochastic volatility in our data, we develop the theoretical framework which can accommodate σ with a short memory. We suggest using a (stationary) Ornstein–Uhlenbeck process:

Assumption 2.1. The stochastic volatility process $\sigma = (\sigma(t))_{t\in\mathbb{R}}$ in Eq. (2.2) satisfies $\sigma^2(t) = \int_{-\infty}^{t} e^{-\delta(t-s)} dV(s)$ where $\delta > 0$ and $V = (V(t))_{t\in\mathbb{R}}$ denotes a two-sided Lévy subordinator independent of the Lévy process L.

3. Pricing futures in the multi-factor model

We will now describe how electricity futures contracts can be priced in our new multi-factor model. Here we will follow the approach described in Benth et al. (2014), but we note that alternative approaches for pricing futures have been mentioned in the literature; see Weron (2008) for a detailed discussion.

We will need to find a new probability measure $\mathbb{Q}$ which is equivalent to the physical measure $\mathbb{P}$. Since not all the assets in electricity markets are tradable, the resulting discounted price dynamics do not need to be (local) martingales with respect to the equivalent measure. Hence we can choose any probability measure equivalent to $\mathbb{P}$.

3.1. Specification of an equivalent probability measure

Let $T^* > 0$ denote a finite time horizon and let $\theta(\cdot)$ denote a real-valued function which is integrable w.r.t. L on $[0, T^*]$. As suggested by Barndorff-Nielsen et al. (2013), we define the equivalent martingale measure $\mathbb{Q}_L^\theta$ via the generalised Esscher transform of L as follows

$$\left.\frac{d\mathbb{Q}_L^\theta}{d\mathbb{P}}\right|_{\mathcal{F}(t)} = \exp\left(\int_0^t \theta(s)dL(s) - \int_0^t \phi_L(\theta(s))\,ds\right), \quad t \in [0, T^*],$$

where $\phi_L(x) = \log(\mathbb{E}[\exp(xL(1))])$ denotes the log-moment generating function of $L(1)$ (if it exists for suitable $x \in \mathbb{R}$). By analogy, let $\eta(\cdot)$ denote a real-valued function which is integrable w.r.t. V on $[0, T^*]$. Then we define $\mathbb{Q}_V^\eta$ by

$$\left.\frac{d\mathbb{Q}_V^\eta}{d\mathbb{P}}\right|_{\mathcal{F}(t)} = \exp\left(\int_0^t \eta(s)dV(s) - \int_0^t \phi_V(\eta(s))\,ds\right), \quad t \in [0, T^*], \tag{3.1}$$

where $\phi_V(x) = \log(\mathbb{E}[\exp(xV(1))])$ denotes the log-moment generating function of $V(1)$ (if it exists for suitable $x \in \mathbb{R}$).

Similarly, let $\kappa(\cdot)$ denote a real-valued function which is integrable w.r.t. Z on $[0, T^*]$. Then we define a measure change for Z as

$$\left.\frac{d\mathbb{Q}_Z^\kappa}{d\mathbb{P}}\right|_{\mathcal{F}(t)} = \exp\left(\int_0^t \kappa(s)dZ(s) - \int_0^t \phi_Z(\kappa(s))\,ds\right), \quad t \in [0, T^*]. \tag{3.2}$$

We will then define the equivalent measure as $\mathbb{Q} = \mathbb{Q}^{\theta,\eta,\kappa} = \mathbb{Q}_L \times \mathbb{Q}_V \times \mathbb{Q}_Z$.

If we choose the function $\theta(\cdot), \eta(\cdot)$ and $\kappa(\cdot)$ to be constant, then the change of measure will preserve the desirable Lévy properties and the independence between the processes.

In order to guarantee the existence of the Esscher transforms above, we need the existence of the corresponding exponential moments of L, V, Z, see Barndorff-Nielsen et al. (2018, p. 347) for a detailed discussion.

3.2. Pricing futures contracts

Suppose that $0 \le t \le T \le T^*$ and consider a futures contract $f(t,T)$ where T denotes the instantaneous delivery time. Using standard arguments, see Duffie (1992), we set $f(t,T) = \mathbb{E}_{\mathbb{Q}}[S(T)|\mathcal{F}(t)]$ – assuming throughout that $Z(T), Y(T)$ and hence $S(T)$ have finite $\mathbb{Q}$-expectation. Then

$$\begin{aligned} f(t,T) &= \mathbb{E}_{\mathbb{Q}}[S(T)|\mathcal{F}(t)] = \mathbb{E}_{\mathbb{Q}}[\Lambda(T) + Z(T) + Y(T) \\ &\quad + \mathbf{a}^\top \bar{\mathbf{x}}(T)|\mathcal{F}(t)] \\ &= \Lambda(T) + Z(t) + (T-t)\mathbb{E}_{\mathbb{Q}}[Z(1)] + \mathbb{E}_{\mathbb{Q}}[Y(T)|\mathcal{F}(t)] \\ &\quad + \mathbf{a}^\top \bar{\mathbf{x}}(T), \qquad \text{where} \end{aligned}$$

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[Y(T)|\mathcal{F}(t)] &= \int_{\infty}^t g(T-s)\sigma(s-)dL(s) + \mathbb{E}_{\mathbb{Q}}[L(1)]\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}} \\ &\quad \times [\sigma(s)|\mathcal{F}(t)]\,ds. \end{aligned}$$

Due to its nonstorability, electricity is delivered over a time period rather than at one specific moment. Thus for all $0 \le t \le T_1 < T_2 \le T^*$ we define the price of a futures contract with a delivery period $[T_1, T_2]$ by

$$F(t,T_1,T_2) := \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{T_2-T_1}\int_{T_1}^{T_2} S(T)dT \middle| \mathcal{F}_t\right]. \tag{3.3}$$

If following Benth et al. (2014) we define *time to maturity* as $u := u(t,T_1,T_2) = \frac{1}{2}(T_1+T_2) - t$, then Eq. (3.3) becomes

$$\begin{aligned} F(t,T_1,T_2) &= \frac{1}{T_2-T_1}\int_{T_1}^{T_2} \Lambda(T)dT + Z(t) + u\mathbb{E}_{\mathbb{Q}}[Z(1)] \\ &\quad + G(t,T_1,T_2) + \frac{1}{T_2-T_1}\int_{T_1}^{T_2} \mathbf{a}^\top \bar{\mathbf{x}}(T)dT, \end{aligned}$$

where we define

$$\begin{aligned} G(t,T_1,T_2) &:= \frac{1}{T_2-T_1}\int_{T_1}^{T_2} \mathbb{E}_{\mathbb{Q}}[Y(T)|\mathcal{F}(t)]dT \\ &= \frac{1}{T_2-T_1}\left(\int_{T_1}^{T_2}\int_{-\infty}^t g(T-s)\sigma(s-)\,dL(s)\,dT \right. \\ &\quad \left. + \mathbb{E}_{\mathbb{Q}}[L(1)]\int_{T_1}^{T_2}\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\,dT\right). \end{aligned}$$

We define the deseasonalised futures price as

$$\begin{aligned} \tilde{F}(t,T_1,T_2) &:= F(t,T_1,T_2) - \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\Lambda(T)dT \\ &= Z(t) + u\mathbb{E}_{\mathbb{Q}}[Z(1)] + G(t,T_1,T_2) + \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\mathbf{a}^\top\bar{\mathbf{x}}(T)dT. \end{aligned}$$

We want to show that the impact of Y on the (deseasonalised) futures prices as described by $G(t,T_1,T_2)$ is negligible when t is much smaller than T_1. Hence we characterise the asymptotic behaviour of G next.

Assumption 3.1. Suppose that Assumption 2.1 holds with $\lim_{x\to\infty}\int_0^x g(y)e^{-\frac{\delta}{2}(x-y)}dy = 0$.

Proposition 3.1. *Consider the multi-factor model and assume that either σ is constant or Assumption 3.1 holds. Then for $\Delta > 0$ and fixed $t > 0$,*

$$\begin{aligned} &\lim_{T_1\to\infty} G(t,T_1,T_1+\Delta) = \mathbb{E}_{\mathbb{Q}}[Y(0)], \text{ with } \mathbb{E}_{\mathbb{Q}}[Y(0)] \\ &= \mathbb{E}_{\mathbb{Q}}[L(1)]\,\mathbb{E}_{\mathbb{Q}}[\sigma(0)]\int_0^\infty g(y)dy, \end{aligned}$$

where the limit is in the L^2–sense w.r.t. $\mathbb{Q}$.

The proof of Proposition 3.1 is given in Appendix.

Hence, with $u = \frac{1}{2}(T_1+T_2) - t$ and for $t \ll T_1$, we can approximate the deseasonalised futures price by

$$\tilde{F}(t,T_1,T_2) \approx Z(t) + u\beta_1 + \beta_0, \tag{3.4}$$

where $\beta_0 = \mathbb{E}_{\mathbb{Q}}[Y(0)] + \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\mathbf{a}^\top\bar{\mathbf{x}}(T)dT$, $\beta_1 = \mathbb{E}_{\mathbb{Q}}[Z(1)]$.

Also,

$$\tilde{F}(t,T_1,T_2) - \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\mathbf{a}^\top\bar{\mathbf{x}}(T)dT \approx Z(t) + u\beta_1 + \beta_0^*, \tag{3.5}$$

where $\beta_0^* = \mathbb{E}_{\mathbb{Q}}[Y(0)]$, $\beta_1 = \mathbb{E}_{\mathbb{Q}}[Z(1)]$.

These approximations will prove useful for inference, see Section 4.

4. Inference in the multi-factor model

Let us now turn towards inference in the multi-factor model. Our new model consists of (up to) four factors which we will estimate in four separate steps. In doing so, we modify and extend the algorithm proposed by Benth et al. (2014), pp. 398–9, to cope with our more general model setting.

4.1. Step 1: Dealing with seasonality and trend

First, we formulate a parametric model for the seasonality and trend function denoted by Λ. We shall use a linear combination of a first order polynomial and a cosine function with yearly frequency and dummies for the day of the week and holidays. The corresponding parameters are then estimated by least squares. This results in an estimate $\widehat{\Lambda}$.

We proceed in a similar way to deseasonalise any exogenous variables $\mathbf{x}$ and we denote the deseasonalised and detrended variables by $\overline{\mathbf{x}}$, see Section 5 in Rowińska et al. (2021) for details. Finally, we compute the deseasonalised and detrended spot price denoted by $\overline{S}$ as $\overline{S}(t) = S(t) - \widehat{\Lambda}(t) = Z(t) + Y(t) + \mathbf{a}^\top \overline{\mathbf{x}}(t)$. In our notation we disregard the estimation error to simplify the exposition.

4.2. Step 2: Filtering out the non-stationary long-term factor

The second step consists of filtering out the long-term factor Z. Here we modify the algorithm by Benth et al. (2014, pp. 398–9). Their key idea was to use futures prices to split spot prices into the long and short-term factors. In Proposition 3.1 we showed that in the long end deseasonalised futures prices are (stochastically) influenced only by the long-term factor Z.

Hence, we first need to determine a threshold $u^\star$ such that deseasonalised prices of contracts with $u = u(t, T_1, T_2) = \frac{1}{2}(T_1 + T_2) - t \geq u^\star$ are described approximately by Z alone. Benth et al. (2014) recommend carrying out the estimation for various choices of the threshold $u^\star$ until an optimal level has been found.

Let us now denote the sample mean by $\widehat{\mathbb{E}}_{\mathbb{P}}[\cdot]$. For $u \geq u^\star$, using Proposition 3.1 and $\mathbb{E}_{\mathbb{P}}[Z(t)] = 0$, we can approximate

$$\mu_{\tilde{F}}(u) := \widehat{\mathbb{E}}_{\mathbb{P}}\left[\tilde{F}(t, T_1, T_2)\right] \approx u\beta_1 + \beta_0, \tag{4.1}$$

and by linear regression find the estimates $\widehat{\beta}_0$ and $\widehat{\beta}_1$.

Given these parameter estimates, we can recover a realisation of the long-term process by setting

$$\begin{aligned}\widehat{Z}(t) &= \widehat{Z}\left(\frac{1}{2}(T_1 + T_2) - u\right) \\ &= \frac{1}{\mathrm{card} U(t, u^\star)} \sum_{(u, T_1, T_2) \in U(t, u^\star)} \left[\tilde{F}(t, T_1, T_2) - u\widehat{\beta}_1 - \widehat{\beta}_0\right], \end{aligned} \tag{4.2}$$

where $U(t, u^\star) := \{(u, T_1, T_2) \in \mathbb{R}^3 : u \geq u^\star \text{ and } \exists\ F(t, T_1, T_2) : \frac{1}{2}(T_1 + T_2) - t = u\}$.

Remark 4.1. In the absence of exogenous variables, we have that $\beta_0 = \mathbb{E}_{\mathbb{Q}}(Y(0))$ which does neither depend on T_1, T_2 nor on t and the Eqs. (3.4) and (3.5) are identical. However, in the presence of exogenous variables, the parameter β_0 is technically still a function of T_1 and T_2, but not of t. In the following, we have ignored this dependence in the estimation. In future work, one could investigate an alternative approach: one could estimate $\mathbf{a}$ first, then approximate the left hand side of Eq. (3.5) and run the corresponding regression and recover Z via an adjusted version of Eq. (4.2).

Since futures contracts are traded only from Monday to Friday, $\widehat{Z}$ does not include estimates for the weekends. Following Benth et al. (2014), we suggest extrapolating by setting $\widehat{Z}$ on Saturdays and Sundays equal to the preceding Friday's value.

Finally, we should check whether the recovered process $\widehat{Z}$ can be well modelled by the suggested Lévy process.

4.3. Step 3: Estimating the parameters associated with the exogenous variables

Since $Y(t) + \mathbf{a}^\top \mathbf{x}(t) = S(t) - \Lambda(t) - Z(t)$, we obtain the estimated short-term factor by

$$\widehat{(Y(t) + \mathbf{a}^\top \overline{\mathbf{x}}(t))} = \overline{S}(t) - \widehat{Z}(t).$$

We then estimate $\mathbf{a}$ by linear regression, resulting in the estimate $\widehat{\mathbf{a}}$. Now we can recover the estimated stationary process Y by setting

$$\widehat{Y}(t) = \widehat{(Y(t) + \mathbf{a}^\top \overline{\mathbf{x}}(t))} - \widehat{\mathbf{a}}^\top \overline{\mathbf{x}}(t).$$

4.4. Step 4: Estimating the stationary short-term factor

We can now fit a model to the recovered LSS process $\widehat{Y}$. In order to simplify the notation, we will suppress the hat and just write Y for $\widehat{Y}$. Recall that $Y(t) = \int_{-\infty}^t g(t-s)\,dM(s)$, for $dM(s) = \sigma(s-)dL(s)$.

4.4.1. Estimating a CARMA process

A very popular choice for the LSS process Y is a continuous-time autoregressive moving average (CARMA) process, see Section 3 in Rowińska et al. (2021) for details. An LSS becomes a (volatility modulated) CARMA(p,q) process with $0 \leq q < p$ and parameters $a_1, \ldots, a_p, b_1, \ldots, b_q$, if the kernel function is given by $g(x) = \boldsymbol{b}^T e^{\boldsymbol{A}t} \boldsymbol{e} \mathbb{1}_{[0,\infty)}(x)$ with

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{pmatrix}, \boldsymbol{e} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix}.$$

As in Brockwell et al. (2011) all eigenvalues of $\boldsymbol{A}$ are assumed to have negative real parts.

The parameters of a CARMA process can be estimated using the algorithm described by García et al. (2011). In particular, Brockwell et al. (2011) showed that for a fixed sampling interval the mean corrected sampled CARMA(p,q) process is a weak ARMA(p,p-1) process, so we can estimate ARMA(p,p-1) parameters and map them to the continuous setting as outlined in Subsection 3.2 in Rowińska et al. (2021).

What is more, the algorithm also enables us to recover the increments of the driving process of the CARMA process M which, in the general LSS case, satisfies $dM(t) = \sigma(t-)dL(t)$.

What type of process is M? Is stochastic volatility present? Suppose we have estimated the CARMA parameters and have recovered the increments of $M = (M(t))_{t \geq 0}$. We should first check whether M could simply be modelled by a Lévy process. If this is not the case, e.g. since we found significant auto-correlation in the squared increments of M, we could consider the choice $M(t) := \int_0^t \sigma(s-)dB(s)$ for a standard Brownian motion B. We know that, in this case, $[M](t) = \int_0^t \sigma^2(s)ds$. Let N denote the number of observations, then we denote by $\Delta_n^h M = M(nh) - M((n-1)h)$ for $n = 1, \ldots, N$ the recovered increments of the driving process over intervals of length $h > 0$. Let $C(t) := \int_0^t \sigma(s)^2 ds$ and $\Delta_n^h C := C(nh) - C((n-1)h) = \int_{(n-1)h}^{nh} \sigma(s)^2 ds$. We note that $\Delta_n^h M$ is mixed Gaussian with zero mean and conditional variance given by $\Delta_n^h C$.

Suppose that the volatility process satisfies Assumption 3.1. Then, we find that, for any $t, s \in \mathbb{R}$,

$$\kappa_1 := \mathbb{E}_{\mathbb{Q}}\left[\sigma(t)^2\right] = \mathbb{E}_{\mathbb{Q}}[V(1)] \int_0^\infty e^{-\delta x} dx = \frac{\mathbb{E}_{\mathbb{Q}}[V(1)]}{\delta};$$

$$\kappa_2 := \mathrm{Var}_{\mathbb{Q}}\left[\sigma(t)^2\right] = \mathrm{Var}_{\mathbb{Q}}[V(1)] \int_0^\infty e^{-2\delta x} dx = \frac{\mathrm{Var}_{\mathbb{Q}}[V(1)]}{2\delta};$$

$$\begin{aligned}\mathrm{Cov}_{\mathbb{Q}}\left(\sigma(t+s)^2, \sigma(t)^2\right) &= \mathrm{Var}_{\mathbb{Q}}[V(1)] \int_0^\infty e^{-\delta(x+s)} e^{-\delta x} dx \\ &= \frac{\mathrm{Var}_{\mathbb{Q}}[V(1)]}{2\delta} e^{-\delta s} = \kappa_2 e^{-\delta|s|}.\end{aligned}$$

Hence $\mathbb{E}_{\mathbb{Q}}\left[\Delta_n^h C\right] = \kappa_1 h$ and, for $k \in 0, \ldots, N-n$, we have

$$\mathrm{Cov}_{\mathbb{Q}}\left(\Delta_{n+k}^h C, \Delta_n^h C\right) = \frac{\kappa_2}{\delta^2} e^{-\delta h k}\left(e^{\delta h} + e^{-\delta h} - 2\right);$$

$$\mathrm{Cor}_{\mathbb{Q}}\left(\Delta_{n+k}^h C, \Delta_n^h C\right) = e^{-\delta k h}.$$

Consider the case when $L = B$ is a Brownian motion, and suppose we want to infer the parameter $\delta > 0$. Then we approximate the empirical autocorrelation function of the squared increments $(\Delta_n^h M)^2$ by $\mathrm{Cor}_{\mathbb{Q}}\left(\Delta_{n+k}^h C, \Delta_n^h C\right)$ and estimate δ by least squares.

4.4.2. General LSS processes

Other parametric choices for the LSS process could be considered. E.g. one popular class is a BSS process where g takes the form of the so-called gamma kernel given by

$$g(x) = \frac{\lambda^{\nu-0.5}}{\Gamma(2\nu-1)^{0.5}} x^{\nu-1}\exp(-0.5\lambda x), \text{ for } x>0, \nu>0.5,$$

and σ is chosen such that Y has generalised hyperbolic marginal law, see Section 1.9 in Barndorff-Nielsen et al. (2018).

In the absence of high-frequent observations, the estimation theory for such processes, however, is not yet very advanced, and least squares and (generalised) methods of moments have mainly been used so far, see Barndorff-Nielsen et al. (2013) and Chapter 10 in Barndorff-Nielsen et al. (2018). Alternatively, a regression-based approach based on the log-variogram (Bennedsen, 2017) seems to be a promising estimation method as well.

5. Empirical results

We analyse daily baseload electricity prices (i.e. averages of hourly spot prices) and monthly (from one month up to six months ahead) baseload futures prices for Germany and Austria, traded between 1 January 2011 and 31 December 2015. The corresponding time series are depicted in Fig. 1.

In addition, we consider the daily time series of day-ahead forecasts of daily wind energy production forecast (WD) and of daily load forecast (LD), the daily wind penetration index (WPI) computed as the ratio WD/LD and the residual demand (RD) computed as $LD - WD$, depicted in Fig. 2.

Remark 5.1. When we write $WD(t)$ and $LD(t)$, we refer to the wind energy production forecast and the load forecast, respectively, for day t, which is available on day $t-1$.

A detailed data description and exploratory study of the data is provided in Section 4 in the supplementary material (Rowińska et al., 2021).

5.1. Seasonality and trend

First of all, we de-seasonalise and de-trend all time series using linear combinations of (first order) polynomials, trigonometric functions and dummy variables. The details of this part of the data analysis is described in Section 5 in the supplementary material (Rowińska et al., 2021).

5.2. Long-term component

Next, we need to find an optimal threshold $u^\star$ such that deseasonalised forward contracts with a *time to maturity* of $u = \frac{1}{2}(T_1+T_2)-t \geq u^\star$ are no longer stochastically influenced by Y. In our empirical analysis we found that the estimation results were not very sensitive to the threshold choice and we will in the following report the results when we fix $u^* = 16$ days.

We run the linear regression in Eq. (4.1), see Fig. 3a, and use the estimates $\widehat{\beta}_0$ and $\widehat{\beta}_1$ in Eq. (4.2) to recover Z. The recovered process $\widehat{Z}$ and its increments are depicted in Figs. 3b and 3c, respectively.

The process $\widehat{Z}$ appears to be non-stationary, therefore we study its increments, plotted in Fig. 3c. At first glance $\widehat{Z}$ has stationary increments, as confirmed at 0.01 significance level by the augmented Dickey–Fuller (ADF) test, computed using the function **adf.test** from the **R** package **tseries** (Trapletti and Hornik, 2018). We observe that the $\widehat{Z}$ has (mostly) uncorrelated increments, as shown in Fig. 4a, and the autocorrelation of the squared increments, depicted in Fig. 4b, exhibits a significant autocorrelation for only a few days. Thus we model $\widehat{Z}$ by a Lévy process with increments following a suitable infinitely divisible distribution (although a model of the type $\widehat{Z}(t) = \int_0^t \omega(s-)dW(s)$ for a standard Brownian motion W and a stochastic volatility process ω might describe the second order structure slightly more precisely).

Inspired by Barndorff-Nielsen et al. (2013), we fit distributions from the class of generalised hyperbolic distributions, using the **R** package **ghyp** provided by Lüthi and Breymann (2016). We find that the symmetric normal inverse Gaussian distribution (with parameters $\overline{\alpha} = 0.236$, $\mu = -0.005$, $\sigma = 0.385$) fits the data well, as shown in Figs. 4c and 4d.

5.3. Short-term component

Finally, we estimate the short-term component. We fill first consider the two scenarios: one with and one without exogenous variables.

5.3.1. Case 1: No exogenous variables

Let us first consider the case when we do not have any exogenous variables, i.e. when $\mathbf{a} = \mathbf{0}$. Then we set $\widehat{Y}(t) = \overline{S}(t) - \widehat{Z}(t)$. Figs. 5a and 5b show the resulting process and its autocorrelation function, respectively. The augmented Dickey–Fuller (ADF) test (Trapletti and Hornik, 2018) allows us to reject the unit root hypothesis at significance level 0.01, suggesting $\widehat{Y}$ is stationary.

Inspired by Benth et al. (2014), we start by modelling $\widehat{Y}$ with a (possibly volatility modulated) CARMA(2,1) process, a special case of an Lévy semi-stationary processes, see Barndorff-Nielsen et al. (2018), pp. 31–32. A volatility modulated CARMA(2,1) process takes the form

$$Y(t) = \alpha_1 \int_{-\infty}^{t} e^{\lambda_1(t-u)} dM(u) + \alpha_2 \int_{-\infty}^{t} e^{\lambda_2(t-u)} dM(u), \quad t \in \mathbb{R}, \tag{5.1}$$

where $dM(u) = \sigma(u-)dL(u)$. The parameters are given by $\alpha_1, \alpha_2, \lambda_1, \lambda_2$, where λ_1, λ_2 are negative (in general, have negative real parts) and are not identical.

Remark 5.2. The CARMA(2,1) process without stochastic volatility (i.e. with $\sigma \equiv 1$) clearly satisfies the assumptions of Proposition 3.1. In the presence of stochastic volatility, in order for Assumption 3.1 in Proposition 3.1 to hold, we require that $\lambda_i \neq -\frac{\delta}{2}$ for $i = 1, 2$. Then $\lim_{x\to\infty} \int_0^x g(y) e^{-\frac{\delta}{2}(x-y)} dy = \lim_{x\to\infty} \sum_{i=1}^2 \alpha_i \int_0^x e^{\lambda_i y} e^{-\frac{\delta}{2}(x-y)} dy = \lim_{x\to\infty} \sum_{i=1}^2 \frac{\alpha_i}{\lambda_i + \frac{\delta}{2}} \left(e^{\lambda_i x} - e^{-\frac{\delta}{2}x} \right) = 0$.

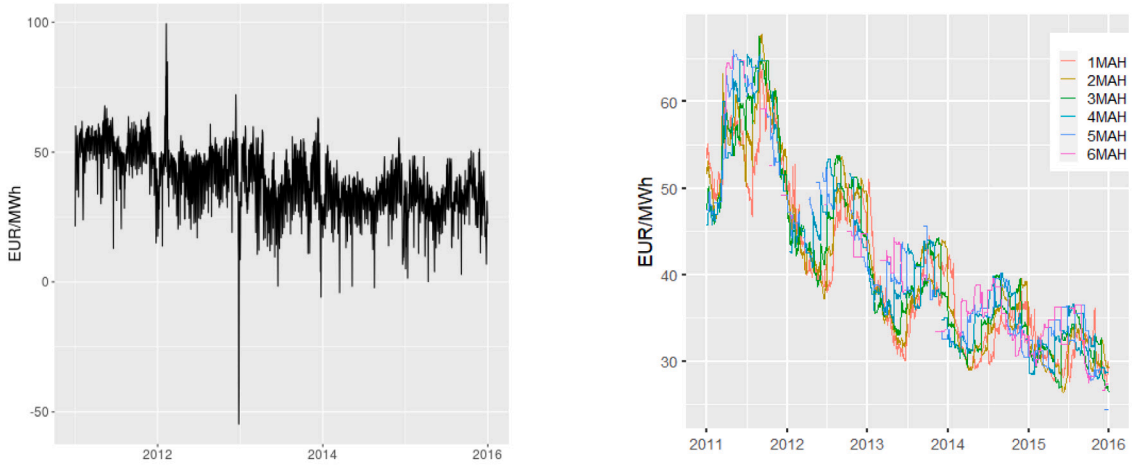
As described above, we implement the algorithm from García et al. (2011) to fit a CARMA(2,1) model to $\widehat{Y}$, see Subsection 3.2 in the supplementary material (Rowińska et al., 2021) for the details. First, we use the function **arima** from the **R** package **stats** (R Core Team, 2018) to fit an ARMA(2,1) process. We estimate the ARMA(2,1) model with parameters $(\phi_1, \phi_2, \theta) = (1.413, -0.446, -0.826)$. In Fig. 5c we can see that the estimated discrete model describes $\widehat{Y}$ accurately, although it cannot capture some extreme values. Next, we map the estimated ARMA(2,1) parameters into CARMA(2,1) kernel parameters, which gives us the process

$$(D^2 + 0.809D + 0.048)\widehat{Y}(t) = (0.194 + D)DM(t),$$

where D denotes "differentiation" with respect to t (for details see Section 3 in the supplementary material Rowińska et al., 2021). Next, we recover the driving noise M.

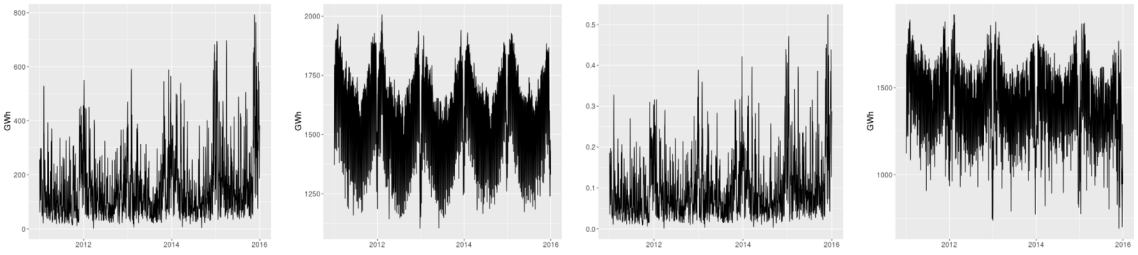
As a first modelling approach, we assume that $M \equiv Z$, i.e. that there is no stochastic volatility and M is given by a Lévy process. Then, as with the increments of Z, we choose the best fit to the increments ΔM of the driving noise among 11 generalised hyperbolic distributions. From Table 1 we learn that the best fit is given by the asymmetric Student's-t distribution. Below we describe the quality of these estimates, from now on called "true" estimates.

Remark 5.3. We note that we assessed the quality of the estimation for the CARMA process driven by a Lévy process in a simulation study, see Subsection 3.5 in Rowińska et al. (2021).



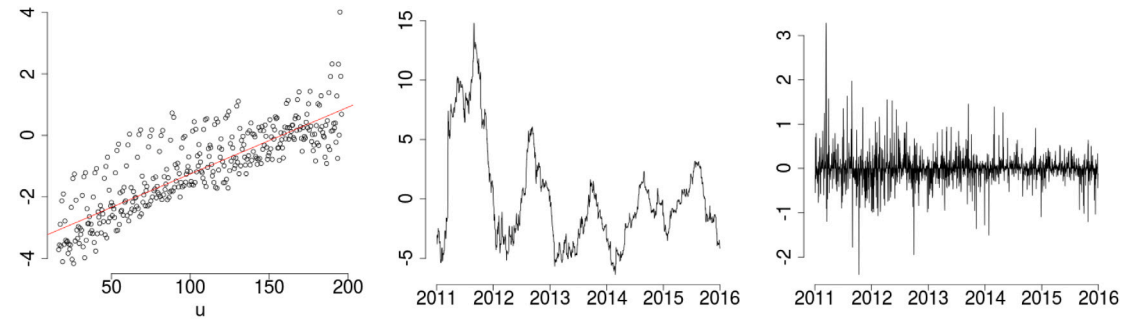
(a) (b)

Fig. 1. Time series of the daily electricity baseload prices and daily futures prices: (a) Time series plot of electricity spot prices S, (b) Time series of futures prices data: One month (1MAH) to six months (6MAH) ahead.



(a) Wind energy production forecast (WD) (b) Load forecast (LD) (c) Wind penetration index (WPI) (d) Residual demand (RD)

Fig. 2. Daily time series of day-ahead forecasts of (a) the wind energy production forecast (WD), (b) the load forecast (LD), (c) the wind penetration index (WPI), and (d) the residual demand (RD).



(a) (b) (c)

Fig. 3. The plots shows (a) $\mu_{\hat{F}}(u)$ with the fitted regression line, (b) $\hat{Z}$ and (c) the increments of $\hat{Z}$.

As a second modelling approach, we now consider the case when $dM(t) = \sigma(t-)dB(t)$ for a standard Brownian motion B. Figs. 6a and 6b present the autocorrelation functions of the increments and squared increments of the recovered driving noise process M. Fig. 6b indicates that the squared increments of M are not independent, so the model could benefit from including a short-memory stochastic volatility in the definition of Y. We notice the highest peak at lag 7, which could indicate that some traces of weekly behaviour remained after the deseasonalisation.

In a next step, we estimated the stochastic volatility model described in Section 4.4.1. We use the function **nls** from the **R** package **stats** (R Core Team, 2018), with the first six lags and the starting value equal zero. It gives us an estimate of $\hat{\delta} = 2.24$, at significance level 0.01. Fig. 6c presents the empirical and estimated autocorrelation functions.

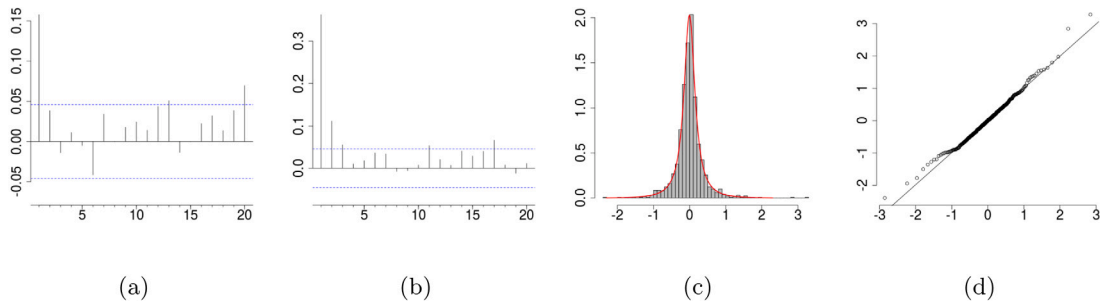


Fig. 4. The autocorrelation function of the (a) increments and of the (b) squared increments of the estimated Z. (c) Histogram and (d) q-q plot of the generalised hyperbolic distribution fitted to the increments of the estimated Z.

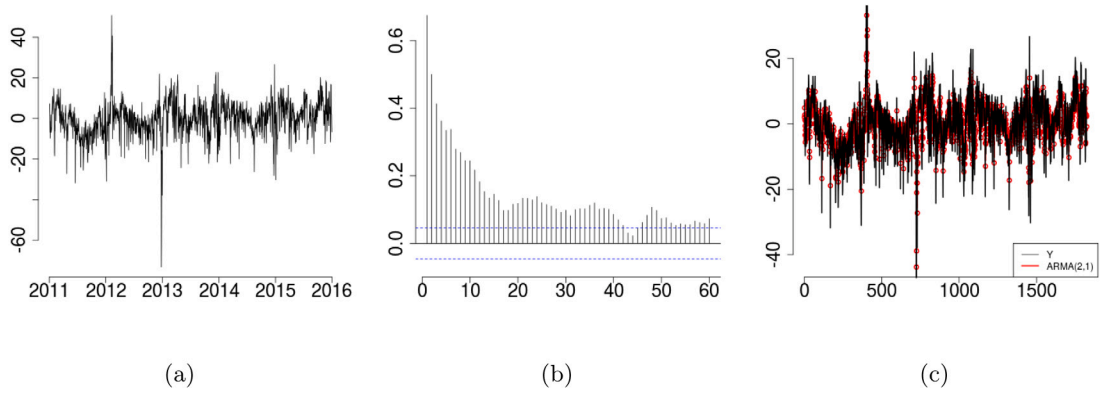


Fig. 5. (a) Time series plot and (b) autocorrelation function of the estimated Y. (c) The ARMA(2,1) model fitted to the estimated Y.

Table 1
Generalised hyperbolic distributions fitted to the recovered increments ΔM of the driving noise of $\hat{Y}$.

	Model	Symmetric	lambda	alpha.bar	mu	sigma	gamma	aic	llh
5	t	FALSE	−2.131	0.000	1.178	7.708	−1.215	12 403.016	−6197.508
1	ghyp	FALSE	−2.055	0.319	1.186	7.654	−1.191	12 405.297	−6197.649
10	t	TRUE	−2.044	0.000	0.293	7.854	0.000	12 412.389	−6203.194
3	NIG	FALSE	−0.500	0.962	1.310	7.606	−1.342	12 413.597	−6202.799
6	ghyp	TRUE	−2.047	0.005	0.293	7.850	0.000	12 414.389	−6203.194
8	NIG	TRUE	−0.500	0.908	0.306	7.724	0.000	12 425.469	−6209.734
2	hyp	FALSE	1.000	0.770	1.463	7.500	−1.498	12 429.264	−6210.632
4	VG	FALSE	1.528	0.000	1.444	7.517	−1.479	12 438.049	−6215.025
7	hyp	TRUE	1.000	0.724	0.286	7.604	0.000	12 442.264	−6218.132
9	VG	TRUE	1.503	0.000	0.285	7.614	0.000	12 451.107	−6222.553
11	gauss	TRUE		Inf	−0.034	7.949	0.000	12 741.822	−6368.911

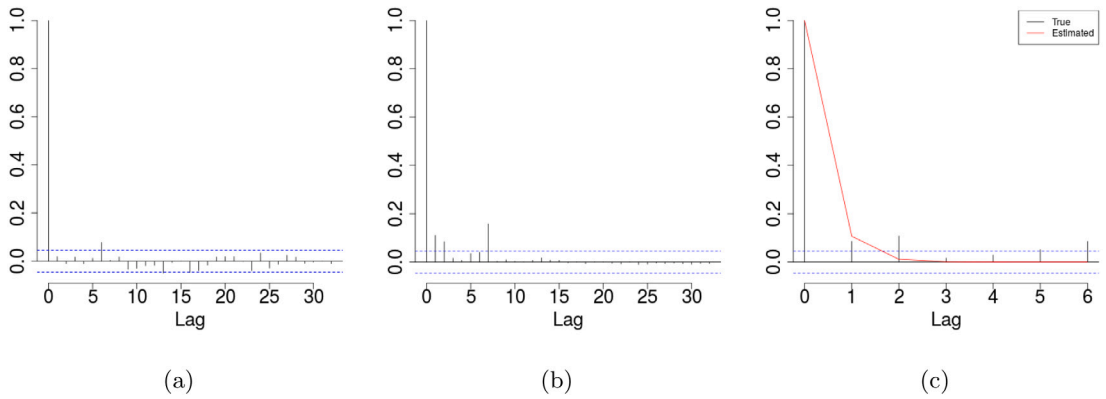


Fig. 6. (a) The autocorrelation function of the recovered increments of M. (b) The autocorrelation function of the squared recovered increments of M. (c) True and estimated autocorrelation functions of the squared recovered increments of M.

Table 2
Estimated coefficients of $\mathbf{a}^\top \bar{\mathbf{x}}(t) = a_1 + a_2 \cdot \overline{WD} + a_3 \cdot \overline{WPI} + a_4 \cdot \overline{RD} + a_5 \cdot \overline{WD}^2 + a_6 \cdot \overline{WPI}^2 + a_7 \cdot \overline{RD}^2$. If the estimation method (EM) is indicated as "r", then robust linear regression (function **rlm** in the **R** package **MASS**) was used, if it is indicated as "o", then ordinary linear regression (function **lm** in the **R** package **stats**) was used. With "x" we denote variables absent in a given model. Coefficients in bold are significant at level 0.05.

Model number	EM	Intercept	$\overline{WD}$	$\overline{WPI}$	$\overline{RD}$	$\overline{WD}^2$	$\overline{WPI}^2$	$\overline{RD}^2$
1		x	x	x	x	x	x	x
2	r	**−0.849**	0.0109	**−54.3**	**0.0267**	**5.28e−05**	−17.3	5.29e−06
2	o	**−0.854**	0.00704	**−45.5**	**0.0303**	**4.19e−05**	17.9	5.93e−06
3	r	−0.193	**0.0169**	**−53.9**	**0.0266**	x	x	x
3	o	−0.115	0.0106	**−38.6**	**0.0302**	x	x	x
4	r	−0.0913	**−0.0441**	x	x	x	x	x
4	o	−0.116	**−0.0451**	x	x	x	x	x
5	r	−0.114	x	**−73.4**	x	x	x	x
5	o	−0.116	x	**−74.7**	x	x	x	x
6	r	−0.191	x	x	**0.038**	x	x	x
6	o	−0.116	x	x	**0.0395**	x	x	x
7	r	**−0.806**	**−0.0514**	x	x	**6.11e−05**	x	x
7	o	**−0.912**	**−0.0542**	x	x	**6.38e−05**	x	x
8	r	**−0.777**	x	**−84.1**	x	x	**143**	x
8	o	**−0.839**	x	**−87.7**	x	x	**145**	x
9	r	**−0.652**	x	x	**0.0409**	x	x	**2.31e−05**
9	o	**−0.577**	x	x	**0.0429**	x	x	**2.17e−05**

We conclude that the volatility modulated Brownian motion-driven CARMA process is a valid alternative to a Lévy-driven CARMA process without stochastic volatility, as indicated by correlated squared increments of the driving process M as well as by the presence of a statistically significant Ornstein–Uhlenbeck memory parameter. However, while the evidence of stochastic volatility in our data set is relatively weak, we expect the volatility modulated Brownian motion-driven CARMA process to be the preferred choice in data sets with even more pronounced volatility.

5.3.2. Case 2: Exogenous variables based on wind energy production

Now we will include exogenous variables based on wind energy production in the multi-factor model. To this end, recall that we denote by WD the daily time series of day-ahead forecasts of daily wind energy production forecast, by LD the daily load forecast, by WPI the daily wind penetration index computed as the ratio WD/LD and by RD the residual demand computed as $LD - WD$, see Fig. 2. WPI is a dimensionless quantity with values between zero (no wind energy production) and one (all energy production coming from wind energy). Because of energy export, on rare occasions the value of WPI might exceed one. RD corresponds to energy from sources other than wind, expressed in GWh. In the discussed period (2011–2015), on average 10% of total energy produced in Austria and Germany came from wind and WPI appears to be very volatile and oscillates a lot around its mean level.

The wind related variables WD, WPI and RD have clear seasonal patterns which we removed by fitting suitable seasonality and trend functions via least squares and then subtracting the fitted functions from the observations, see Subsection 5.1 in Rowińska et al. (2021) for details. We start by regressing the recovered short-term factor $\overline{S}(t) - \widehat{Z}(t) = (Y(t) + \mathbf{a}^\top \bar{\mathbf{x}}(t))$ on combinations of these deseasonalised and detrended variables ($\overline{WD}$, $\overline{LD}$ and $\overline{WPI}$) and their squares. More precisely, we choose $M = 7$ and

$$\bar{\mathbf{x}}^\top(t) = (1, \overline{WD}(t), \overline{WPI}(t), \overline{RD}(t), \overline{WD}^2(t), \overline{WPI}^2(t), \overline{RD}^2(t)).$$

Remark 5.4. Elberg and Hagspiel (2015) suggest a similar approach, with some significant differences. Firstly, they only consider the residual demand RD, whereas we also include the wind energy generation WD and the wind penetration index WPI. Secondly, in order to capture the non-linear relationship between spot prices and the residual demand, they suggest using splines. However, as shown in Figure 20 in Rowińska et al. (2021), the relationships between deseasonalised wind variables and spot prices do not seem to require a polynomial level higher than quadratic. Moreover, linear regression results in a more parsimonious and intuitive model than splines, which is important for practitioners. Thirdly, instead of varying model coefficients with time like Elberg and Hagspiel (2015), we deseasonalise the data; again, the resulting model is easier to use in practice.

We are considering nine model specifications, denoted by Model 1–9, which are summarised in Table 2. First we consider models with one wind-related variable each (models 4, 5 and 6), to check which one helps explain electricity spot prices. We then extend them by adding square terms (models 7, 8 and 9) to allow for non-linear relationships. We also look into all wind variables together (model 3) as well as all wind variables in both linear and squared forms (model 2). These two models let us compare the significance of variables of interest directly. Since the relationships between explanatory variables are not linear, we do not expect any collinearity issues. However, we include these two models to check if the impact of some variables on the prices changes due to the presence of the others. In practice, we suggest using one of the remaining models to avoid overfitting. For that reason and to keep the interpretability we do not include any models with interaction terms between explanatory variables. Finally, we consider a model without any information about the wind energy production for comparison.

In Table 2, we report the regression coefficients for all nine proposed models, estimated with robust linear regression and ordinary least squares, respectively. In the remainder of the article we keep the coefficients estimated with robust regression, as this approach should reduce the bias towards outliers. The results of both estimation methods differ slightly, but almost only quantitatively: only $\overline{WD}$ in the model 3 is significant at level 0.05 when we use the robust regression, while insignificant with the standard linear regression.

The intercept becomes significant if we include squared variables. We are particularly interested in model 2, where we include all three variables ($\overline{WD}$, $\overline{WPI}$ and $\overline{RD}$) as well as their squares. The results indicate that all three wind-related variables help explain the short-term factor: $\overline{WPI}$ and $\overline{RD}$ in their basic form, while $\overline{WD}$ is squared. Since $\overline{WD}^2$ and $\overline{RD}^2$ take large values, the coefficients in front of them are very small. We also remark that there is a change in sign for $\overline{WPI}^2$ depending on the estimation method. However, the coefficient is insignificant and $\overline{WPI}^2$ takes on average very small values.

Table 3
Best and worst models according to different statistical divergencies (in sample).

Measure	Euclidean	Manhattan	Minkowski	Chebyshev
min distance	9	9	9	7
max distance	1	1	1	8

After having estimated $\hat{\mathbf{a}}$, we recover $\widehat{Y(t)} = (Y(t) + \widehat{\mathbf{a}^\top \mathbf{x}(t)}) - \widehat{\mathbf{a}(t)}^\top \overline{\mathbf{x}}(t)$ and estimate a Lévy-driven CARMA(2,1) process (without stochastic volatility).

5.3.3. Model comparisons

In this section we try to answer two questions: First, does the model fit for electricity spot prices improve when exogenous variables based on wind energy production are included? Second, which combination of wind energy variables gives the best results?

Since statistical significance gives us valuable information about the usefulness of different wind-related variables in explaining the electricity prices, we do not remove coefficients insignificant at level 0.05. This choice allows us to interpret the compared models in a meaningful way.

We compare the models in sample, i.e. on the data set the model was fitted to (2011–2015). As the difference between models lies only in the short-term factor $Y(t)+\hat{\mathbf{a}}^\top\overline{\mathbf{x}}(t)$, we focus on this part of the model. Using the parameters estimated in Section 5.3.2, for each model we simulate a path of Y and then compute $Y(t) + \hat{\mathbf{a}}^\top \overline{\mathbf{x}}(t)$.

In Fig. 7 we present the densities of the simulated and empirical $Y(t) + \hat{\mathbf{a}}^\top \overline{\mathbf{x}}(t)$. All models replicate the density relatively well. Additionally, for all models we measure the distances between the true and simulated densities with four statistical distances provided in the **R** package **philentropy** (Drost, 2018): (1) Euclidean with $d = \sqrt{(\sum |P_i - Q_i|^2)}$; (2) Manhattan with $d = \sum |P_i - Q_i|$; (3) Minkowski with $d = (\sum |P_i - Q_i|^p)^{\frac{1}{p}}$; (4) Chebyshev with $d = \max |P_i - Q_i|$. Table 3 presents the models whose densities are the closest and the furthest from the empirical $Y(t) + \hat{\mathbf{a}}^\top \overline{\mathbf{x}}(t)$. The model 1 (with no wind information) has the largest distance for three out of four distances out of the models considered. This confirms our intuition, since the model without any wind information should explain smaller parts of prices than other models. On the other hand, model 9 (with $\overline{RD}$ linear and squared) has the smallest distance for three out of four distances, which would confirm the insights of Elberg and Hagspiel (2015).

Similarly to Cartea et al. (2009), we also compare the true and simulated summary statistics. Precisely, we simulate 1000 paths of $Y(t)$ of the same length as the original data ($N = 1824$) and then compute $Y(t) + \hat{\mathbf{a}}^\top \overline{\mathbf{x}}(t)$. We compute the mean (m), variance (v), skewness (s) and kurtosis (k) of each path, and average them over all paths to obtain Monte Carlo estimates of summary statistics. In Table 4 we present squared differences between true and Monte Carlo statistics and the squared differences normalised by the value of the appropriate true statistics. Additionally, for all models we compute (normalised) Euclidean distances:

$$\sqrt{\left(\mathrm{m}^{\mathrm{true}} - \mathrm{m}^{\mathrm{MC}}\right)^2 + \left(\mathrm{v}^{\mathrm{true}} - \mathrm{v}^{\mathrm{MC}}\right)^2 + \left(\mathrm{s}^{\mathrm{true}} - \mathrm{s}^{\mathrm{MC}}\right)^2 + \left(\mathrm{k}^{\mathrm{true}} - \mathrm{k}^{\mathrm{MC}}\right)^2}$$

and

$$\sqrt{\left(1 - \frac{\mathrm{m}^{\mathrm{MC}}}{\mathrm{m}^{\mathrm{true}}}\right)^2 + \left(1 - \frac{\mathrm{v}^{\mathrm{MC}}}{\mathrm{v}^{\mathrm{true}}}\right)^2 + \left(1 - \frac{\mathrm{s}^{\mathrm{MC}}}{\mathrm{s}^{\mathrm{true}}}\right)^2 + \left(1 - \frac{\mathrm{k}^{\mathrm{MC}}}{\mathrm{k}^{\mathrm{true}}}\right)^2},$$

i.e. square roots of all columns of Table 4. From Tables 5 and 6 we learn that models 8 and 1 perform best according to these distances, while models 4 and 5 worst.

Additionally, we look at the squared differences between the first ten moments of $Y(t) + \hat{\mathbf{a}}^\top \overline{\mathbf{x}}(t)$: true $(\mu_1^{\mathrm{true}}, \ldots, \mu_{10}^{\mathrm{true}})$ and averaged over 1000 Monte Carlo simulations $(\mu_1^{\mathrm{MC}}, \ldots, \mu_{10}^{\mathrm{MC}})$. We also compute the same metric, but normalised by true moments. Thus for $i = 1, \ldots, 10$ we compute $\left(\mu_i^{\mathrm{true}} - \mu_i^{\mathrm{MC}}\right)^2$ and $\left(1 - \frac{\mu_i^{\mathrm{MC}}}{\mu_i^{\mathrm{true}}}\right)^2$, respectively. For each

Table 4
Squared differences (squared normalised differences) between true and averaged (over 1000 MC simulations) summary statistics of $Y(t) + \hat{\mathbf{a}}^\top \overline{\mathbf{x}}(t)$. Square brackets denote the minimum value for a given statistics, while underlined numbers represent the corresponding maximum value.

Model	Mean (m)	Variance (v)	Skewness (s)	Kurtosis (k)
1	0.00 (0.12)	14.39 (0.00)	0.09 (0.30)	17.85 (0.20)
2	0.09 (6.99)	32.45 (0.01)	0.13 (0.43)	39.41 (0.45)
3	0.02 (1.17)	25.62 (0.01)	**[0.02]** (**[0.06]**)	35.04 (0.40)
4	**[0.00]** (**[0.00]**)	**533.92** (**0.11**)	**0.32** (**1.05**)	56.07 (0.63)
5	**0.47** (**34.71**)	6.81 (0.00)	0.05 (0.15)	35.99 (0.41)
6	0.25 (18.40)	24.76 (0.01)	0.04 (0.12)	35.69 (0.40)
7	0.00 (0.02)	279.02 (0.06)	0.29 (0.96)	**77.63** (**0.88**)
8	0.01 (0.58)	**[0.09]** (**[0.00]**)	0.06 (0.20)	**[14.95]** (**[0.17]**)
9	0.36 (26.94)	43.09 (0.01)	0.17 (0.57)	40.06 (0.45)

Table 5
Euclidean distances between summary statistics of $Y(t)+\hat{\mathbf{a}}^\top\overline{\mathbf{x}}(t)$: true and averaged over 1000 Monte Carlo simulations. Square brackets denote the minimum value for a given statistics, while underlined numbers represent the corresponding maximum value.

Model	1	2	3	4	5	6	7	8	9
Distance	5.69	8.49	7.79	**24.30**	6.58	7.79	18.89	**[3.89]**	9.15

Table 6
Normalised Euclidean distances between summary statistics of $Y(t) + \hat{\mathbf{a}}^\top \overline{\mathbf{x}}(t)$: true and averaged over 1000 Monte Carlo simulations. Square brackets denote the minimum value for a given statistics, while underlined numbers represent the corresponding maximum value.

Model	1	2	3	4	5	6	7	8	9
Distance	**[0.79]**	2.80	1.28	1.34	**5.94**	4.35	1.38	0.97	5.29

Table 7
Best and worst models: models minimising and maximising (normalised) squared distances between the moments of empirical and simulated data (1000 simulations). Models are described in Table 2.

Moment	Best models		Worst models	
	Non-normalised	Normalised	Non-normalised	Normalised
1	4	4	5	5
2	8	8	4	4
3	3	3	4	4
4	5	5	4	4
5	3	3	4	4
6	5	5	4	4
7	3	3	4	4
8	5	5	4	4
9	5	5	4	4
10	5	5	4	4

distance and each number of moments we define the best model as the model minimising the distance and the worst as the one maximising the distance. We present the best and worst models in Table 7 (differences between moments) and Table 8 (sums of differences between moments). Clearly model 4 ($\overline{WD}$) has the worst performance. The best-performing models are 3 (all variables without squares) and 5 ($\overline{WPI}$).

6. Outlook: A model for the exogenous factors

In Section 5.3.3 we showed that wind-related variables ($\overline{WD}$, $\overline{RD}$ and $\overline{WPI}$) help modelling electricity spot prices. So far, we have treated such variables as exogenous. In this section, we will give an outlook on how these three variables could be modelled.

While wind *speed* has been widely modelled using the Weibull distribution, see e.g. Weisser (2003), we will show that the class of generalised hyperbolic distributions is suitable for describing wind energy production data well.

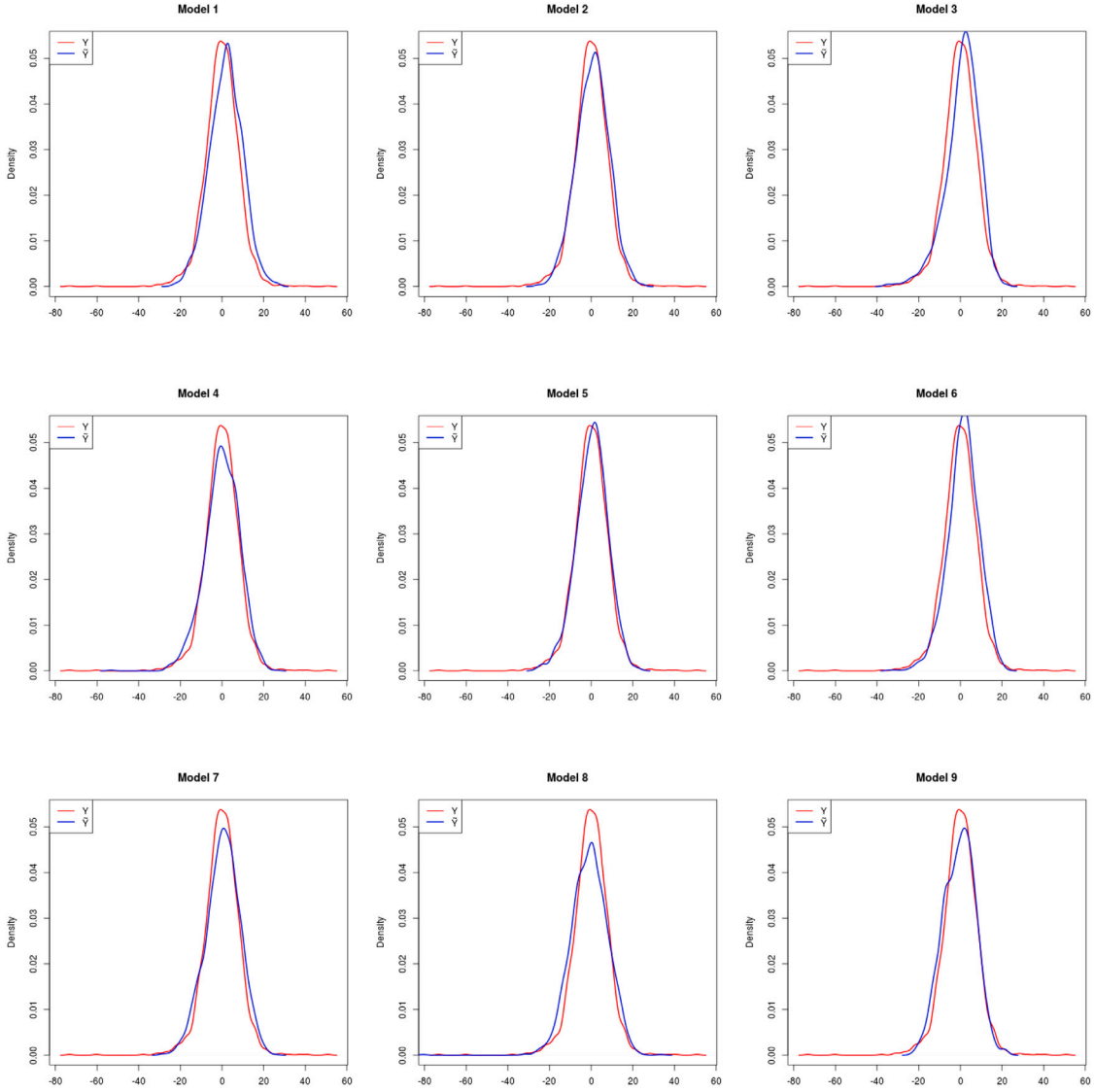


Fig. 7. Densities of the true and simulated short-term factors $Y(t) + \hat{\mathbf{a}}^{\top}\bar{\mathbf{x}}(t)$ for nine different model variations: in sample.

Table 8
Best and worst models: models minimising and maximising the sum of (normalised) squared distances between the moments of empirical and simulated data (1000 simulations), for different numbers of moments. Models are described in Table 2.

Highest moment	Best models		Worst models	
	Non-normalised	Normalised	Non-normalised	Normalised
1	4	4	5	5
2	8	4	5	5
3	8	8	4	5
4	5	8	4	4
5	3	3	4	4
6	5	3	4	4
7	3	3	4	4
8	5	3	4	4
9	3	3	4	4
10	5	3	4	4

6.1. Brownian semi-stationary processes with gamma kernels

Inspired by Barndorff-Nielsen et al. (2013), we propose to model wind-related variables by a Brownian semi-stationary (BSS) process of the form

$$X(t) = \mu + c\int_{-\infty}^{t} g(t-s)\omega(s)dB(s) + \gamma\int_{-\infty}^{t} q(t-s)\omega^2(s)ds, \tag{6.1}$$

where μ, c and γ are real constants, ω is a stationary volatility process and B denotes a standard Brownian motion independent of ω.

We will now fully specify the BSS process such that its marginal distribution is given by the class of generalised hyperbolic distributions, see Section 2 in Rowińska et al. (2021). To this end, we assume that the kernel functions g and q are of gamma-type. I.e. we denote the gamma density with parameters $\bar{\nu} > 0$ and $\bar{\lambda} > 0$ by

$$\bar{g}\left(t;\bar{\nu},\bar{\lambda}\right) = \frac{\bar{\lambda}^{\bar{\nu}}}{\Gamma(\bar{\nu})}t^{\bar{\nu}-1}e^{-\bar{\lambda}t},$$

and define

$$g(t) = \frac{\bar{\lambda}^{\bar{\nu}-\frac{1}{2}}}{\sqrt{\Gamma(2\bar{\nu}-1)}}t^{\bar{\nu}-1}\exp\left(-\frac{\bar{\lambda}}{2}t\right),$$

and

$$q(t) = g^2(t) = \bar{g}\left(t;2\bar{\nu}-1,\bar{\lambda}\right).$$

To ensure the existence of Eq. (6.1), we assume that $\frac{1}{2} < \bar{\nu} < 1$. We also let

$$\omega^2(t) = \int_{-\infty}^{t} p(t-s)dU(s),$$

for

$$p(t) = \frac{1}{\bar{\lambda}} \bar{g}\left(t; 2-2\bar{\nu}, \bar{\lambda}\right),$$

and a subordinator $U(t)$, which we will specify later. We observe that

$$\begin{aligned} X(t)|\omega \sim & N\left(\mu + \gamma \int_{-\infty}^{t} q(t-s)\omega^2(s)ds, c^2 \int_{-\infty}^{t} g^2(t-s)\omega^2(s)ds\right) \\ & = N\left(\mu + \gamma\sigma^2(t), c^2\sigma^2(t)\right), \end{aligned}$$

with

$$\sigma^2(t) = \int_{-\infty}^{t} e^{-\bar{\lambda}(t-s)}dU(s).$$

We can compute $\mathbb{E}\left[\sigma^2(0)\right] = \frac{\mathbb{E}[U(1)]}{\bar{\lambda}}$ and $\mathrm{Var}\left(\sigma^2(0)\right) = \frac{\mathrm{Var}(U(1))}{2\bar{\lambda}}$. Barndorff-Nielsen et al. (2013) showed that if U denotes a subordinator such that $\sigma^2 \sim GIG(\lambda, \chi, \psi)$, then $X \sim GH\left(\lambda, \chi, \psi, \mu, c^2, \gamma\right)$, so the marginal distribution of X uniquely determines the distribution of σ^2. So the above construction leads to a stationary stochastic process which has generalised hyperbolic marginal law.

Based on the explicit formula for the autocovariance of Brownian semi-stationary processes provided by Barndorff-Nielsen et al. (2018, p. 24), for $s > 0$, we get that

$$\begin{aligned} \mathrm{Cov}\left(X(t+s), X(t)\right) &= c^2\mathbb{E}\left[\omega^2(0)\right]\int_0^\infty g(x+s)g(x)dx \\ &+\gamma^2\int_0^\infty\int_0^\infty q(x+s)q(y)\mathrm{Cov}\left(\omega^2(|x-y|), \omega^2(0)\right)dxdy. \end{aligned}$$

Let us study each term separately. By Barndorff-Nielsen et al. (2013),

$$\begin{aligned} \mathbb{E}\left[\omega^2(0)\right] &= \mathbb{E}\left[U(1)\right]\int_0^\infty p(x)dx = \mathbb{E}\left[U(1)\right]\int_0^\infty \frac{1}{\bar{\lambda}}\frac{\bar{\lambda}^{2-2\bar{\nu}}}{\Gamma(2-2\bar{\nu})}x^{1-2\bar{\nu}}e^{-\bar{\lambda}x}dx \\ &= \frac{\mathbb{E}\left[U(1)\right]}{\bar{\lambda}} = \mathbb{E}\left[\sigma^2(0)\right], \end{aligned}$$

and

$$\begin{aligned} \mathrm{Cov}\left(\omega^2(t+s), \omega^2(t)\right) &= \mathrm{Var}\left(U(1)\right)\int_0^\infty p(x+s)p(x)dx \\ &= 2\bar{\lambda}\mathrm{Var}\left(\sigma^2(0)\right)\int_0^\infty p(x+s)p(x)dx. \end{aligned}$$

Furthermore,

$$\begin{aligned} \int_0^\infty g(x+s)g(x)dx &= \int_0^\infty \frac{\bar{\lambda}^{2\nu-1}}{\Gamma(2\bar{\nu}-1)}\left((x+s)x\right)^{\bar{\nu}-1}e^{-\frac{\bar{\lambda}}{2}(2x+s)}dx \\ &= \frac{1}{\Gamma(\bar{\nu}-\frac{1}{2})}2^{\frac{3}{2}-\bar{\nu}}\bar{K}_{\bar{\nu}-\frac{1}{2}}\left(\frac{\bar{\lambda}}{2}s\right), \end{aligned}$$

where $\bar{K}_\alpha(x) = x^\alpha K_\alpha(x)$ and $K_\nu(x)$ denotes the modified Bessel function of the third kind. Finally,

$$\begin{aligned} \int_0^\infty p(x+s)p(x)dx &= \int_0^\infty \left(\frac{1}{\bar{\lambda}}\frac{\bar{\lambda}^{2-2\bar{\nu}}}{\Gamma(2-2\bar{\nu})}\right)^2\left((x+s)x\right)^{1-2\bar{\nu}}e^{-\bar{\lambda}(2x+s)}dx \\ &= \frac{1}{\bar{\lambda}^3}\left(\frac{1}{\Gamma(2-2\bar{\nu})}\right)^2\frac{\Gamma(3-2\bar{\nu})}{\Gamma\left(\frac{1}{2}\right)}2^{2\bar{\nu}-\frac{5}{2}}\bar{K}_{\nu-\frac{1}{2}}\left(\bar{\lambda}s\right). \end{aligned}$$

Therefore

$$\mathrm{Cov}\left(X(t+s), X(t)\right) = P + Q,$$

with

$$P = c^2\mathbb{E}\left[\sigma^2(0)\right]\frac{2^{\frac{3}{2}-\bar{\nu}}}{\Gamma(\bar{\nu}-\frac{1}{2})}\bar{K}_{\bar{\nu}-\frac{1}{2}}\left(\frac{\bar{\lambda}}{2}s\right),$$

and

$$\begin{aligned} Q &= \gamma^2 2\bar{\lambda}\mathrm{Var}\left(\sigma^2(0)\right)\frac{1}{\bar{\lambda}^3}\left(\frac{1}{\Gamma(2-2\bar{\nu})}\right)^2\frac{\Gamma(3-2\bar{\nu})}{\Gamma\left(\frac{1}{2}\right)}2^{2\bar{\nu}-\frac{5}{2}}\frac{\bar{\lambda}^{4\bar{\nu}-2}}{\Gamma(2\bar{\nu}-1)^2} \\ &\int_0^\infty\int_0^\infty (x+s)^{2\bar{\nu}-2}e^{-\bar{\lambda}(x+s)}y^{2\bar{\nu}-2}e^{-\bar{\lambda}y}\bar{K}_{\bar{\nu}-\frac{1}{2}}\left(\bar{\lambda}|x-y|\right)dxdy \\ &= c_Q\int_0^\infty\int_0^\infty (x+s)^{2\bar{\nu}-2}e^{-\bar{\lambda}(x+s)}y^{2\bar{\nu}-2}e^{-\bar{\lambda}y}\bar{K}_{\bar{\nu}-\frac{1}{2}}\left(\bar{\lambda}|x-y|\right)dxdy, \end{aligned}$$

where

$$c_Q = \gamma^2\bar{\lambda}^{4\bar{\nu}-4}\mathrm{Var}\left(\sigma^2(0)\right)2^{2\bar{\nu}-\frac{3}{2}}\frac{\Gamma(3-2\bar{\nu})}{\Gamma\left(\frac{1}{2}\right)}\left(\frac{1}{\Gamma(2-2\bar{\nu})\Gamma(2\bar{\nu}-1)}\right)^2.$$

Similarly, we compute the variance of X:

$$\mathrm{Var}\left(X(0)\right) = P_0 + Q_0,$$

with

$$P_0 = c^2\mathbb{E}\left[\sigma^2(0)\right],$$

and

$$Q_0 = c_Q\int_0^\infty\int_0^\infty (xy)^{2\bar{\nu}-2}e^{-\bar{\lambda}(x+y)}\bar{K}_{\bar{\nu}-\frac{1}{2}}\left(\bar{\lambda}|x-y|\right)dxdy.$$

Now we are ready to estimate the parameters $\bar{\lambda}$ and $\bar{\nu}$ of the gamma kernel by matching the first lags of the empirical and theoretical autocorrelation functions, with the latter equal to

$$\mathrm{cor}\left(X(t), X(t+s)\right) = \frac{P+Q}{P_0+Q_0}.$$

6.2. Estimation procedure

We start the estimation procedure by using the function **stepAIC.ghyp** from the **R** package **ghyp** (Lüthi and Breymann, 2016) to fit 11 generalised hyperbolic distributions to X, where X denotes the wind variable of interest ($\overline{WD}$, $\overline{RD}$ or $\overline{WPI}$). Here we find the distribution which minimises the Akaike information criterion (AIC). Finally, we convert the parameters of the best fitting marginal generalised hyperbolic distribution, see Subsection 2.3 in the supplementary material (Rowińska et al., 2021), to the parametrisation $\left(\lambda, \chi, \psi, \mu, c = \sqrt{\Sigma}, \gamma\right)$.

After fitting the marginal distribution, we proceed to estimate kernel parameters $\bar{\nu}$ and $\bar{\lambda}$ by matching first lags of the empirical and theoretical autocorrelation functions of $X(t)$. Barndorff-Nielsen et al. (2013) suggest using the first $\left\lfloor\sqrt{N}\right\rfloor$ lags, where N denotes the number of observations. Therefore we apply the function **optim** from the **R** package **stats** (R Core Team, 2018) to minimise the squared difference between the first $\left\lfloor\sqrt{1824}\right\rfloor = 42$ lags of the empirical and theoretical autocorrelation functions. We set the initial parameters to $\bar{\nu} = 0.99$ and $\bar{\lambda} = 0.01$.

Remark 6.1. The optimisation procedure is not sensitive to reasonable starting values. For example, setting the initial values to $\bar{\nu} = 0.75$ and $\bar{\lambda} = 0.20$ results in the same optimal parameters up to the third decimal place. On the other hand, the number of lags used in the estimation procedure influences the results. In case of the residual demand, the estimation procedure performed on six lags, with the original initial values, resulted in $\bar{\nu} = 0.87$ and $\bar{\lambda} = 0.26$ instead of $\bar{\nu} = 0.82$ and $\bar{\lambda} = 0.20$. This should not surprise us considering the changing behaviour of the autocorrelation function. Therefore the choice of the number of lags must correspond to the application of interest.

6.3. Numerical results

In this subsection we summarise numerical results of the estimation procedure from Section 6.2 run on three wind-related variables. Detailed tables and diagnostic plots are available in the supplementary material Section 6 of Rowińska et al. (2021). Since we have constructed a BSS process with marginal GH law, we can use the Akaike Information criteria to find the best sub-class within the GH class for our three wind-related time series.

6.3.1. Wind energy production

According to the Akaike information criterion (AIC), the asymmetric normal inverse Gaussian provides the best fit to the marginal distribution of the wind energy production data (out of 11 considered subclasses of the generalised hyperbolic distribution). Therefore we assume that the marginal distribution of $\overline{WD}$ is asymmetric hyperbolic with

$$\left(\lambda, \chi, \psi, \mu, c = \sqrt{\Sigma}, \gamma\right) = (-0.50, 2.21, 2.21, -121.45, 8.79, 121.39)\,.$$

The value of the parameter γ is positive, which reflects the positive skewness of $\overline{WD}$.

For the kernel we obtain the parameter estimates $\bar{\nu} = 0.99$ and $\bar{\lambda} = 0.57$, which suggests that the BSS process considered here is not a semimartingale, see Barndorff-Nielsen et al. (2013).

6.3.2. Residual demand

The marginal distribution of $\overline{RD}$ is best described the asymmetric hyperbolic distribution with

$$\left(\lambda, \chi, \psi, \mu, c = \sqrt{\Sigma}, \gamma\right) = (1.00, 0.26, 2.53, 96.44, 11.12, -96.43)\,.$$

The kernel parameters are estimated as $\bar{\nu} = 0.82$ and $\bar{\lambda} = 0.20$, so again we are not dealing with a semimartingale.

6.3.3. Wind penetration index

Similarly to $\overline{WD}$, the deseasonalised wind penetration index is best described by the asymmetric normal inverse Gaussian distribution with

$$\left(\lambda, \chi, \psi, \mu, c = \sqrt{\Sigma}, \gamma\right) = (-0.50, 1.71, 1.71, -0.07, 0.22, 0.07)\,.$$

For the kernel we obtain the parameter estimates $\bar{\nu} = 0.88$ and $\bar{\lambda} = 0.39$, so once again we are outside the semimartingale setting.

7. Summary and outlook

In this article we introduced a four-factor arithmetic model of electricity spot prices, consisting of a deterministic seasonality and trend function, both a short- and a long-term stochastic component and exogenous variables. We modelled the long-term component as a Lévy process with increments belonging to the class of generalised hyperbolic distributions, and the short-term component as a Lévy semi-stationary processes. We included (combinations of) the wind energy production forecast, the residual demand and the wind penetration index as exogenous variables. We showed how futures prices can be derived for our spot price model and carried out an empirical study of electricity prices from the German and Austrian electricity market. In our empirical study, we worked with the discretised counter-parts of our continuous-time models. We refer to Phillips (1974), Sargan (1974) and Hamerle et al. (1993) for details on how discretely observed exogenous variables can be incorporated in a feasible way in continuous-time models.

We showed that even the basic model (without wind energy production data or stochastic volatility) already provides a good fit to the data. We compared this benchmark model with alternatives including wind energy generation. Each suggested model has different benefits and drawbacks, so the choice should depend on the application of interest. For example, the model including residual demand in the linear and squared form (model 9) has the smallest distance between the true and fitted densities according to three out of four tested metrics out of the models we have considered. On the other hand, the model without the information about wind energy production (model 1) attains the largest values of the same distances. If the quantity of interest is the distances between empirical and fitted summary statistics, we suggest using the model with wind penetration index in the linear and squared form (model 8). If one is interested in the goodness of fit of higher moments, the model with wind penetration index (model 5) or the model with all wind variables (model 3) would both be good options.

Moreover, we rounded off this article by modelling wind-related variables — wind energy production, wind penetration index and residual demand — with Brownian semi-stationary processes with gamma kernels. The empirical results showed that this type of model suits our application very well. Models of wind-related variables enable practitioners to apply models of electricity prices without relying on additional data.

Our studies indicate that the inclusion of wind energy production data in electricity pricing models is a promising area of research. We expect this effect to grow as wind energy constitutes an increasingly bigger part of total energy production in many countries — therefore its impact on the prices will increase. Furthermore, we suggest repeating similar experiments with other renewables such as solar, possibly making use of already existing models of solar power generation (Lingohr and Müller, 2019) or (Veraart and Zdanowicz, 2015). Since the participation of solar power in the total energy generation in Austria and Germany is still smaller than of the wind power, its impact might be smaller. However, the inclusion of other renewables would give us the full picture and help explain the prices even better due to their priority in the electricity market (merit-order effect).

Our work on modelling wind-related variables complements recent work by Benth and Pircalabu (2018), who propose a non-Gaussian Ornstein–Uhlenbeck model for pricing wind power futures and by Benth et al. (2021) who develop a multivariate continuous-time model of wind indices which can be used for hedging of wind risk.

Going forward, it would be interesting to further develop joint models for electricity prices and renewable sources of electricity to complement models using exogenous factors. Here the key challenge lies in a suitable concept for modelling the dependence, and various ideas from classical vector-valued time series models, regime-switching models to copula-based approaches would be worth exploring further.

CRediT authorship contribution statement

Paulina A. Rowińska: Conceptualization, Methodology, Software (original code), Validation, Formal analysis, Writing – original draft, Writing – review & editing. **Almut E.D. Veraart:** Conceptualization, Methodology, Software (review), Validation, Formal analysis, Writing – original draft, Writing – review & editing, Supervision. **Pierre Gruet:** Conceptualization, Methodology, Software (review), Formal analysis, Writing – review & editing, Supervision, Funding acquisition.

Acknowledgements

We would like to thank Olivier Féron for valuable comments and insights. PR gratefully acknowledges the financial support from EPSRC under the grant EP/L016613/1 as well as from EDF. We also thank the developers of the **R** package **yuima** for technical support. We would also like to thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the programme *The Mathematics of Energy Systems*. This work was supported by: EPSRC grant number EP/R014604/1.

Appendix. Proofs of the theoretical results

To prove Proposition 3.1 we will use the following lemma.

Lemma A.1. *Suppose that a for stochastic processes $H = (H(t))_{t\geq 0}$, and a random variable $\tilde{H}$, with $H, \tilde{H} \in L^2(\mathbb{Q})$ we have that $H(T) \xrightarrow{L^2(\mathbb{Q})} H^*$ as $T \to \infty$. Then, for a fixed $\tau > 0$, the $L^2(\mathbb{Q})$–limit of $\frac{1}{\tau}\int_{T_1}^{T_1+\tau} H(T)dT$ also exists and equals $\tilde{H}$.*

Proof of Lemma A.1. Our assumption says that $\lim_{T\to\infty}\mathbb{E}_{\mathbb{Q}}\left[(H(T)-\tilde{H})^2\right] = 0$, so that for all $\varepsilon > 0$ there exists $\tilde{T}$ such that for all $T > \tilde{T}$ we have $\mathbb{E}_{\mathbb{Q}}\left[\left(H(T)-\tilde{H}\right)^2\right] < \varepsilon$. Then, for a fixed $\tau > 0$, and for $T_1 > \tilde{T}$,

$$\mathbb{E}_{\mathbb{Q}}\left[\left(\frac{1}{\tau}\int_{T_1}^{T_1+\tau} H(T)dT - H^*\right)^2\right] = \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{1}{\tau}\int_{T_1}^{T_1+\tau}(H(T)-H^*)dT\right)^2\right] \leq \frac{1}{\tau}\int_{T_1}^{T_1+\tau}\mathbb{E}_{\mathbb{Q}}(H(T)-H^*)^2 dT < \varepsilon,$$

where we used the Cauchy–Schwartz inequality and Fubini's theorem. □

Proof of Proposition 3.1. First we will show that $\lim_{T_1\to\infty}\frac{1}{\tau}\int_{T_1}^{T_1+\tau}\int_{-\infty}^{t} g(T-s)\sigma(s-)\,dL(s)\,dT = 0$ in the $L^2(\mathbb{Q})$–sense. As all considered functions are measurable and non-negative, we can use Tonelli's theorem (Tao, 2011, p. 171) to compute

$$\begin{aligned}
&\lim_{T\to\infty}\mathbb{E}_{\mathbb{Q}}\left[\left(\int_{-\infty}^{t} g(T-s)\sigma(s-)dL(s)\right)^2\right]\\
&= \mathbb{E}_{\mathbb{Q}}\left[L(1)^2\right]\mathbb{E}_{\mathbb{Q}}\left[\sigma(0)^2\right]\lim_{T\to\infty}\int_{-\infty}^{t} g(T-s)^2 ds\\
&= \mathbb{E}_{\mathbb{Q}}\left[L(1)^2\right]\mathbb{E}_{\mathbb{Q}}\left[\sigma(0)^2\right]\lim_{T\to\infty}\int_{T-t}^{\infty} g^2(u)dy = 0,
\end{aligned}$$

which by Lemma A.1 proves this statement.

Note that $\mathbb{E}_{\mathbb{Q}}(Y(0)) = \mathbb{E}_{\mathbb{Q}}[L(1)]\,C$, where $C := \mathbb{E}_{\mathbb{Q}}[\sigma(0)]\int_0^\infty g(y)dy$. Now we need to prove that $\lim_{T_1\to\infty}\int_{T_1}^{T_1+\tau}\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\,dT = C$ in the $L^2(\mathbb{Q})$–sense. By Lemma A.1 it is enough to prove that $\lim_{T\to\infty}\mathbb{E}_{\mathbb{Q}}\left[\left(\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds - C\right)^2\right] = 0$. The case when σ is a constant is trivial, so we shall focus on the case when Assumption 3.1 holds.

We observe that

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[\left(\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds - C\right)^2\right]\\
&= \mathbb{E}_{\mathbb{Q}}\left[\left(\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\right)^2\right]\\
&\quad -2C\mathbb{E}_{\mathbb{Q}}\left[\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\right] + C^2.
\end{aligned} \tag{A.1}$$

Using Jensen's inequality and Tonelli's theorem, we obtain the following lower bound

$$\begin{aligned}
&\lim_{T\to\infty}\mathbb{E}_{\mathbb{Q}}\left[\left(\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\right)^2\right]\\
&\geq \lim_{T\to\infty}\left(\mathbb{E}_{\mathbb{Q}}\left[\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\right]\right)^2\\
&= \lim_{T\to\infty}\left(\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\right]ds\right)^2\\
&= \left(\mathbb{E}_{\mathbb{Q}}[\sigma(0)]\lim_{T\to\infty}\int_0^{T-t} g(y)dy\right)^2 = C^2.
\end{aligned} \tag{A.2}$$

On the other hand, almost surely, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)] &= \mathbb{E}_{\mathbb{Q}}\left[\sqrt{\int_{-\infty}^{s} e^{-\delta(s-x)}dV(x)}\,\middle|\,\mathcal{F}(t)\right]\\
&\leq \mathbb{E}_{\mathbb{Q}}\left[\sqrt{\int_{-\infty}^{t} e^{-\delta(s-x)}dV(x)} + \sqrt{\int_{t}^{s} e^{-\delta(s-x)}dV(x)}\,\middle|\,\mathcal{F}(t)\right]\\
&= \sqrt{\int_{-\infty}^{t} e^{-\delta(s-x)}dV(x)} + \mathbb{E}_{\mathbb{Q}}\left[\sqrt{\int_{t}^{s} e^{-\delta(s-x)}dV(x)}\right]\\
&= \sigma(t)e^{-\frac{\delta}{2}(s-t)} + \mathbb{E}_{\mathbb{Q}}\left[\sqrt{\int_{t}^{s} e^{-\delta(s-x)}dV(x)}\right],
\end{aligned} \tag{A.3}$$

where to non-negative processes we apply the inequality $\sqrt{a+b} \leq \sqrt{a}+\sqrt{b}$ as well as the identity $\int_{-\infty}^{t} e^{-\delta(s-x)}dV(x) = e^{-\delta(s-t)}\int_{-\infty}^{t} e^{-\delta(t-x)}dV(x) = \sigma(t)^2 e^{-\delta(s-t)}$. We remark that $\sigma(s)$ is stationary in mean, i.e. $\mathbb{E}_{\mathbb{Q}}[\sigma(s)] = \mathbb{E}_{\mathbb{Q}}[\sigma(0)]$ for all $s \in \mathbb{R}$. Furthermore,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[\sqrt{\int_{t}^{s} e^{-\delta(s-x)}dV(x)}\right] = \mathbb{E}_{\mathbb{Q}}\left[\sqrt{\int_{0}^{s-t} e^{-\delta u}dV(u)}\right]\\
&\leq \mathbb{E}_{\mathbb{Q}}\left[\sqrt{\int_{0}^{\infty} e^{-\delta u}dV(u)}\right] = \mathbb{E}_{\mathbb{Q}}[\sigma(0)],
\end{aligned} \tag{A.4}$$

as we integrate the non-negative exponential function w.r.t. a subordinator. Combining Eqs. (A.3) and (A.4) leads to $\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)] \leq \sigma(t)e^{-\frac{\delta}{2}(s-t)} + \mathbb{E}_{\mathbb{Q}}[\sigma(0)]$. Therefore

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,\mathbb{E}_{\mathbb{Q}}[\sigma(u)|\mathcal{F}(t)]\right]\\
&\leq \mathbb{E}_{\mathbb{Q}}\left[\sigma(0)^2\right]e^{-\frac{\delta}{2}(s-t)}e^{-\frac{\delta}{2}(u-t)} + \mathbb{E}_{\mathbb{Q}}[\sigma(0)]^2\left(e^{-\frac{\delta}{2}(s-t)} + e^{-\frac{\delta}{2}(u-t)} + 1\right).
\end{aligned} \tag{A.5}$$

This implies that

$$\begin{aligned}
&\lim_{T\to\infty}\mathbb{E}_{\mathbb{Q}}\left[\left(\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\right)^2\right]\\
&= \lim_{T\to\infty}\int_t^T\int_t^T g(T-s)g(T-u)\mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,\mathbb{E}_{\mathbb{Q}}[\sigma(u)|\mathcal{F}(t)]\right]dsdu\\
&\leq \mathbb{E}_{\mathbb{Q}}\left[\sigma(0)^2\right]\left(\lim_{T\to\infty}\int_t^T g(T-s)e^{-\frac{\delta}{2}(s-t)}ds\right)^2\\
&+ 2\mathbb{E}_{\mathbb{Q}}[\sigma(0)]^2\lim_{T\to\infty}\left(\int_t^T g(T-s)ds\int_t^T g(T-s)e^{-\frac{\delta}{2}(s-t)}ds\right)\\
&+ \mathbb{E}_{\mathbb{Q}}[\sigma(0)]^2\left(\lim_{T\to\infty}\int_t^T g(T-s)ds\right)^2\\
&= \mathbb{E}_{\mathbb{Q}}\left[\sigma(0)^2\right]\left(\lim_{T\to\infty}e^{-\frac{\delta}{2}(T-t)}\int_0^{T-t} g(y)e^{\frac{\delta}{2}y}dy\right)^2\\
&+ 2\mathbb{E}_{\mathbb{Q}}[\sigma(0)]^2\lim_{T\to\infty}\int_0^{T-t} g(y)dy\lim_{T\to\infty}\left(e^{-\frac{\delta}{2}(T-t)}\int_0^{T-t} g(y)e^{\frac{\delta}{2}y}dy\right)\\
&+ \mathbb{E}_{\mathbb{Q}}[\sigma(0)]^2\left(\lim_{T\to\infty}\int_0^{T-t} g(y)dy\right)^2 = C^2,
\end{aligned} \tag{A.6}$$

where we used Assumption 3.1. From Eqs. (A.2) and (A.6) we deduce that $\lim_{T\to\infty}\mathbb{E}_{\mathbb{Q}}\left[\left(\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\right)^2\right] = C^2$. Also, by Tonelli's theorem $\lim_{T\to\infty}\mathbb{E}_{\mathbb{Q}}\left[\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\right] = \mathbb{E}_{\mathbb{Q}}[\sigma(0)]\lim_{T\to\infty}\int_t^T g(T-s)ds = C$. Hence, taking the limit in Eq. (A.1) together with Lemma A.1 implies that $\lim_{T_1,\to\infty}\int_{T_1}^{T_1+\tau}\int_t^T g(T-s)\mathbb{E}_{\mathbb{Q}}[\sigma(s)|\mathcal{F}(t)]\,ds\,dT = C$ in the $L^2(\mathbb{Q})$–sense. □

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eneco.2021.105640.

References

Barndorff-Nielsen, O.E., Benth, F.E., Veraart, A.E.D., 2013. Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. Bernoulli 19 (3), 803–845, URL: http://projecteuclid.org/euclid.bj/1372251144.

Barndorff-Nielsen, O.E., Benth, F.E., Veraart, A.E.D., 2015. Recent advances in ambit stochastics with a view towards tempo-spatial stochastic volatility/intermittency. In: Advances in Mathematics of Finance. In: Banach Center Publ., vol. 104, Polish Acad. Sci. Inst. Math., Warsaw, pp. 25–60. http://dx.doi.org/10.4064/bc104-0-2.

Barndorff-Nielsen, O., Benth, F., Veraart, A., 2018. Ambit Stochastics. In: Probability Theory and Stochastic Modelling, Springer International Publishing, URL: https://books.google.co.uk/books?id=j-11DwAAQBAJ.

Bennedsen, M., 2017. A rough multi-factor model of electricity spot prices. Energy Econ. 63, 301–313, URL: https://www.sciencedirect.com/science/article/pii/S0140988317300610.

Benth, F.E., Christensen, T.S., Rohde, V., 2021. Multivariate continuous-time modeling of wind indexes and hedging of wind risk. Quant. Finance 21 (1), 165–183. http://dx.doi.org/10.1080/14697688.2020.1804606.

Benth, F.E., Klüppelberg, C., Müller, G., Vos, L., 2014. Futures pricing in electricity markets based on stable CARMA spot models. Energy Econ. 44, 392–406. http://dx.doi.org/10.1016/j.eneco.2014.03.020.

Benth, F.E., Paraschiv, F., 2018. A space-time random field model for electricity forward prices. J. Bank. Financ. 95, 203–216, URL: https://www.sciencedirect.com/science/article/pii/S0378426617300766.

Benth, F.E., Pircalabu, A., 2018. A non-Gaussian Ornstein–Uhlenbeck model for pricing wind power futures. Appl. Math. Finance 25 (1), 36–65. http://dx.doi.org/10.1080/1350486X.2018.1438904.

Borovkova, S., Geman, H., 2006. Analysis and modelling of electricity futures prices. Stud. Nonlinear Dyn. Econom. 10 (3), 239–263. http://dx.doi.org/10.2202/1558-3708.1372.

Brockwell, P.J., Davis, R.A., Yang, Y., 2011. Estimation for non-negative Lévy-driven CARMA processes. J. Bus. Econ. Stat. 29 (2), 250–259, URL:: http://www.tandfonline.com/doi/abs/10.1198/jbes.2010.08165.

Carmona, R., Coulon, M., Schwarz, D., 2013. Electricity price modeling and asset valuation: A multi-fuel structural approach. Math. Financ. Econ. 7 (2), 167–202. http://dx.doi.org/10.1007/s11579-012-0091-4.

Cartea, A., Figueroa, M.G., Geman, H., 2009. Modelling electricity prices with forward looking capacity constraints. Appl. Math. Finance 16 (2), 103–122, URL: http://www.tandfonline.com/doi/abs/10.1080/13504860802351164.

Deschatre, T., Féron, O., Gruet, P., 2021. A survey of electricity spot and futures price models for risk management applications. Preprint available on arXiv, 2103.16918. URL: https://arxiv.org/abs/2103.16918.

Deschatre, T., Veraart, A.E.D., 2018. A joint model for electricity spot prices and wind penetration with dependence in the extremes. In: Drobinski, P., Mougeot, M., Picard, D., Plougonven, R., Tankov, P. (Eds.), Renewable Energy: Forecasting and Risk Management. Springer International Publishing, Cham, pp. 185–207. http://dx.doi.org/10.1007/978-3-319-99052-1_10.

Drost, H., 2018. Philentropy: Information theory and distance quantification with R. J. Open Source Softw. 3 (26), 765, URL: http://joss.theoj.org/papers/10.21105/joss.00765.

Duffie, D., 1992. Dynamic Asset Pricing Theory. Princeton University Press, Princeton, USA.

Elberg, C., Hagspiel, S., 2015. Spatial dependencies of wind power and interrelations with spot price dynamics. European J. Oper. Res. 241 (1), 260–272. http://dx.doi.org/10.1016/j.ejor.2014.08.026.

García, I., Klüppelberg, C., Müller, G., 2011. Estimation of stable CARMA models with an application to electricity spot prices. Stat. Model. 11 (5), 447–470. http://dx.doi.org/10.1177/1471082X1001100504.

Hamerle, A., Singer, H., Nagl, W., 1993. Identification and estimation of continuous time dynamic systems with exogenous variables using panel data. Econom. Theory 9 (2), 283–295. http://dx.doi.org/10.1017/S0266466600007544.

Jónsson, T., Pinson, P., Madsen, H., 2010. On the market impact of wind energy forecasts. Energy Econ. 32 (2), 313–320. http://dx.doi.org/10.1016/j.eneco.2009.10.018.

Ketterer, J.C., 2014. The impact of wind power generation on the electricity price in Germany. Energy Econ. 44, 270–280. http://dx.doi.org/10.1016/j.eneco.2014.04.003.

Lingohr, D., Müller, G., 2019. Stochastic modeling of intraday photovoltaic power generation. Energy Econ. 81, 175–186. http://dx.doi.org/10.1016/j.eneco.2019.03.007.

Lingohr, D., Müller, G., 2021. Conditionally independent increment processes for modeling electricity prices with regard to renewable power generation. Energy Econ. 105244, URL: https://www.sciencedirect.com/science/article/pii/S0140988321001493.

Lüthi, D., Breymann, W., 2016. Ghyp: A package on generalized hyperbolic distribution and its special cases. R package version 1.5.7. URL: https://CRAN.R-project.org/package=ghyp.

Müller, G., Seibert, A., 2019. Bayesian estimation of stable CARMA spot models for electricity prices. Energy Econ. 78, 267–277. http://dx.doi.org/10.1016/j.eneco.2018.10.016.

Nicolosi, M., Fürsch, M., 2009. The impact of an increasing share of RES-e on the conventional power market — the example of Germany. Z. Energiewirtschaft 33 (3), 246–254. http://dx.doi.org/10.1007/s12398-009-0030-0.

Phillips, P.C.B., 1974. The estimation of some continuous time models. Econometrica 42 (5), 803–823. http://dx.doi.org/10.2307/1913790.

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, URL: https://www.R-project.org/.

Rowińska, P.A., 2020. Stochastic Modelling and Statistical Inference for Electricity Prices, Wind Energy Production and Wind Speed (PhD thesis). Imperial College London.

Rowińska, P.A., Veraart, A.E., Gruet, P., 2021. A multifactor approach to modelling the impact of wind energy on electricity spot prices: Supplementary material. Preprint.

Sargan, J.D., 1974. Some discrete approximations to continuous time stochastic models. J. R. Stat. Soc. Ser. B Stat. Methodol. 36 (1), 74–90. http://dx.doi.org/10.1111/j.2517-6161.1974.tb00987.x.

Tao, T., 2011. An Introduction to Measure Theory Vol. 126. American Mathematical Society, http://dx.doi.org/10.1090/gsm/126.

Trapletti, A., Hornik, K., 2018. Tseries: Time series analysis and computational finance. R package version 0.10-46. URL: https://CRAN.R-project.org/package=tseries.

Veraart, A.E.D., 2016. Modelling the impact of wind power production on electricity prices by regime-switching Lévy semistationary processes. In: Benth, F.E., Di Nunno, G. (Eds.), Stochastics of Environmental and Financial Economics. Springer International Publishing, Cham, pp. 321–340. http://dx.doi.org/10.1007/978-3-319-23425-0_13.

Veraart, A.E.D., Zdanowicz, H., 2015. Modelling and predicting photovoltaic power generation in the EEX market. Preprint Available At SSRN. URL: https://ssrn.com/abstract=2691906.

Weisser, D., 2003. A wind energy analysis of Grenada: an estimation using the 'Weibull' density function. Renew. Energy 28 (11), 1803–1812, URL: https://www.sciencedirect.com/science/article/pii/S0960148103000168.

Weron, R., 2008. Market price of risk implied by Asian-style electricity options and futures. Energy Econ. 30 (3), 1098–1115, URL:: https://www.sciencedirect.com/science/article/pii/S014098830700076X.