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

Paulina A. Rowińska

Imperial College London

p.rowinska15@imperial.ac.uk

Almut E.D. Veraart

Imperial College London

a.veraart@imperial.ac.uk

Pierre Gruet

EDF Research and Development

pierre.gruet@edf.fr

July 2, 2021

Abstract

We present the supplementary material for the article Rowińska et al. (2021). In particular, we provide background material on the class of generalised hyperbolic distributions and continuous-time autoregressive moving average processes. Moreover, we include a detailed description of the data used in our empirical work and expand on some of our empirical results.

Contents

1 Introduction

This supplementary material provides additional material for the article Rowińska et al. (2021). It is structured as follows.

First, Section 2 reviews the definition of the generalised hyperbolic distribution and its important subclasses and specifies the parameterisation used in our empirical study. Then, Section 3 provides background material on CARMA processes and their simulation and estimation. Such processes lie at the heart of the short-term factor in our new stochastic model. Next, Section 4 provides a detailed data description and exploratory data analysis. Moreover, Section 5 describes how we have dealt with the seasonal behaviour of the various time series. Finally, Section 6 presents the detailed empirical results from our new models proposed for wind-related variables, which have been summarised in Section 6 in the main article Rowińska et al. (2021).

This supplementary material contains material from the PhD thesis Rowińska (2020).

2 Generalised hyperbolic distributions

In our article, we consider distributions from the generalised hyperbolic (GH) class. Their properties such as fat tails and skewness make them suitable for our applications. Most importantly, they are infinitely divisible, which in particular means that they can generate a Lévy process $(L(t))_{t \geq 0}$ such that $L(1)$ has a generalised hyperbolic density.
Lüthi & Breymann (2016) provided a concise introduction to these distributions and developed an **R** package **ghyp**, which came in useful many times during our research.

2.1 Definitions and properties

We begin by defining the building block of these Gaussian mixture distributions: the generalised inverse Gaussian distribution (GIG).

Definition 2.1 (**Generalised inverse Gaussian distribution**). The density of a *generalised inverse Gaussian* variable $W \sim GIG(\lambda, \chi, \psi)$ with parameters satisfying one of the following

- $\chi > 0, \psi \geq 0, \lambda < 0$,

- $\chi > 0, \psi > 0, \lambda = 0$,

- $\chi \geq 0, \psi > 0, \lambda > 0$,

is given by

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

Now we are ready to define generalised hyperbolic distributions as Gaussian mixtures.

Definition 2.2 (**Generalised hyperbolic distribution**). A random vector X follows a *univariate generalised hyperbolic (GH) distribution* if

$$X \stackrel{law}{=} \mu + W\gamma + \sqrt{W}\sigma Z, \tag{2.1}$$

where $Z \sim N(0,1)$, $\sigma, \mu, \gamma \in \mathbb{R}$ and W is a scalar-valued random variable, independent of Z, whose distribution is generalised inverse Gaussian: $GIG(\lambda, \chi, \psi)$.

We interpret the parameters of generalised hyperbolic distribution as follows.

- μ corresponds to the location.

- $\Sigma = \sigma^2$ measures the dispersion.

- λ, χ and ψ determine the shape, i.e. the relationship between the tails and the center: as a rule of thumb, their large values indicate that the distribution resembles a Gaussian distribution.

- γ indicates the skewness: $\gamma = 0$ means that the distribution is symmetric.

We also observe that

$$X|W = w \sim N\left(\mu + w\gamma, w\Sigma\right).$$

We recall one more desirable property of this distribution class.

Proposition 2.1 (Lüthi & Breymann (2016), Proposition 1). *The generalised hyperbolic class is closed under linear transformations: if*

$$X \sim GH\left(\lambda, \chi, \psi, \mu, \Sigma, \gamma\right)$$

and

$$Y = aX + b$$

for $a, b \in \mathbb{R}$, then

$$Y \sim GH\left(\lambda, \chi, \psi, a\mu + b, a^2\Sigma, a\gamma\right).$$

2.2 Special cases of generalised hyperbolic distributions

In the class of generalised hyperbolic distributions we distinguish special subclasses.

- *Hyperbolic (hyp)* for $\lambda = 1$.

- *Normal inverse Gaussian (NIG)* for $\lambda = \frac{1}{2}$.

- *Variance gamma (VG)* for $\chi = 0$ and $\lambda > 0$ (a limiting case).

- *Generalised hyperbolic Student's-t* (often called *Student's-t*) for $\psi = 0$ and $\lambda < 0$.

In our research we use the **R** package **ghyp** (Lüthi & Breymann (2016)) to fit 11 different distributions to the increments of the Lévy process of interest and rank them according to the Akaike information criterion (AIC). Namely, we compare symmetric and asymmetric versions of hyp, NIG, VG, Student's-t and GH distributions, as well as a Gaussian distribution, to find the best fitting one.

2.3 Parametrisations

The $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ parametrisation has alternatives; the **R** package **ghyp** (Lüthi & Breymann (2016)) lets the user choose between three options. The other two are $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma,)$ and $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ parametrisations. Lüthi & Breymann (2016) provided ranges of parameters for special cases listed in Subsection 2.2 as well as explicit functions switching between parametrisations.

3 Background on CARMA processes

We will now review the definition and key properties of continuous-time autoregressive moving average processes.

Definition 3.1 (**CARMA process**). Let $\{L(t) : t \geq 0\}$ be a Lévy process and $0 \leq q < p$. We define the *L-driven CARMA(p,q) process* with parameters $[a_1, \ldots, a_p; b_1, \cdots, b_q]$ as a stationary solution of the system of formal stochastic differential equations

$$a(D)Y(t) = b(D)DL(t), \tag{3.1}$$

where D denotes "differentiation" with respect to t,

$$a(z) := z^p + a_1 z^{p-1} + \cdots + a_p, \tag{3.2}$$

$$b(z) := b_0 + b_1 z + \cdots + b_{p-1}, \tag{3.3}$$

$a(z)$ and $b(z)$ have no common roots, $b_q = 1$ and $b_j = 0$ for $q < j < p$. Since we cannot define the derivative $DL(t)$ in the usual sense, we interpret Equation (3.1) as the state-space representation, i.e.

$$Y(t) = \boldsymbol{b}^T \boldsymbol{X}(t), \tag{3.4}$$

$$d\boldsymbol{X}(t) = \boldsymbol{A}\boldsymbol{X}(t)dt + \boldsymbol{e}dL(t), \tag{3.5}$$

where

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{pmatrix}, \tag{3.6}$$

$$\boldsymbol{e} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix}. \tag{3.7}$$

Following Brockwell et al. (2011), we assume that $\boldsymbol{X}(0)$ is independent of $\{L(t), t \geq 0\}$, all eigenvalues of $\boldsymbol{A}$ (easily calculated as roots of the autoregressive polynomial $a(z)$) have negative real parts and

$$\boldsymbol{X}(0) \overset{law}{=} \int_0^\infty e^{\boldsymbol{A}(t-u)} \boldsymbol{e} dL(u). \tag{3.8}$$

Then $\{\boldsymbol{X}(t), t \geq 0\}$ is a strictly stationary solution of Equation (3.5).

In order to define CARMA(p,q) on the real line, we can define a Lévy process $\{L'(t), t \geq 0\}$ such that $L(t)$ and $L'(t)$ are iid. Then we extend $L(t)$ to the whole real line as

$$\tilde{L}(t) := L(t)\mathbb{1}_{[0,\infty)}(t) - L'(-t-)\mathbb{1}_{(-\infty,0]}(t), \quad t \in \mathbb{R}. \tag{3.9}$$

In the remaining part of this chapter we will denote $\tilde{L}$ by L and call it the *background driving Lévy process* of $Y(t)$. Now we are ready to provide the necessary and sufficient conditions for the existence of a covariance stationary solution of Equation (3.5).

Proposition 3.1 (Brockwell et al. (2011), Proposition 1). *Equation (3.5) has a covariance stationary solution $\boldsymbol{X}$ such that $\boldsymbol{X}(t)$ is independent of $\{L(s) - L(t)\}_{s>t}$ for all $t \in \mathbb{R}$ if and only if the real parts of the roots of $a(z) = 0$ are negative. This solution has the form*

$$\boldsymbol{X}(t) = \int_{-\infty}^t e^{\boldsymbol{A}(t-u)} \boldsymbol{e} dL(u) \overset{law}{=} \int_0^\infty e^{\boldsymbol{A}u} \boldsymbol{e} dL(u).$$

This leads us to an alternative definition of the continuous autoregressive moving average process (CARMA).

Definition 3.2 (CARMA(p,q) process). Let us assume that L is a Lévy process, the roots of $a(z) = 0$ are distinct and their real parts are negative. Then we define the *L-driven CARMA(p,q) process* with parameters $a_1, \ldots, a_p; b_1, \ldots, b_q$ as the process satisfying

$$Y(t) = \boldsymbol{b}^T \boldsymbol{X}(t) = \int_{-\infty}^\infty g(t-u) dL(u)$$

with

$$g(t) = \boldsymbol{b}^T e^{\boldsymbol{A}t} \boldsymbol{e} \mathbb{1}_{[0,\infty)}(t)$$

called the *kernel of the CARMA process* $\{Y(t)\}$. Y is a *causal* function of L because $Y(t)$ does not depend on $\{L(s) - L(t), s \geq t\}$ for all t.

To assure the existence and uniqueness of our CARMA processes, we need to state the following additional assumptions.

Assumption 3.1. Unless stated otherwise, we assume that roots of $a(z) = 0$ are distinct and their real parts are negative. Furthermore, $a(z)$ and $b(z)$ have no common roots.

CARMA processes have an intuitive and very useful representation as a sum of Lévy-driven Ornstein-Uhlenbeck processes, described in more detail in Subsection 3.4. This observation helps us estimate the parameters of these objects as well as simulate them.

Proposition 3.2 (Brockwell et al. (2011), Proposition 2). *Under Assumption 3.1, any CARMA(p,q) process $Y(t)$ can be expressed as a sum of dependent and possibly complex-valued CAR(1) processes, i.e.*

$$Y(t) = \sum_{r=1}^{p} Y^{(r)}(t), \tag{3.10}$$

where for $r = 1, \ldots, p$

$$Y^{(r)}(t) = \int_{-\infty}^{t} \alpha_r e^{\lambda_r (t-u)} dL(u), \tag{3.11}$$

$$\alpha_r = \frac{b(\lambda_r)}{a'(\lambda_r)}, \tag{3.12}$$

with $a'(\cdot)$ denoting the derivative of $a(\cdot)$.

In our article Rowińska et al. (2021), we focus on one particular example: the CARMA(2,1) process, whose properties we will review in the following.

Example 3.1 (CARMA(2,1)). The Lévy-driven CARMA(2,1) process is defined as the strictly stationary solution to the system of equations

$$(D^2 + a_1 D + a_2)Y(t) = (b_0 + D)DL(t), \tag{3.13}$$

where $t \in \mathbb{R}$, $b(z) = b_0 + z$ and $a(z) = z^2 + a_1 z + a_2 = (z - \lambda_1)(z - \lambda_2)$, $\lambda_1 \neq \lambda_2$, $Re(\lambda_1) < 0$ and $Re(\lambda_2) < 0$. Furthermore, for $u \geq 0$ the kernel of $Y(t)$ equals

$$g(u) = \alpha_1 e^{\lambda_1 u} + \alpha_2 e^{\lambda_2 u}, \tag{3.14}$$

where $\alpha_1 = \frac{b_0 + \lambda_1}{\lambda_1 - \lambda_2}$ and $\alpha_2 = \frac{b_0 + \lambda_2}{\lambda_2 - \lambda_1}$. Therefore we can represent this process as

$$Y(t) = \alpha_1 \int_{-\infty}^{t} e^{\lambda_1 (t-u)} dL(u) + \alpha_2 \int_{-\infty}^{t} e^{\lambda_2 (t-u)} dL(u). \tag{3.15}$$

3.1 The sampled CARMA process

In practice, we typically observe the continuous-time process $Y(t)$ only in discrete time. Precisely, we denote by $\{Y_n := Y(nh), n = 0, 1, \ldots, N\}$ the sampled process, where N is the number of available observations and $h > 0$ is a small, fixed interval between the consecutive observations. The following proposition explores the relationship between the continuous and discrete processes, which prepares us for the estimation of CARMA(p,q) parameters.

Proposition 3.3 (Brockwell et al. (2011), Proposition 3). *Under Assumption 3.1 the following hold.*

1. *The sampled process $\{Y_n := Y(nh), n = 0, 1, \ldots, N\}$, with a fixed $h > 0$, can be represented as $Y_n = \sum_{r=1}^{p} Y_n^{(r)}$, where $n \in \mathbb{Z}$. For each $r = 1, \ldots, p$, the discrete-time process $\{Y_n^{(r)}\}$ is obtained by sampling the component CAR(1) process $\{Y^{(r)}(t)\}$ (cf. Subsection 3.4) at spacing $h > 0$. As Y is strictly stationary,*

$$\forall n \in \mathbb{Z} \quad Y_n^{(r)} = e^{\lambda_r} Y_{n-1}^{(r)} + Z_n^{(r)}, \tag{3.16}$$

with the iid noise

$$\forall n \in \mathbb{Z} \quad Z_n^{(r)} = \alpha_r \int_{(n-1)h}^{nh} e^{\lambda_r(nh-u)} dL(u). \tag{3.17}$$

2. *The sampled process $\{Y_n := Y(nh), n = 0, 1, \ldots, N\}$ satisfies*

$$\phi(B)Y_n = \sum_{r=1}^{p} V_{n-r+1}^{r} =: U_n, \tag{3.18}$$

where

$$\phi(z) := \prod_{r=1}^{p} \left(1 - e^{\lambda_r h} z\right) = 1 - \sum_{r=1}^{p} \phi_r z^r \tag{3.19}$$

and B denotes the backshift operator, i.e. $B^j Y_n := Y_{n-j}$. For each $r = 1, \ldots, p$, we define the iid sequence $\{V_n^{(r)}\}$ by

$$V_n^{(r)} := \int_{(n-1)h}^{nh} \sum_{k=1}^{p} \alpha_k \left(e^{(r-1)h\lambda_k} - \sum_{j=1}^{r-1} \phi_j e^{(r-1-j)h\lambda_k} \right) \times e^{(nh-u)\lambda_k} dL(u). \tag{3.20}$$

3. *We can represent the right-hand side of Equation (3.18) as an invertible moving average*

$$\theta(B)W_n := W_n + \theta_1 W_{n-1} + \cdots + \theta_{p-1} W_{n-p+1}, \tag{3.21}$$

where $\{W_n\}$ is a sequence of white noise (possibly not iid) and $\theta_1, \ldots, \theta_q$ are moving average constants depending on the CARMA process. Therefore $\{Y_n\}$ can be represented as a weak ARMA(p,p-1) process (an ARMA(p,p-1) process allowing for not iid driving white noise) such that

$$\phi(B)Y_n = \theta(B)W_n \tag{3.22}$$

and

$$W_n = \theta(B)^{-1} \sum_{r=1}^{p} V_{n-r+1}^{p}. \tag{3.23}$$

3.2 Estimation

We can now estimate the CARMA parameters, following the algorithm described by García et al. (2011). Brockwell et al. (2011) showed that for a fixed sampling interval $h > 0$ 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 below.

Since we are particularly interested in the CARMA(2,1) process, we specify the estimation procedure for this particular process. We slightly abuse the notation by denoting variables and their estimates with the same symbols; the meaning should be clear from the context.

1. We estimate the ARMA(2,1) parameters $\boldsymbol{\beta} = (\phi_1, \phi_2, \theta)^T$, using the quasi-maximum likelihood approach (alternatively one could use least squares).

2. Equation (3.18) has the form

$$Y_n - \phi_1 Y_{n-1} - \phi_2 Y_{n-2} = \left(1 - e^{\lambda_1 h} B\right)\left(1 - e^{\lambda_2 h} B\right) Y_n. \tag{3.24}$$

By multiplying through and matching coefficients, we obtain

$$\phi_1 = e^{\lambda_1 h} + e^{\lambda_2 h}, \quad \phi_2 = -e^{(\lambda_1 + \lambda_2)h}. \tag{3.25}$$

This gives us a nonlinear system of two equations for the estimators of λ_1 and λ_2, whose solutions are

$$\lambda_1 = \log\left(\frac{\phi_1}{2} + \sqrt{\left(\frac{\phi_1}{2}\right)^2 + \phi_2^2}\right), \tag{3.26}$$

$$\lambda_2 = \log\left(\frac{\phi_1}{2} - \sqrt{\left(\frac{\phi_1}{2}\right)^2 + \phi_2^2}\right). \tag{3.27}$$

From there we immediately calculate $a_1 = -(\lambda_1 + \lambda_2)$ and $a_2 = \lambda_1 \lambda_2$.

Remark 3.1. We implicitly assume that $\phi_1 > 0$ and $\phi_2 < 0$, so this method is not suitable for some CARMA processes. Alternative estimation methods include quasi-maximum likelihood implemented in the **R** package **yuima**, see Iacus & Mercuri (2015).

3. The right-hand side of Equation (3.18) implies the form of the autocovariance of the process $\phi(B)Y_n$:

$$\forall s \in \mathbb{R} \quad \gamma_U(s) = \text{Cov}\left(\phi(B)Y_n, \phi(B)Y_{n-s}\right). \tag{3.28}$$

Furthermore, using Corollary 3 by Barndorff-Nielsen et al. (2013), for all $s \in \mathbb{R}$ we calculate the autocovariance of $Y(t)$

$$\gamma_Y(s) = \text{Cov}\left(Y(s), Y(0)\right) = \int_0^\infty g(x)g(x+s)dx = w_1 e^{\lambda_1 s} + w_2 e^{\lambda_2 s}, \tag{3.29}$$

where

$$w_1 = \frac{\alpha_1^2 \lambda_1 \lambda_2 + {\alpha_1}^2 \lambda_2^2 + 2\lambda_1 \lambda_2 \alpha_1 \alpha_2}{2\lambda_1 \lambda_2 (\lambda_1 + \lambda_2)}, \tag{3.30}$$

$$w_2 = \frac{\alpha_2^2 \lambda_1 \lambda_2 + {\alpha_2}^2 \lambda_1^2 + 2\lambda_1 \lambda_2 \alpha_1 \alpha_2}{2\lambda_1 \lambda_2 (\lambda_1 + \lambda_2)}. \tag{3.31}$$

Thus for all $s \in \mathbb{R}$ the autocorrelation of $Y(t)$ equals

$$\delta_Y(s) = \frac{\gamma_Y(s)}{\gamma_Y(0)} = \frac{w_1 e^{\lambda_1 s} + w_2 e^{\lambda_2 s}}{w_1 + w_2}. \tag{3.32}$$

For CARMA(2,1) we can rewrite Equation (3.28) as

$$\gamma_U(0) = (1 + \phi_1^2 + \phi_2^2)\gamma_Y(0) + (2\phi_2\phi_2 - 2\phi_1)\gamma_Y(1) - 2\phi_2\gamma_Y(2),$$

$$\gamma_U(1) = -\phi_2\gamma_Y(3) + \phi_1(\phi_2 - 1)\gamma_Y(2) \tag{3.33}$$

$$+ (1 + \phi_1^2 + \phi_2^2 - \phi_2)\gamma_Y(1) + \phi_1(\phi_2 - 1)\gamma_Y(0),$$

where we use explicit formulae for $\gamma_Y(\cdot)$ given by Equation (3.29). Since they depend on a_0, a_1 and b_0, we plug in the estimates of the first two parameters.

On the other hand, the autocorrelation function at the first lag of a moving average process with coefficient θ can be expressed as

$$\delta_U(1) = \frac{\gamma_U(1)}{\gamma_U(0)} = \frac{\theta}{1 + \theta^2}. \tag{3.34}$$

Now we can replace the left-hand side of Equation (3.34) by expressions from Equation (3.33) to get a non-linear equation for b_0, which we solve numerically.

4. Having estimated the parameters of CARMA(2,1), we need to recover the background driving Lévy process L(t), using results by (Brockwell et al. 2011, Section 5):

$$X^{(0)}(t) = X^{(0)}(0)e^{-b_0 t} + \int_0^t e^{-b_0(t-s)}Y(s)ds, \tag{3.35}$$

$$X^{(1)}(t) = DX^{(0)}(t) = -b_0 X^{(0)}(t) + Y(t). \tag{3.36}$$

The canonical state vector $\mathbf{Y}(t)$ is given by

$$\begin{pmatrix} Y^{(1)}(t) \\ Y^{(2)}(t) \end{pmatrix} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} -\lambda_2(b_0 + \lambda_1) & (b_0 + \lambda_1) \\ \lambda_1(b_0 + \lambda_2) & -b_0 + \lambda_2 \end{pmatrix} \begin{pmatrix} X^{(0)}(t) \\ X^{(1)}(t) \end{pmatrix}. \tag{3.37}$$

To recover the background driving Lévy process L(t) we can choose one of two equations, either with $r = 1$ or $r = 2$:

$$L(t) = \frac{1}{\alpha_r}\left[Y^{(r)}(t) - Y^{(r)}(0) - \lambda_r \int_0^t Y^{(r)}(s)ds \right]. \tag{3.38}$$

We approximate the integral using the trapezoidal rule. Brockwell et al. (2011) recommend choosing r such that $|\lambda_r|$ is minimal, which minimizes the contribution of $\lambda_r \int_0^t Y^{(r)}(s)ds$ compared to $Y^{(r)}(t) - Y^{(r)}(0)$.

3.3 Simulation

To simulate any CARMA(p,q) process, we use its representation as a sum of CAR(1) processes stated in Proposition 3.2. For brevity, we restrict our attention to simulating a CARMA(2,1)

process $Y(t)$ with parameters a_1, a_2 and b_0 and eigenvalues of the matrix $\boldsymbol{A}$: λ_1 and λ_2. We start by simulating two Ornstein-Uhlenbeck processes $Y^{(1)}(t)$ and $Y^{(2)}(t)$ with λ_1 and λ_2, driven by the same Lévy process (cf. Section 3.4.3). Then we compute

$$Y(t) = \alpha_1 Y^{(1)}(t) + \alpha_2 Y^{(2)}(t),$$

where

$$\alpha_1 = \frac{b_0 + \lambda_1}{\lambda_1 - \lambda_2}$$
$$\alpha_2 = \frac{b_0 + \lambda_2}{\lambda_1 - \lambda_2}.$$

3.4 Example: Lévy-driven CAR(1) processes

One of the most important processes belonging to the CARMA class is the Lévy-driven CAR(1) process, also called the Lévy-driven Ornstein-Uhlenbeck process. For more details we refer the reader to Brockwell et al. (2007).

Definition 3.3 (**Lévy-driven CAR(1) process**)**.** Let $L(t)$ be a Lévy process. The *L-driven Ornstein-Uhlenbeck process* with the parameter $\lambda > 0$ is a strictly stationary solution of the stochastic differential equation

$$dY(t) = -\lambda Y(t)dt + dL(t), \quad t > 0.$$

As in Equation (3.9), we extend $L(t)$ to the whole real line. If $\lambda > 0$, for all $t > s \in \mathbb{R}$, the process $Y(t)$ defined by

$$Y(t) = \int_{-\infty}^{t} e^{-\lambda(t-u)} dL(u)$$

is a strictly stationary solution to

$$dY(t) = -\lambda Y(t)dt + dL(t).$$

The process $L(t)$ is called the *background driving Lévy process* of $Y(t)$.

3.4.1 The sampled CAR(1) process

In this section we point out the relationship between the Ornstein-Uhlenbeck process and the discrete autoregressive process of order one. We denote by $\{Y_n := Y(nh), n = 0, 1, \ldots, N\}$ the sampled CAR(1) process, where N is the number of available observations and $h > 0$ is a small, fixed interval between the consecutive observations. The sampled CAR(1) process is the discrete time autoregressive AR(1) process:

$$\forall n \in \mathbb{Z} \quad Y_n = \phi Y_{n-1} + Z_n,$$

where

$$\phi = e^{-\lambda h}$$

and

$$\forall n \in \mathbb{Z} \quad Z_n = \int_{(n-1)h}^{nh} e^{\lambda(nh-u)} dL(u)$$

is the iid and positive noise.

3.4.2 Estimation

To estimate the parameters of a CAR(1) process $Y(t)$ we follow the procedure suggested by Brockwell et al. (2007) and the references therein. Again, we slightly abuse the notation by denoting variables and their estimates with the same symbols, the meaning of which is clear from the context. We proceed as follows.

1. We estimate the parameter ϕ of the sampled process Y_n, for example by maximising the likelihood function.

2. We compute

$$\lambda = -\log(\phi).$$

3. For all $n = 1, \ldots N$ we define the increments of the driving process $L(t)$ on intervals $((n-1)h), nh]$ by

$$\Delta_n^h L := L_{nh} - L_{(n-1)h} = Y_{nh} - Y_{(n-1)h} + \lambda \int_{(n-1)h}^{nh} Y_u du.$$

We recover these increments using the trapezoidal rule

$$\Delta_n^h L \approx Y_n - Y_{n-1} + \frac{\lambda h}{2} \left(Y_n + Y_{n-1} \right).$$

3.4.3 Simulation

We simulate a CAR(1) process $Y(t)$ as follows.

1. We partition the interval $[0, T]$ into N equal intervals of length $h = \frac{T}{N}$.

2. For $i = 1, \ldots, N$ we simulate the increments of the driving process $\Delta_n^h L$ following the appropriate centred distribution, for example from the generalised hyperbolic class. If the mean μ of Lévy increments differs from zero, after simulating the increments we subtract the mean from the resulting time series.

3. We set the initial value of Y to 0.

4. For $n = 1, \ldots, (N-1)$ we use the Euler-Maruyama method to compute

$$Y_{n+h} = Y_n + \lambda h Y_n + (L_{n+h} - L_n) + h\mu.$$

In practice we first simulate a CAR(1) process of length $2N$. We further discard the first half of the time series as a burn-in period. We finally resample the resulting time series every $\frac{1}{h}$ points to obtain a process corresponding to the original data, if needed.

3.5 Simulation study

In order to test the quality of the estimation procedure used for the CARMA process driven by a Lévy process we simulated 1,000 paths of Y with $N = 1,824$ observations. The simulation method requires simulating two Ornstein-Uhlenbeck processes driven by the same Lévy noise. We use the Euler-Maruyama method with the step size $h = 0.01$. We first simulate two Ornstein-Uhlenbeck processes of length $\frac{2N}{h} = 364,800$. We discard the first half of each time series as a burn-in period. Finally, we resample every $\frac{1}{h} = 100$ points to obtain two time series of length $N = 1,824$.

For each path we re-estimate the parameters $\mathrm{par}^{\mathrm{true}}$, which after averaging over all paths give us Monte Carlo estimates $\mathrm{par}^{\mathrm{MC}}$. We compute the bias and standard error of all parameters: $\mathrm{par}^{\mathrm{MC}} - \mathrm{par}^{\mathrm{true}}$ and $\sqrt{\frac{\sum_{i=1}^{1000}\left(\mathrm{par}_i^{\mathrm{MC}} - \mathrm{par}^{\mathrm{true}}\right)^2}{1000}}$, respectively. We additionally report the relative bias and error defined as $\frac{\mathrm{par}^{\mathrm{MC}} - \mathrm{par}^{\mathrm{true}}}{\mathrm{par}^{\mathrm{MC}}}$ and $\frac{\sqrt{\frac{\sum_{i=1}^{1000}\left(\mathrm{par}_i^{\mathrm{MC}} - \mathrm{par}^{\mathrm{true}}\right)^2}{1000}}}{\mathrm{par}^{\mathrm{MC}}}$, respectively. From Table 1 we learn that the (relative) biases as well as (relative) standard errors of estimates are small. We present histograms of resulting estimates in Figure 1.

	a_1	a_2	b_0
True estimate	0.809	0.048	0.194
Monte Carlo estimate	0.811	0.045	0.191
Standard error	0.116	0.023	0.066
Relative error	0.143	0.519	0.348
Bias	0.002	-0.003	-0.004
Relative bias	0.002	-0.059	-0.019

Table 1: Parameters of CARMA(2,1) kernel estimated via bootstrapping with $1,000$ simulations.
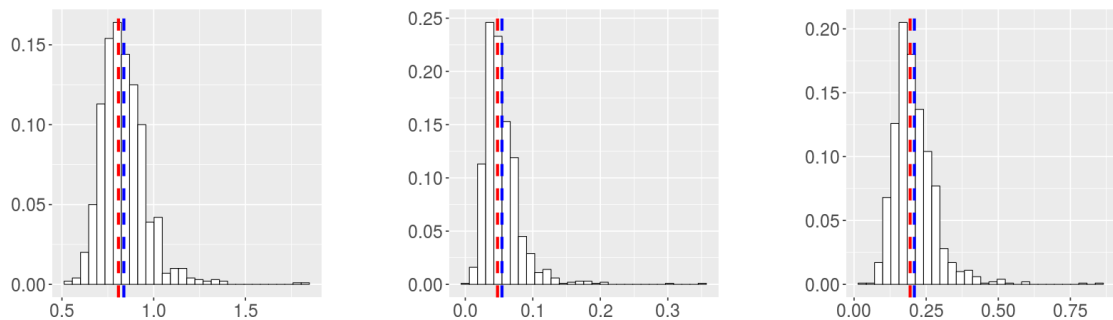

Figure 1: Parameters of CARMA(2,1) kernel estimated from $1,000$ simulated paths (from left to right: a_1, a_2 and b_0). Red vertical lines indicate true estimates, while red lines Monte Carlo estimates.

Remark 3.2. It is crucial to simulate the Ornstein-Uhlenbeck processes on a sufficiently refined grid. For example, the Euler-Maruyama step size $ht = 1.0$ results in large (relative) biases of estimated parameters.

Similarly we re-estimate the parameters of the generalised hyperbolic noise. In Table 2 we observe large standard and relative errors as well as (relative) biases. However, we recall that to each path of the noise's increments we fit 11 distributions from the generalised hyperbolic class and choose the one with the smallest value of AIC (Akaike information criterion). Slightly different parameters might result in a distribution belonging to a different class. Therefore the distributions fitted to similar paths might have very different parameters. Figure 2 helps explain this phenomenon. For example, the estimates of λ and γ are clearly bimodal, so the average of parameters resulting from both cases matches neither of two distributions. However, this should not worry us because these different parameters result in very similar distributions.

	λ	α	δ	β	μ
True estimate	-2.131	0.020	11.592	-0.020	1.178
Monte Carlo estimate	0.631	0.415	29.214	-0.153	5.802
Standard error	38.008	1.212	20.699	1.062	9.926
Relative error	60.282	2.923	0.709	-6.944	1.711
Bias	2.761	0.394	17.622	-0.132	4.624
Relative bias	4.379	0.951	0.603	0.866	0.797

Table 2: Parameters of the generalised hyperbolic noise estimated via bootstrapping with $1,000$ simulations.
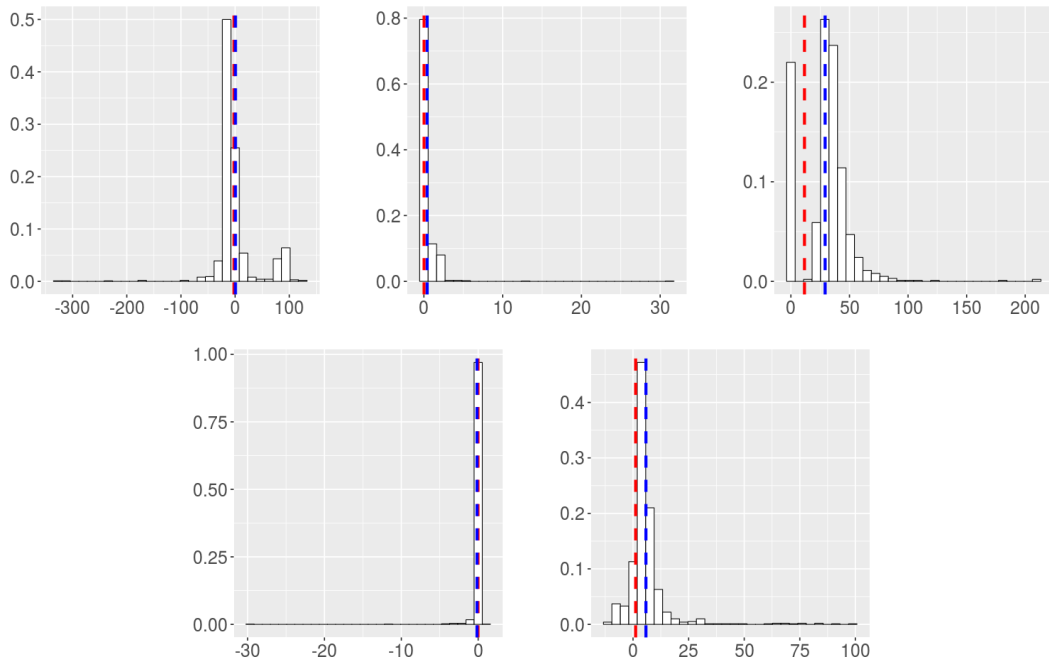

Figure 2: Parameters of the generalised hyperbolic noise estimated from $1,000$ simulated paths (from left to right: λ, α, δ, β and μ). Red vertical lines indicate true estimates, while blue lines Monte Carlo estimates.

4 Exploratory data analysis

4.1 Data description

In our empirical work, we consider daily data from the Austrian and German energy markets: electricity spot and monthly futures prices, total load (total energy generation) as well as wind energy generation. Unless stated otherwise, we estimate our models on five-year-long time series of daily data between 01.01.2011 and 31.12.2015. We downloaded hourly time series from the website of the Open Power System Data project (Open Power System Data (2019)) and Austrian Power Grid (Austrian Power Grid (2019)). We use the following time series.

- Austrian Power Grid (APG):

    - Day-ahead forecasts of wind energy production in Austria [MW].

    - Day-ahead forecasts of total load (total energy production) in Austria [MW].

- Open Power System Data:

    - Day-ahead forecasts of wind energy production in Germany, summed over all four energy providers: TransnetBW, TenneT, 50Hertz and Amprion [MW].

    - Actual total load (total energy production) in Germany [MW].

    - Spot prices (day-ahead baseload prices) in the joint Austrian-German market [EUR].

These two data sources are not fully consistent with regard to the load as they provide values from different areas of Austria. The Open Power System Data gather data not only from the control area APG, but also from the rest of the country, including industrial production units and railroad consumption, not connected to the APG (European Commission (2018)). However, in our research we are mostly interested in the impact of wind energy generation on energy prices. Since wind energy should not influence commercial production units, these inconsistencies do not cause problems in our analyses — especially as Austrian load is a few orders of magnitude smaller than German one, so any differences in the Austrian data become negligible when we look at both countries together. Figure 3 presents the control areas of four German transmission system operators.

Remark 4.1. The meaning of the term "load" varies across the literature: it can denote either power (the rate of electricity consumption over time) in kW or MW or consumed energy in kWh or MWh (Hong (2014)). Although hourly data sets will report the same magnitude regardless of the meaning, in our research we consider also other time periods, so for consistency we report the data in MWh and EUR/MWh. We recall that we define a watt as one joule per second, so it measures the power produced or used in a given time in other words, the rate of energy generation
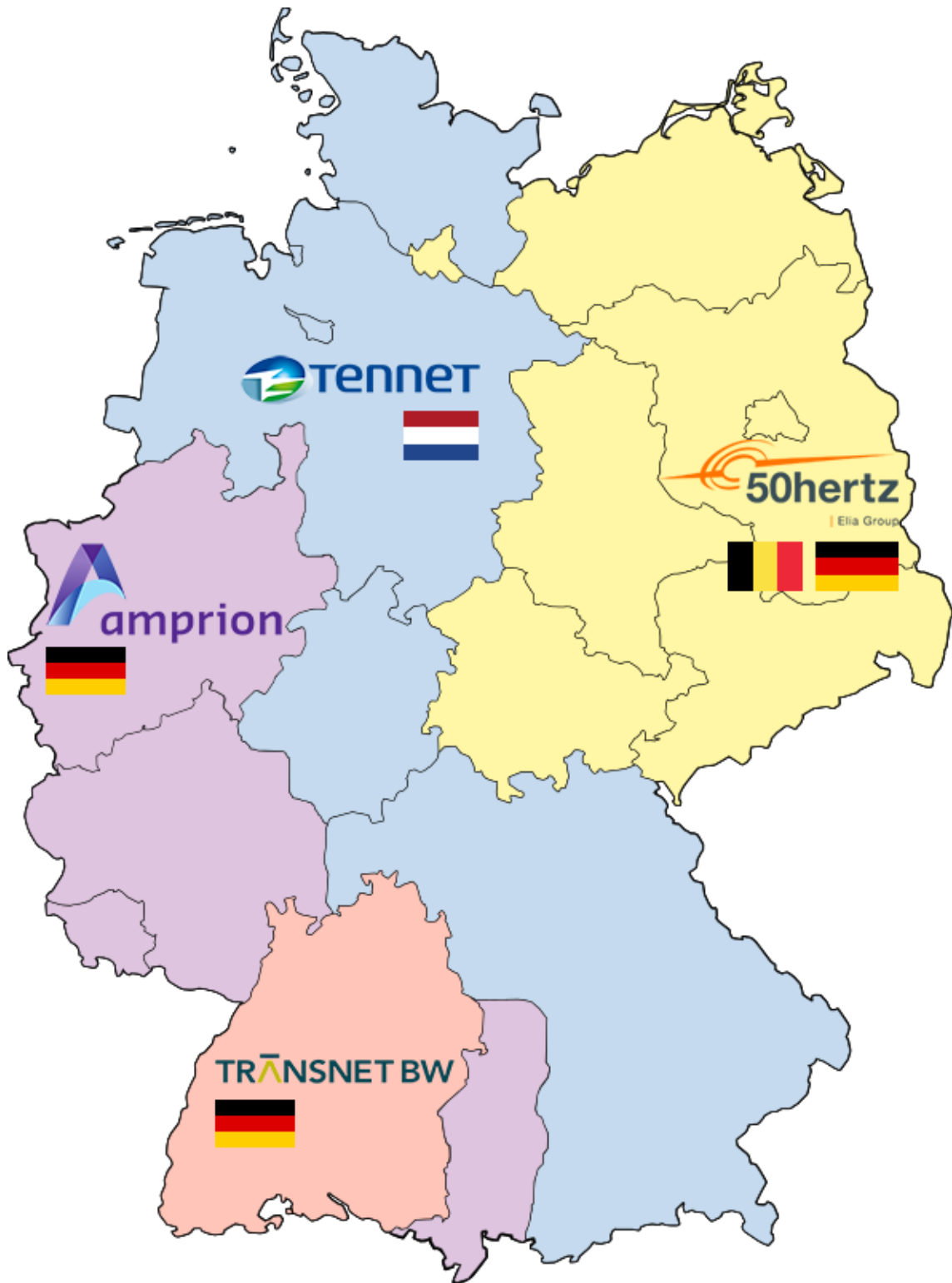
Figure 3: Control areas of transmission system operators in Germany (McLloyd (2013)).

or consumption. We also recall that

$$1\,\mathrm{kW} = 10^3\,\mathrm{W}$$
$$1\,\mathrm{MW} = 10^6\,\mathrm{W}$$
$$1\,\mathrm{GW} = 10^9\,\mathrm{W}.$$

On the other hand, watt-seconds are units of energy defined as one joule. For example, a country with hourly demand of 20 GW will use 20 GWh in one hour. A country with hourly demand of 10 GW will use the same amount of energy in two hours. Conversions of energy units are similar to power units:

$$1\,\mathrm{kWh} = 10^3\,\mathrm{W\,h}$$
$$1\,\mathrm{MWh} = 10^6\,\mathrm{W\,h}$$
$$1\,\mathrm{GWh} = 10^9\,\mathrm{W\,h}.$$

We aggregate the data to daily values: we average the prices and sum up total wind energy productions as well as total loads. Similarly to Jónsson et al. (2010), we approximate the unavailable day-ahead forecasts of total load in Germany by randomly perturbing the actual total load. Precisely, we assume that for each day $t = 1, \ldots, N$ (N denotes the number of observed days)

$$\text{forecasted load}(t) = \text{actual load}(t) + \varepsilon(t),$$

where $\varepsilon \sim N\left(0, \sigma^2\right)$ and σ equals approximately 2% of averaged actual load. While this method works very well in practice (Jónsson et al. (2010), Veraart (2016)), Jónsson et al. (2010) point out some flaws of this approach. First, it cannot account for possibly autocorrelated residuals of forecasted load. Second, it assumes the independence of forecast errors of wind energy production and total load, which is unrealistic as they both stem from weather forecasts. However, the effect of these simplifications on our applications is negligible. Alternatively, we could obtain the forecasted load by smoothing the actual load, as we expect the forecasts to be smoother than the data.

We sum up the values of forecasted load and wind energy production in both countries to get figures for the whole region. Finally, we convert the units to gigawatts. In the rest of this supplementary material, "wind" denotes the sum of day-ahead forecasts (true or perturbed) of wind energy production levels in Austria and Germany (in GWh).

4.2 Futures prices

In Rowińska et al. (2021), we propose a model calibration procedure which additionally requires the prices of monthly futures contracts: one month ahead (1MAH) up to six months ahead (6MAH), which we downloaded from the webpage of the European Energy Exchange (2017). Figure 4 shows the empirical data for 1MAH and 6MAH, the two most extreme time horizons considered. Since

the trading periods of futures contracts are limited, we can see discontinuities in the beginning and the end of trading periods. We also note seasonal patterns in the data as well as the downward trend over time. Finally, as opposed to the spot prices, we do not observe large spikes in the futures prices. ter Haar (2010) explains this finding by the smaller influence of demand-supply imbalances on longer term contracts. Table 3 presents the summary statistics of considered futures data. For more details about futures contracts, including a thorough data analysis, we refer the reader to ter Haar (2010) and the references therein.
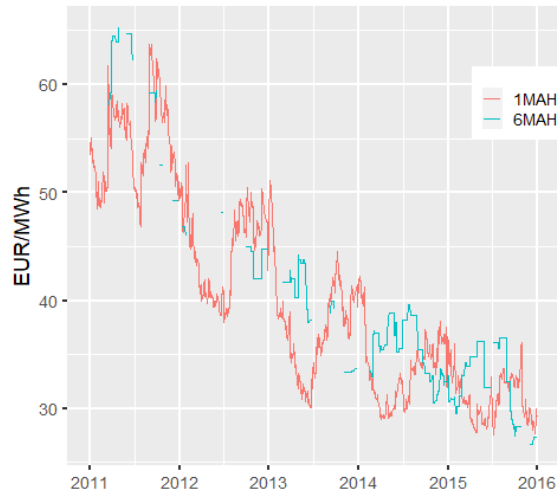
Figure 4: Time series of futures prices data: one-month-ahead (red) and six-months-ahead (blue).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1MAH	27.60	32.45	38.45	40.33	48.16	63.70
2MAH	26.40	32.54	38.50	40.96	48.56	67.80
3MAH	26.40	33.60	38.58	41.35	47.50	67.50
4MAH	28.00	33.17	38.15	41.33	46.88	65.50
5MAH	24.40	33.58	38.65	40.85	47.00	66.00
6MAH	26.65	33.25	36.55	39.40	42.00	65.25

Table 3: Summary statistics of monthly futures prices computed after removing NA values.

4.3 Summary statistics

Table 4 presents the summary statistics of the daily data: the day-ahead electricity prices, the day-ahead forecasts of wind energy production and the total load, denoted by S, WD and LD, respectively.

When we write $WD(t)$, we refer to the wind energy production forecast for day t, which is available on day $t - 1$. Similarly, $LD(t)$ denotes the forecasted daily load for day t.

First we note the high variability of all time series. We would like to point out the occurrence of daily negative prices and the strict positivity of daily wind energy production forecasts. Figure 5a, Figure 6a and Figure 7a illustrate the same data with time series plots, in which we can clearly see the seasonal patterns. Furthermore, during the time of observations, the spot prices have significantly decreased while the wind energy production increased. Also, the plots of the empirical autocorrelation functions, see Figure 5b, Figure 6b and Figure 7b, indicate strong seasonal patterns in all variables.

In Figure 8 we observe the densities as well as the relationships (scatter plots and correlation coefficients) between variables. We also see that wind energy production forecasts are clearly right-skewed. Finally, the correlation coefficients between S and WD as well as between S and LD indicate some level of dependency between these variables.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
S	-54.70	31.26	38.76	39.18	48.35	99.43
WD	2.21	67.44	117.93	156.30	204.42	792.91
LD	1103.53	1441.13	1617.32	1586.58	1727.01	2005.90

Table 4: Summary statistics of day-ahead electricity prices S [EUR/MWh], day-ahead forecasts of daily wind energy production WD [GWh] and day-ahead forecasts of daily load LD [GWh].

(a) Time series plot of electricity spot prices S.

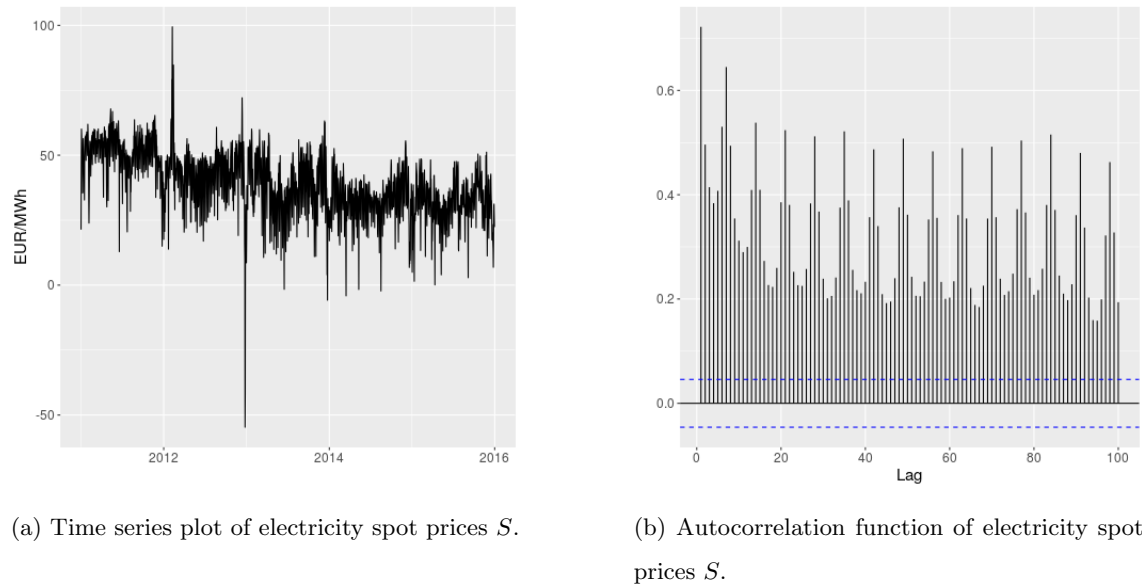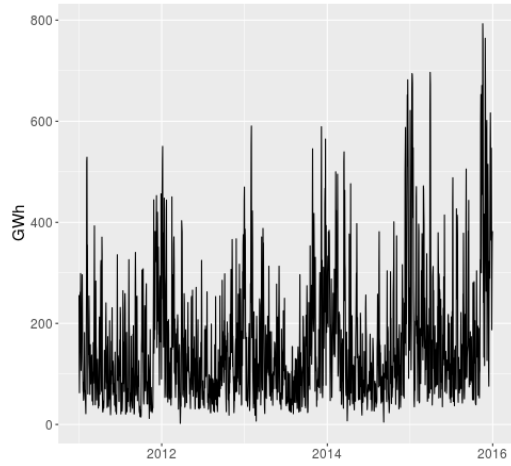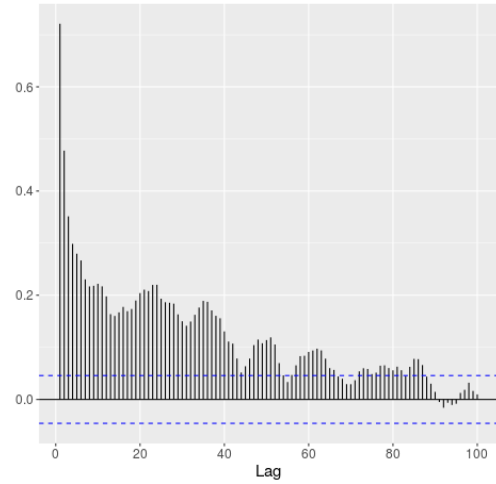(b) Autocorrelation function of electricity spot prices S.

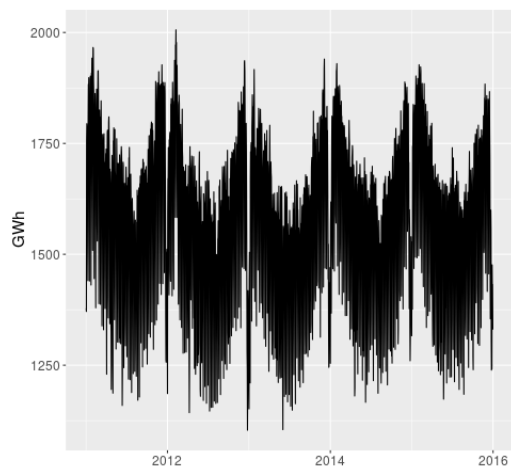Figure 5: Time series and autocorrelation function plots of electricity spot prices S.

(a) Time series plot of day-ahead forecasts of daily wind energy production WD.
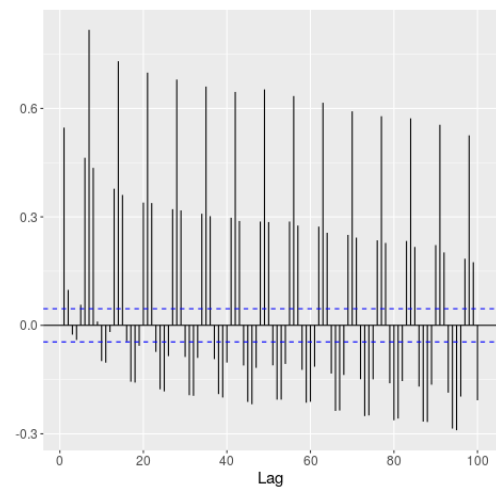

(b) Autocorrelation function of day-ahead forecasts of daily wind energy production WD.

Figure 6: Time series and autocorrelation function plots of day-ahead forecasts of daily wind energy production WD.

(a) Time series plot of day-ahead forecasts of daily load LD.

(b) Autocorrelation function of day-ahead forecasts of daily load LD.

Figure 7: Time series and autocorrelation function plots of day-ahead forecasts of daily load LD.
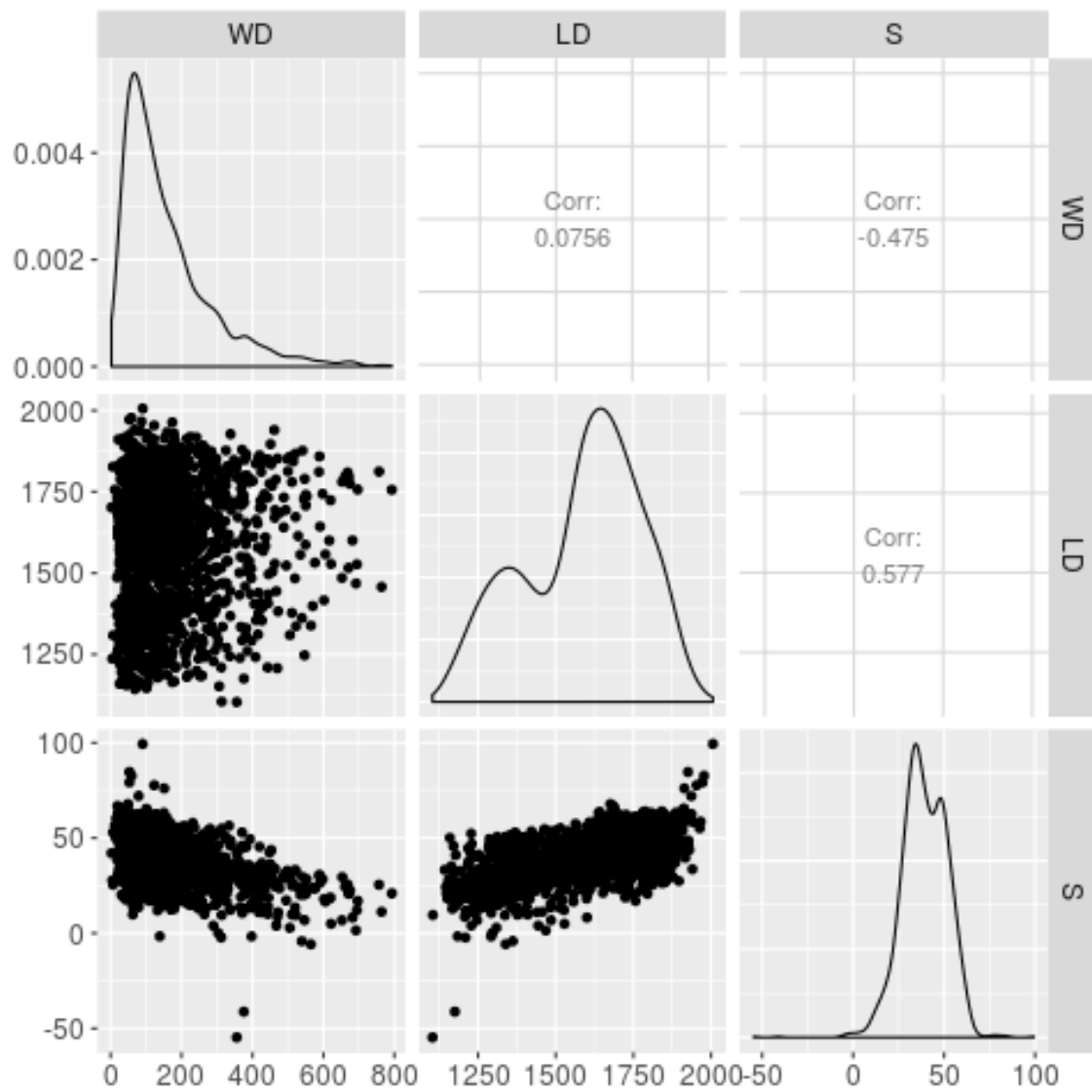
Figure 8: Relationships between day-ahead forecasts of wind energy generation WD and total load LD as well as electricity price S.

4.3.1 Wind penetration index and residual demand

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
WPI	0.00	0.04	0.07	0.10	0.13	0.52
RD	693.18	1267.06	1471.30	1430.28	1590.27	1919.02

Table 5: Summary statistics of day-ahead wind penetration index WPI [-] and day-ahead residual demand RD [GWh].
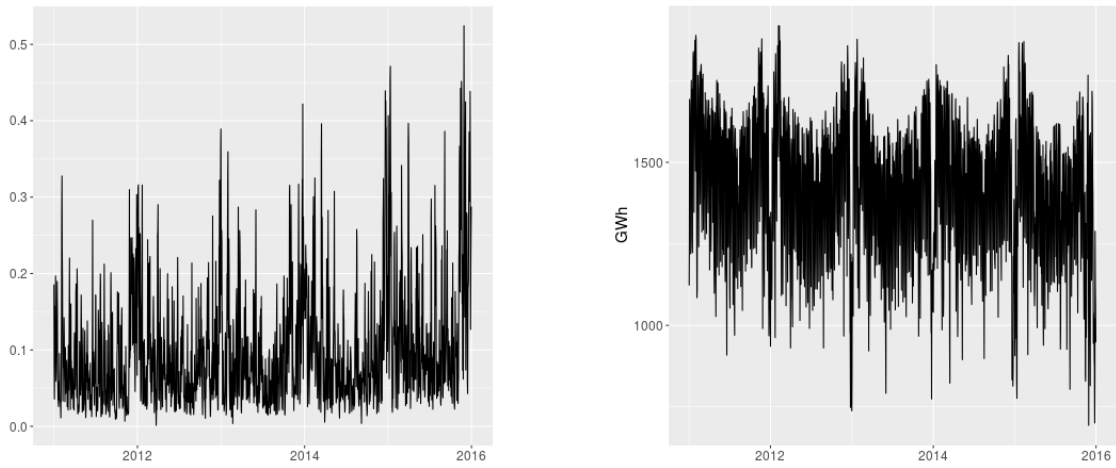
(a) Time series plot of day-ahead forecasts of the wind penetration index WPI.

(b) Time series plot of day-ahead forecasts of the residual demand RD.

Figure 9: Time series and autocorrelation function plots of day-ahead forecasts of WPI and RD.

Table 5 shows the summary statistics of the two time series of the wind penetration index (WPI) and the residual demand (RD). In the discussed period (2011–2015) on average 10% of total energy produced in Austria and Germany came from wind. Figure 9a and Figure 9b present both time series. We note that WPI is very volatile and oscillates a lot around its mean level.

5 Dealing with seasonality

5.1 Seasonal patterns in the data

All variables considered in our study exhibit seasonal patterns. According to Figure 10, the on average lowest prices occur during the weekends and in the summer months. We also see that the prices decreased from 2011 to 2015. Figure 11 shows that, while (unsurprisingly) we generate similar amounts of wind energy regardless of the weekday, the production is much higher in winter than in summer. Moreover, we notice a significant increase in wind energy production as the years progressed, which might explain the decrease in prices. When it comes to the total load and residual demand, in Figure 12 and Figure 13 we can see that similarly to spot prices lowest values occur during the weekends as well as in summer. The total energy generation stayed approximately the same over the years. Finally, from Figure 14 we learn that the highest values of the wind penetration index occur on the weekends and in winter, which is mostly due to lower values of total load in the denominator.
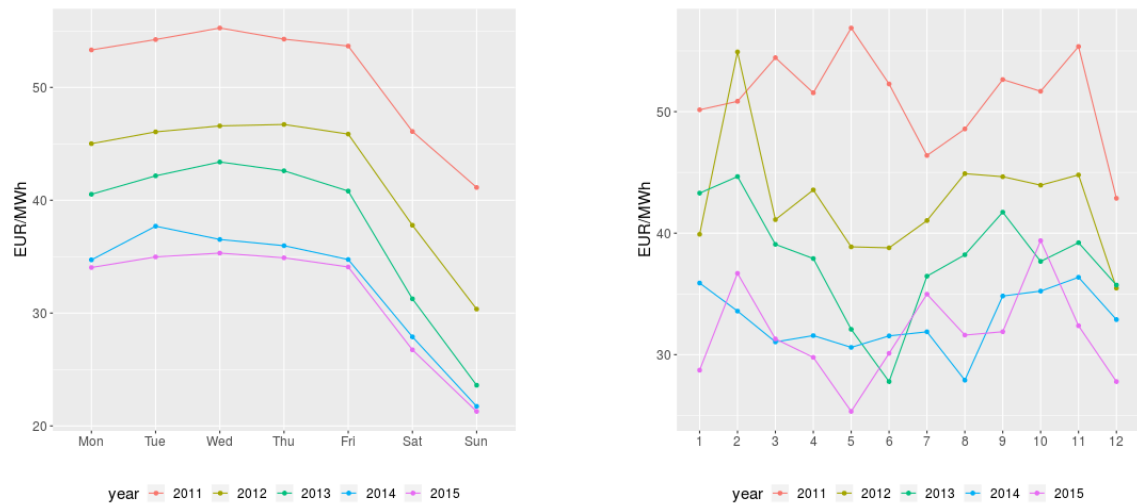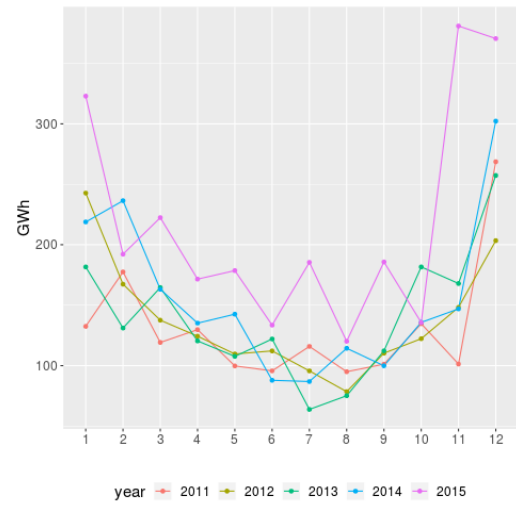

(a) Averaged day-ahead electricity prices S grouped by weekdays.

(b) Averaged day-ahead electricity prices S grouped by months.

Figure 10: Day-ahead electricity prices S grouped by seasons: weekdays (left panel) and months (right panel).
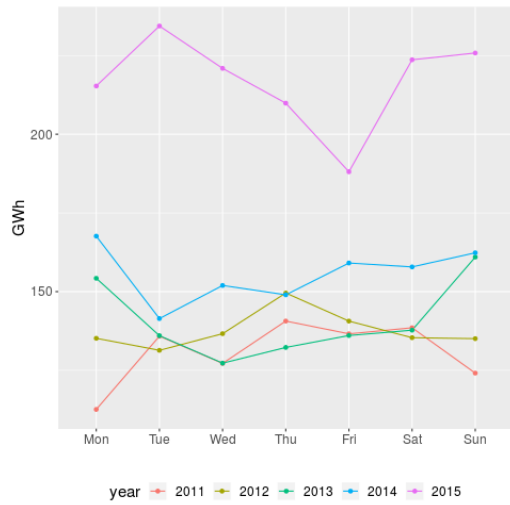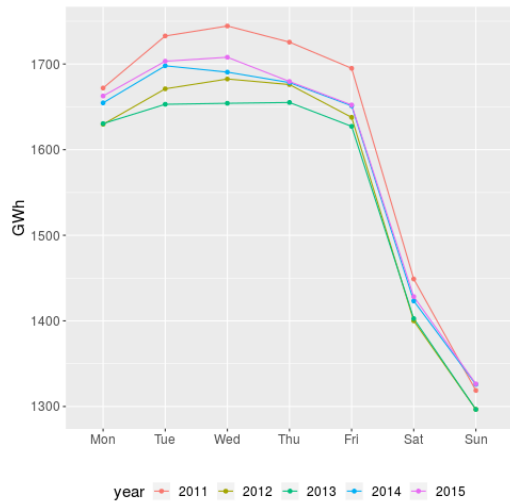
(a) Averaged day-ahead forecasts of daily wind energy production WD grouped by weekdays.

(b) Averaged day-ahead forecasts of daily wind energy production WD grouped by months.

Figure 11: Day-ahead forecasts of daily wind energy production WD grouped by seasons: weekdays (left panel) and months (right panel).
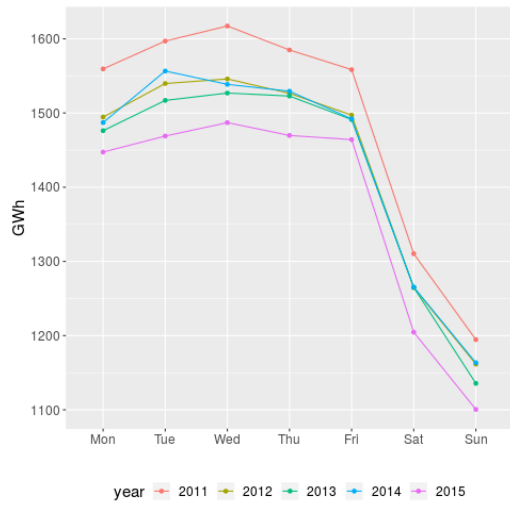

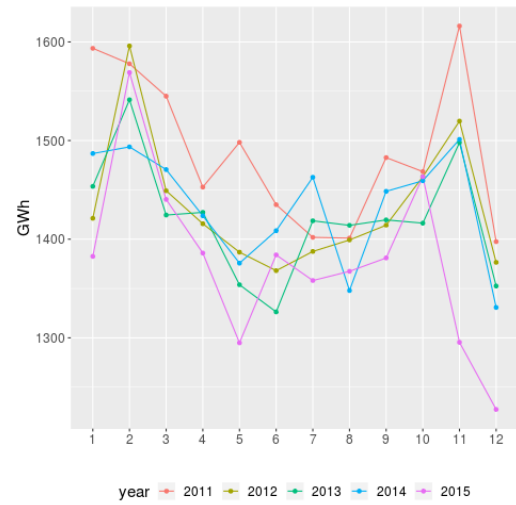(a) Averaged day-ahead forecasts of daily load LD grouped by weekdays.

(b) Averaged day-ahead forecasts of daily load LD grouped by months.

Figure 12: Day-ahead forecasts of daily load LD grouped by seasons: weekdays (left panel) and months (right panel).
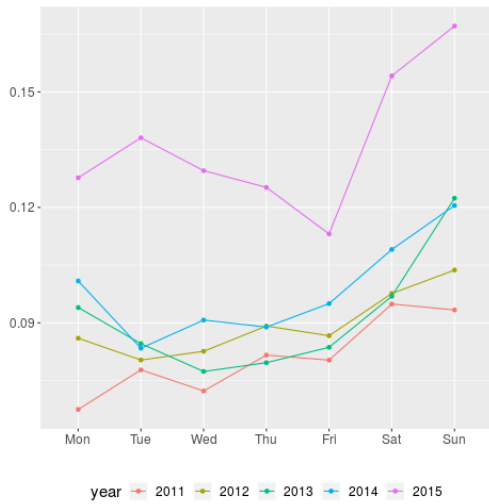
(a) Averaged day-ahead forecasts of the residual demand RD grouped by weekdays.
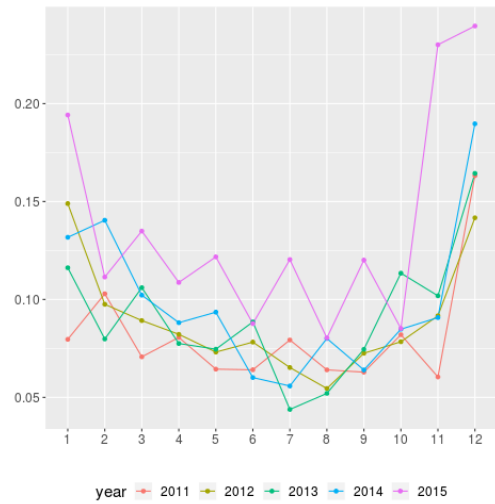
(b) Averaged day-ahead forecasts of the residual demand RD grouped by months.

Figure 13: Day-ahead forecasts of the residual demand RD grouped by seasons: weekdays (left panel) and months (right panel).

(a) Averaged day-ahead forecasts of the wind penetration index WPI grouped by weekdays.

(b) Averaged day-ahead forecasts of the wind penetration index WPI grouped by months.

Figure 14: Day-ahead forecasts of the wind penetration index WPI grouped by seasons: weekdays (left panel) and months (right panel).

5.2 Seasonality functions

Before we propose models for any of the variables, we remove trends and seasonal patterns from all data sets by assuming that

$$S(t) = \Lambda^S(t) + \overline{S}(t),$$

$$WD(t) = \Lambda^{WD}(t) + \overline{WD}(t),$$

$$LD(t) = \Lambda^{LD}(t) + \overline{LD}(t),$$

$$WPI(t) = \Lambda^{WPI}(t) + \overline{WPI}(t),$$

$$RD(t) = \Lambda^{RD}(t) + \overline{RD}(t),$$

where $\Lambda^\star$ denote deterministic seasonality and trend functions for appropriate variables, while bars detrended and deseasonalised variables. Let

$$\Lambda_1^\star(t) = c_0^\star + c_1^\star t + c_2^\star \cos\left(\frac{c_3^\star + 2\pi t}{365}\right),$$

where stars correspond to coefficients of the variable of interest. We additionally define dummy variables for days of the week $(d_{\mathrm{Mon}}, \ldots, d_{\mathrm{Sat}})$ and main holidays ($h(t) = 0$ for Dec 24th, Dec 25th, Dec 26th and Jan 1st; $h(t) = 1$ otherwise). We set

$$\Lambda_2^\star(t) = c_h^\star h(t) + c_{\mathrm{Mon}}^\star d_{\mathrm{Mon}}(t) + \cdots + c_{\mathrm{Sat}}^\star d_{\mathrm{Sat}}(t),$$

and

$$\Lambda_3^\star(t) = c_h^\star h(t) + c_{\mathrm{Mon}}^\star d_{\mathrm{Mon}}(t) + \cdots + c_{\mathrm{Fri}}^\star d_{\mathrm{Fri}}(t),$$

where the latter does not include a dummy variable for Saturday. Now we can define

$$\Lambda^S(t) = \Lambda_1^S(t) + \Lambda_2^S(t),$$

$$\Lambda^{WD}(t) = \Lambda_1^{WD}(t),$$

$$\Lambda^{LD}(t) = \Lambda_1^{LD}(t) + \Lambda_2^{LD}(t),$$

$$\Lambda^{WPI}(t) = \Lambda_1^{WPI}(t) + \Lambda_3^{WPI}(t),$$

$$\Lambda^{RD}(t) = \Lambda^{LD}(t) - \Lambda^{WD}(t).$$

As we do not observe weekly effects in WD, we remove only yearly seasonality. Furthermore, the dummy variable for Saturday is not significant in case of WPI, so we do not include this variable. Finally, since RD is an additive function of LD and WD, we do not introduce any additional seasonal factors for this variable.

We use the function **nls** from the **R** package **stats** (R Core Team (2018)) to estimate the coefficients of seasonality functions Λ. We initialise all parameters with 0.1. We present the estimated

parameters in Table 6, Table 7, Table 8 and Table 9. The estimated coefficients reflect the patterns in the data. For example, c_{Sat}^S has the smallest magnitude of all weekly coefficients of Λ^S, which corresponds to lower spot prices on the weekends. Also, the trend coefficient c_1^{WD} is positive, which reflects the gradual increase in wind energy production over time.

	Estimate	Std. Error	t value	$Pr(> |t|)$
c_0^S	9.757	1.9	5.135	3.119e-07
c_1^S	-0.0134	0.0003584	-37.4	2.038e-227
c_2^S	-3.027	0.2672	-11.33	8.465e-29
c_3^S	8328	32.17	258.9	0
c_{hol}^S	30.33	1.814	16.72	1.795e-58
c_{Mon}^S	13.92	0.698	19.94	3.931e-80
c_{Tue}^S	15.53	0.6981	22.25	3.38e-97
c_{Wed}^S	16.03	0.6982	22.96	1.303e-102
c_{Thu}^S	15.52	0.6982	22.22	5.643e-97
c_{Fri}^S	14.23	0.6987	20.37	3.359e-83
c_{Sat}^S	6.471	0.6981	9.27	5.122e-20

Table 6: Estimated parameters of the seasonality and trend function Λ^S.

	Estimate	Std. Error	t value	$Pr(> |t|)$
c_0^{WD}	106.7	5.282	20.2	4.895e-82
c_1^{WD}	0.05425	0.005024	10.8	2.15e-26
c_2^{WD}	66.29	3.7	17.92	2.923e-66
c_3^{WD}	-4.364e+04	20.62	-2117	0

Table 7: Estimated parameters of the seasonality and trend function Λ^{WD}.

5.3 De-seasonalised and de-trended data

Figure 15, Figure 16, Figure 17, Figure 18 and Figure 19 present plots of deseasonalised time series averaged by weekdays or months. We can see that the deseasonalisation procedure was effective and removed most seasonal effects (please pay attention to the scale in the plots).

Since in in our article Rowińska et al. (2021) we explore the relationship between electricity spot prices and wind energy production, we first need to check if such a relationship exists in deseasonalised data sets. In Figure 20 we observe strong correlations between $\overline{S}$ and three deseasonalised wind-related variables: $\overline{WD}$ (-0.600), $\overline{WPI}$ (-0.635) and $\overline{RD}$ (0.721). We also note that in first

27

	Estimate	Std. Error	t value	$Pr(>|t|)$
c_0^{LD}	868.2	20.77	41.81	4.178e-268
c_1^{LD}	-0.01136	0.003918	-2.899	0.003784
c_2^{LD}	-113	2.916	-38.74	8.935e-240
c_3^{LD}	-5.853e+04	9.441	-6199	0
c_{hol}^{LD}	458.2	19.83	23.11	8.758e-104
c_{Mon}^{LD}	337.1	7.63	44.18	4.403e-290
c_{Tue}^{LD}	380.5	7.631	49.87	0
c_{Wed}^{LD}	386.7	7.632	50.67	0
c_{Thu}^{LD}	373.7	7.632	48.97	0
c_{Fri}^{LD}	340.4	7.637	44.57	1.059e-293
c_{Sat}^{LD}	109.7	7.631	14.38	1.692e-44

Table 8: Estimated parameters of the seasonality and trend function Λ^{LD}.

two cases the correlations are negative, while in the last one positive. This agrees with the expectations, since cheap wind energy tends to lower the prices (in case of the residual demand we subtract wind energy generation from the load, so we swap the sign).
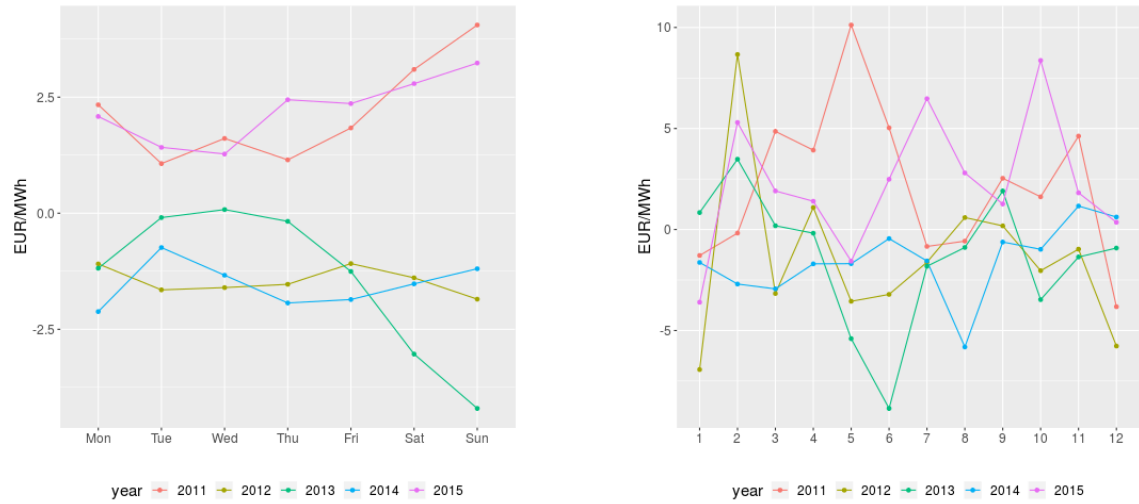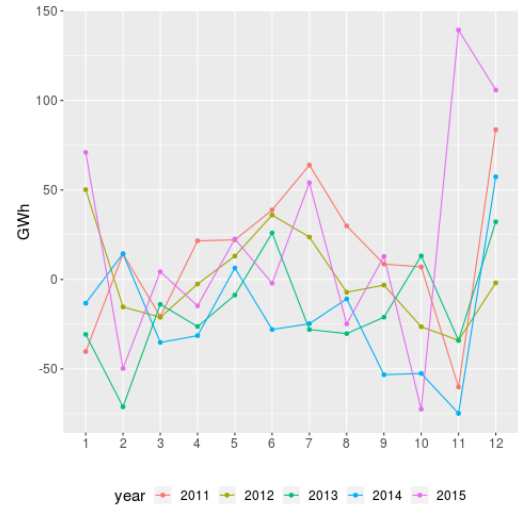

(a) Averaged deseasonalised day-ahead electricity prices $\overline{S}$ grouped by weekdays.

(b) Averaged deseasonalised day-ahead electricity prices $\overline{S}$ grouped by months.

Figure 15: Deseasonalised day-ahead electricity prices $\overline{S}$ grouped by seasons: weekdays (left panel) and months (right panel).
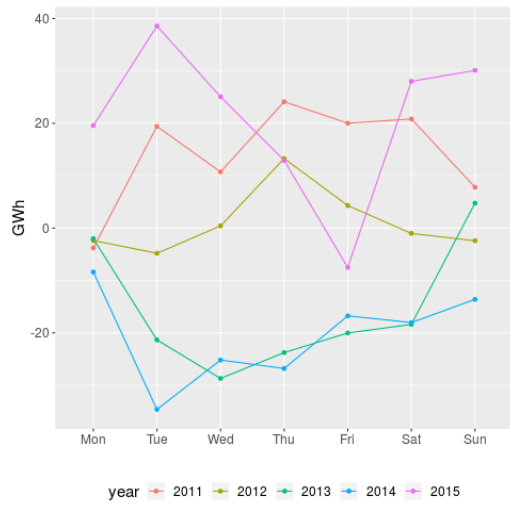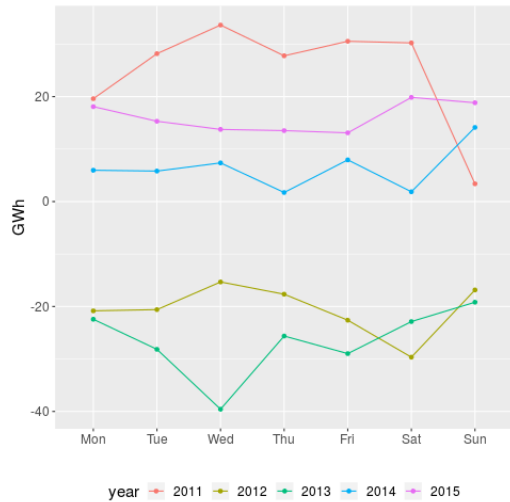
(a) Deseasonalised averaged day-ahead forecasts of daily wind energy production $\overline{WD}$ grouped by weekdays.
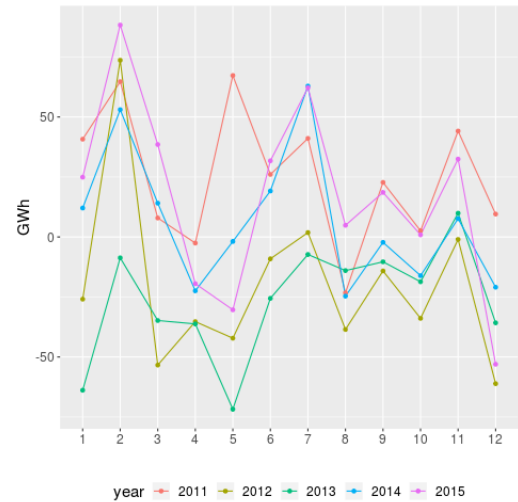
(b) Deseasonalised averaged day-ahead forecasts of daily wind energy production $\overline{WD}$ grouped by months.

Figure 16: Deseasonalised day-ahead forecasts of daily wind energy production $\overline{WD}$ grouped by seasons: weekdays (left panel) and months (right panel).
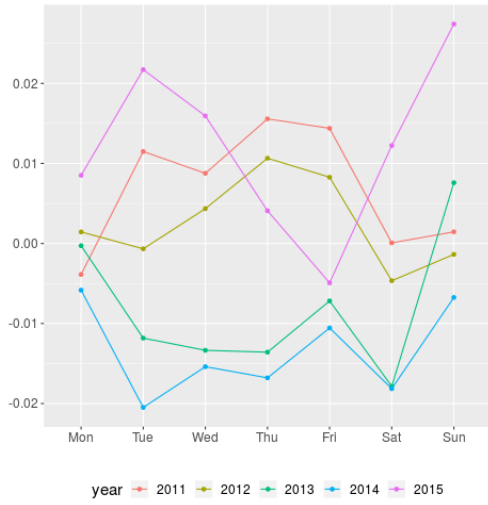

(a) Deseasonalised averaged day-ahead forecasts of daily load $\overline{LD}$ grouped by weekdays.
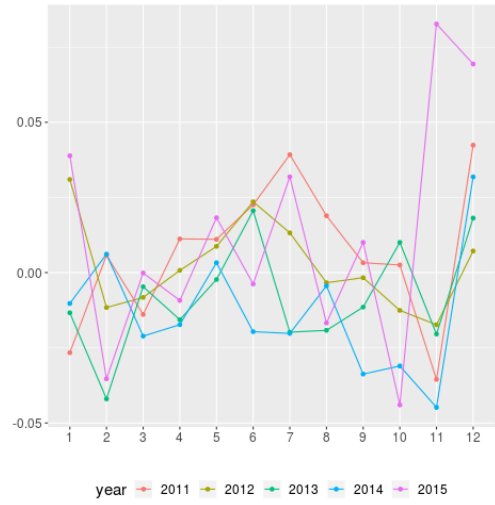
(b) Deseasonalised averaged day-ahead forecasts of daily load $\overline{LD}$ grouped by months.

Figure 17: Deseasonalised day-ahead forecasts of daily load $\overline{LD}$ grouped by seasons: weekdays (left panel) and months (right panel).
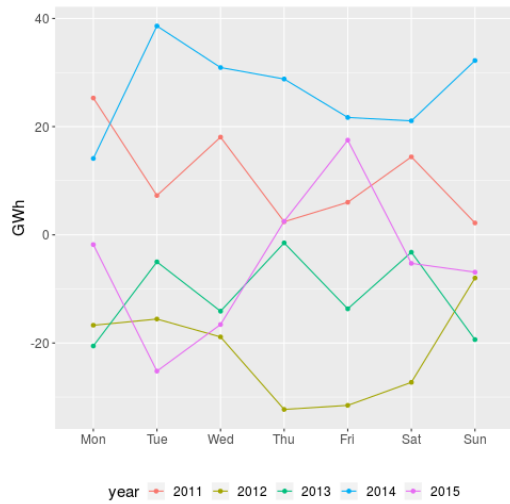
(a) Deseasonalised averaged day-ahead forecasts of the wind penetration index $\overline{WPI}$ grouped by weekdays.

(b) Deseasonalised averaged day-ahead forecasts of the wind penetration index $\overline{WPI}$ grouped by months.

Figure 18: Deseasonalised day-ahead forecasts of the wind penetration index $\overline{WPI}$ grouped by seasons: weekdays (left panel) and months (right panel).

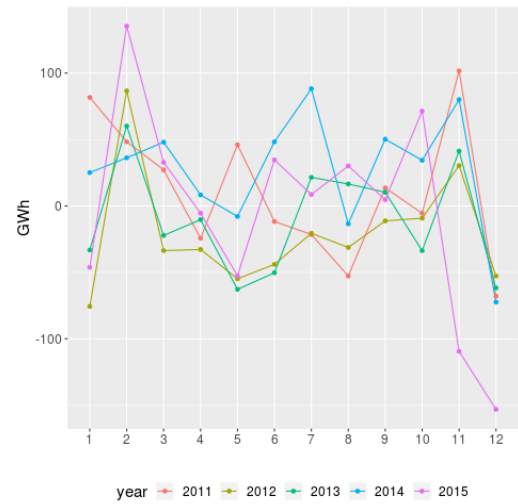(a) Deseasonalised averaged day-ahead forecasts of the residual demand $\overline{RD}$ grouped by weekdays.

(b) Deseasonalised averaged day-ahead forecasts of the residual demand $\overline{RD}$ grouped by months.

Figure 19: Deseasonalised day-ahead forecasts of the residual demand $\overline{RD}$ grouped by seasons: weekdays (left panel) and months (right panel).
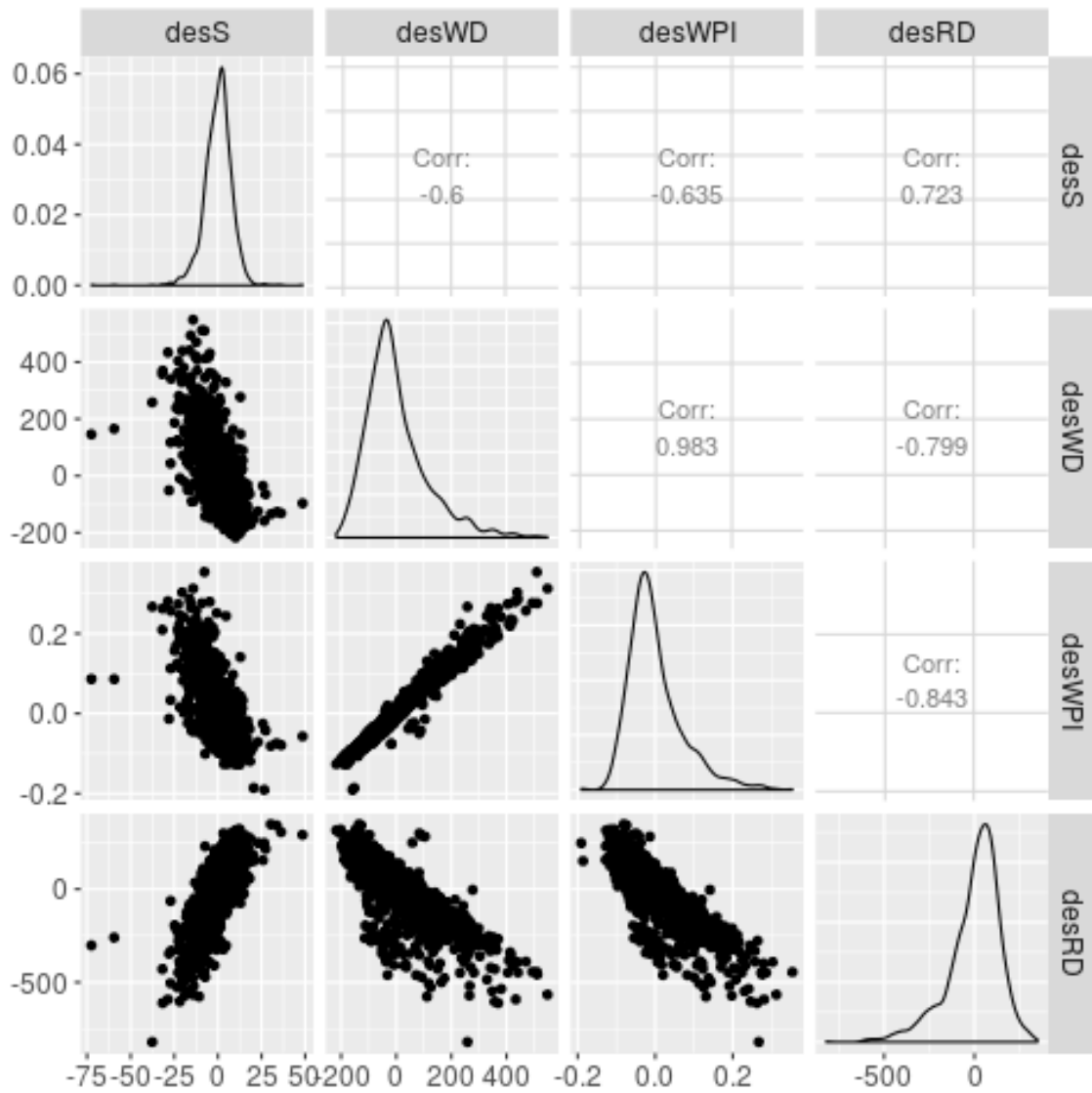
Figure 20: Relationships between deseasonalised electricity spot prices $\overline{S}$ and deseasonalised forecasts of wind energy production $\overline{WD}$, wind penetration index $\overline{WPI}$ and residual demand $\overline{RD}$.

	Estimate	Std. Error	t value	$Pr(> |t|)$
c_0^{WPI}	0.2015	0.01653	12.19	6.527e-33
c_1^{WPI}	3.422e-05	3.178e-06	10.77	2.962e-26
c_2^{WPI}	0.03468	0.002365	14.67	4.102e-46
c_3^{WPI}	-75.88	24.94	-3.043	0.002378
c_{hol}^{WPI}	-0.1179	0.01608	-7.332	3.388e-13
c_{Mon}^{WPI}	-0.02055	0.00536	-3.834	0.0001301
c_{Tue}^{WPI}	-0.02341	0.00536	-4.367	1.331e-05
c_{Wed}^{WPI}	-0.0262	0.005361	-4.887	1.116e-06
c_{Thu}^{WPI}	-0.02371	0.005361	-4.423	1.032e-05
c_{Fri}^{WPI}	-0.02388	0.005367	-4.449	9.142e-06

Table 9: Estimated parameters of the seasonality and trend function Λ^{WPI}.

6 Numerical results for wind-related variables

In this subsection we describe the results of the estimation procedure from Subsection 6.2 in Rowińska et al. (2021) run on three wind-related variables.

6.1 Wind energy production

In Table 10 we can see that the asymmetric normal inverse Gaussian distribution minimises the Akaike information criterion (AIC), therefore it provides the best fit. Figure 21 shows that the fit is indeed very good, apart from some imperfections in the right tail. Therefore we assume that the marginal distribution of $\overline{WD}$ is asymmetric normal inverse Gaussian 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 parameter estimates $\bar{\nu} = 0.99$ and $\bar{\lambda} = 0.57$, so similarly to Barndorff-Nielsen et al. (2013) we are outside the semimartingale setting, but this does not cause any problems since we are not modelling a tradable asset. The fit between autocorrelation functions is good, as shown in Figure 22.

6.2 Residual demand

A quick look at Figure 20 suggests that we expect the marginal distribution of $\overline{RD}$ to be asymmetric. Table 11 shows that indeed the best fit is provided by the asymmetric hyperbolic distribution. Apart from some minor issues in the left tail, this distribution fits very well, as we observe in

	model	symmetric	lambda	alpha.bar	mu	sigma	gamma	aic	llh
3	NIG	FALSE	-0.500	2.212	-121.451	77.228	121.395	21945.436	-10968.718
5	t	FALSE	-6.690	0.000	-212.251	58.792	212.452	21948.790	-10970.395
2	hyp	FALSE	1.000	1.682	-113.757	79.374	113.742	21950.036	-10971.018
1	ghyp	FALSE	-6.715	0.000	-213.113	58.569	213.321	21950.789	-10970.394
4	VG	FALSE	2.067	0.000	-101.971	83.414	101.966	21962.771	-10977.386
8	NIG	TRUE	-0.500	0.716	-20.367	113.947	0.000	22192.167	-11093.084
6	ghyp	TRUE	-0.524	0.715	-20.333	113.970	0.000	22194.166	-11093.083
7	hyp	TRUE	1.000	0.331	-20.978	112.161	0.000	22197.557	-11095.779
10	t	TRUE	-1.802	0.000	-18.542	118.648	0.000	22198.139	-11096.070
9	VG	TRUE	1.171	0.000	-21.571	112.311	0.000	22199.454	-11096.727
11	gauss	TRUE		Inf	-0.000	111.723	0.000	22383.348	-11189.674

Table 10: Generalised hyperbolic distributions fitted to $\overline{WD}$ with parametrisation $(\lambda, \overline{\alpha}, \mu, \Sigma, \gamma)$.
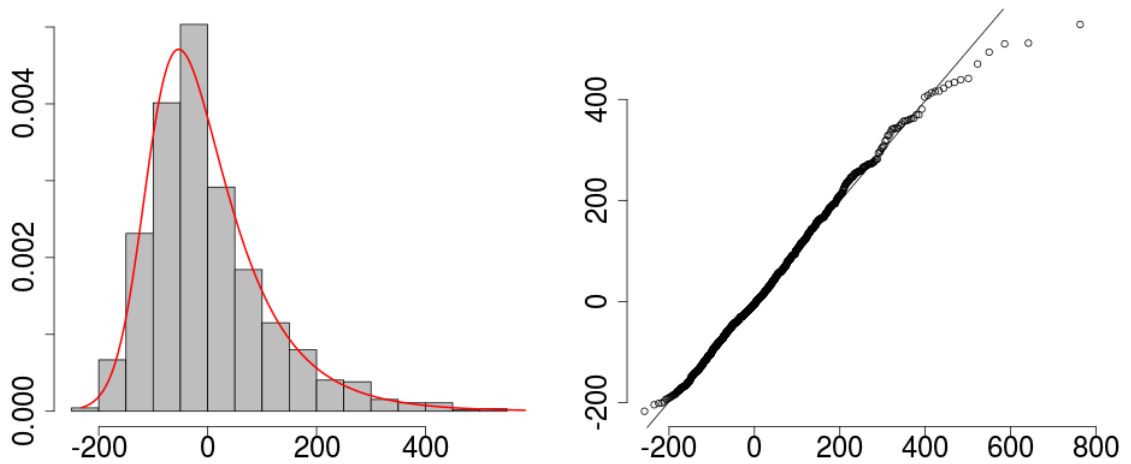

Figure 21: $\overline{WD}$ and the fitted generalised hyperbolic distribution: histogram and q-q plot.
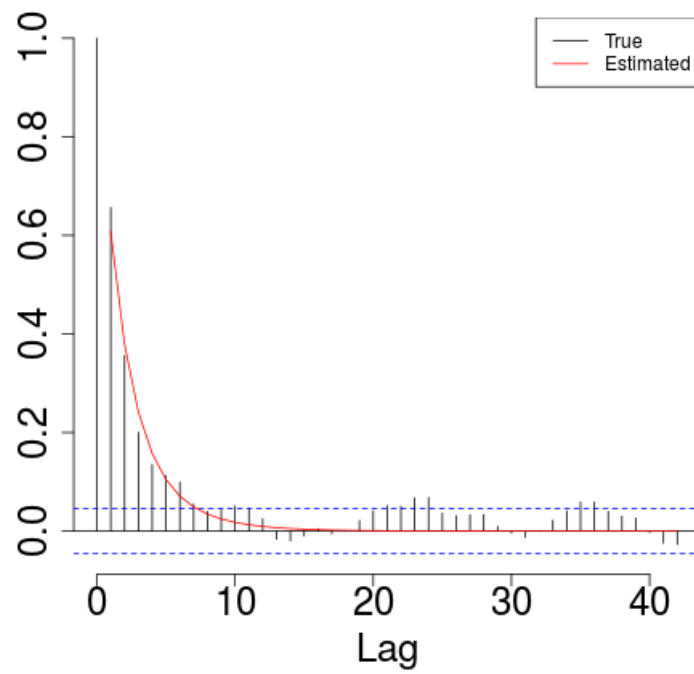
Figure 22: True and estimated autocorrelation functions of $\overline{WD}$ according to the model described in Subsection 6.1 in Rowińska et al. (2021).

Figure 23. Therefore the marginal distribution of $\overline{RD}$ is asymmetric hyperbolic with

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

	model	symmetric	lambda	alpha.bar	mu	sigma	gamma	aic	llh
2	hyp	FALSE	1.000	0.809	96.439	123.598	-96.435	23019.220	-11505.610
4	VG	FALSE	1.568	0.000	95.617	124.152	-95.607	23019.521	-11505.760
1	ghyp	FALSE	1.253	0.595	95.796	123.875	-95.795	23021.045	-11505.522
3	NIG	FALSE	-0.500	1.532	108.093	120.079	-108.117	23024.637	-11508.319
5	t	FALSE	-4.000	0.000	149.128	112.832	-150.102	23041.408	-11516.704
8	NIG	TRUE	-0.500	0.898	21.547	147.663	0.000	23199.988	-11596.994
6	ghyp	TRUE	-0.559	0.895	21.473	147.727	0.000	23201.982	-11596.991
7	hyp	TRUE	1.000	0.655	21.368	145.660	0.000	23205.129	-11599.565
10	t	TRUE	-1.996	0.000	19.800	152.178	0.000	23205.779	-11599.889
9	VG	TRUE	1.475	0.000	21.750	145.478	0.000	23209.951	-11601.976
11	gauss	TRUE		Inf	-0.000	145.640	0.000	23350.474	-11673.237

Table 11: Generalised hyperbolic distributions fitted to $\overline{RD}$ with parametrisation $(\lambda, \overline{\alpha}, \mu, \Sigma, \gamma)$.
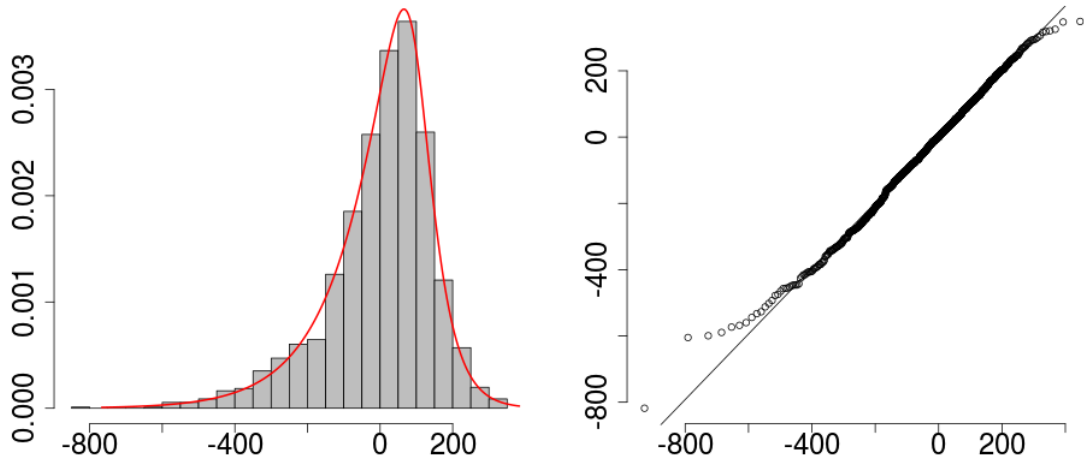

Figure 23: $\overline{RD}$ and the fitted generalised hyperbolic distribution: histogram and q-q plot.

The kernel parameters are estimated as $\bar{\nu} = 0.82$ and $\bar{\lambda} = 0.20$, so again we are not dealing with a semimartingale. Figure 24 presents the good match between the empirical and theoretical autocorrelation functions.
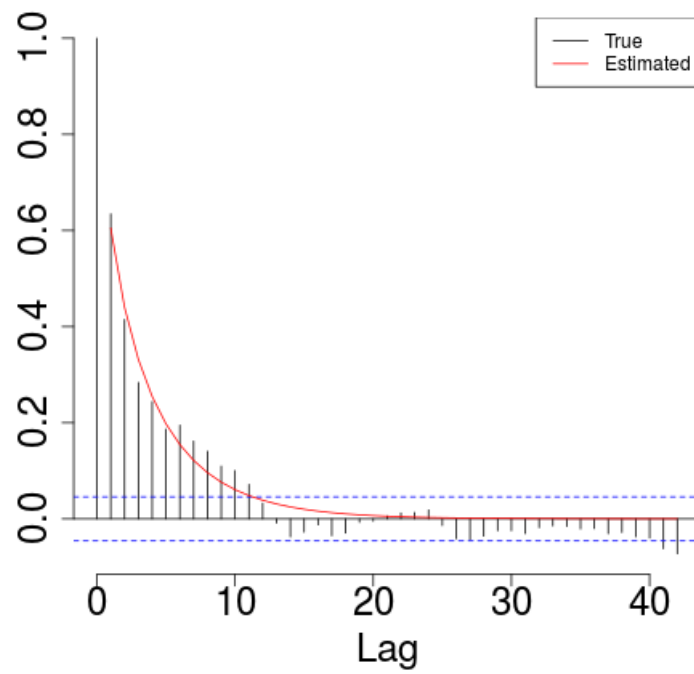
Figure 24: True and estimated autocorrelation functions of $\overline{RD}$ according to the model described in Subsection 6.1 in Rowińska et al. (2021).

6.3 Wind penetration index

Similarly to $\overline{WD}$, the deseasonalised wind penetration index is best described by the asymmetric normal inverse Gaussian distribution. From Table 12 we learn that the skewness parameter γ is positive, but of smaller magnitude than the one of $\overline{WD}$. Again, we are satisfied with the fit presented in Figure 25. The marginal distribution of $\overline{WPI}$ is thus asymmetric hyperbolic with

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

	model	symmetric	lambda	alpha.bar	mu	sigma	gamma	aic	llh
3	NIG	FALSE	-0.500	1.712	-0.067	0.050	0.067	-4956.436	2482.218
2	hyp	FALSE	1.000	1.110	-0.062	0.052	0.062	-4955.579	2481.789
1	ghyp	FALSE	0.227	1.447	-0.063	0.051	0.063	-4955.566	2482.783
4	VG	FALSE	1.757	0.000	-0.059	0.053	0.059	-4945.900	2476.950
5	t	FALSE	-4.411	0.000	-0.095	0.045	0.095	-4940.267	2474.133
8	NIG	TRUE	-0.500	0.686	-0.014	0.072	0.000	-4688.184	2347.092
6	ghyp	TRUE	-0.618	0.676	-0.014	0.072	0.000	-4686.234	2347.117
10	t	TRUE	-1.750	0.000	-0.012	0.076	0.000	-4681.876	2343.938
7	hyp	TRUE	1.000	0.402	-0.014	0.071	0.000	-4678.876	2342.438
9	VG	TRUE	1.206	0.000	-0.014	0.071	0.000	-4674.755	2340.378
11	gauss	TRUE		Inf	0.000	0.071	0.000	-4493.115	2248.557

Table 12: Generalised hyperbolic distributions fitted to $\overline{WPI}$ with parametrisation $(\lambda, \overline{\alpha}, \mu, \Sigma, \gamma)$.
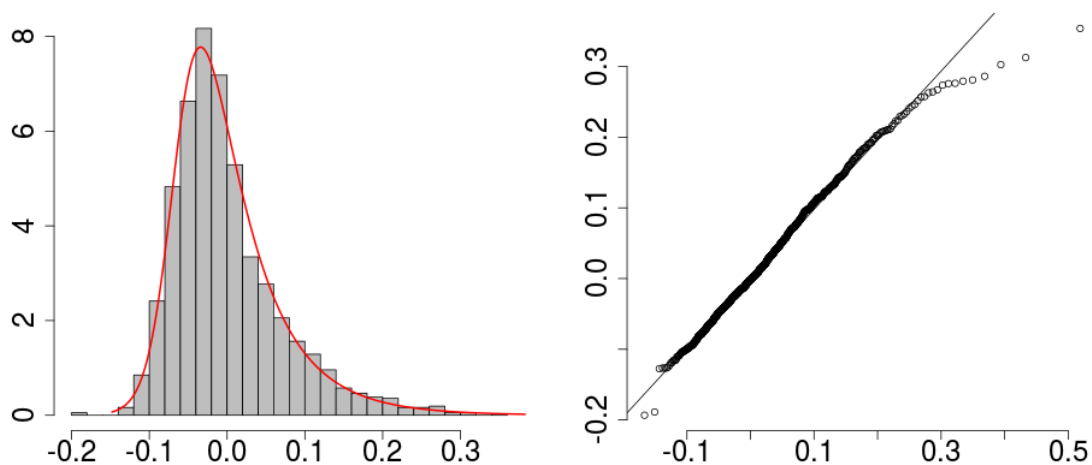

Figure 25: $\overline{WPI}$ and the fitted generalised hyperbolic distribution: histogram and q-q plot.

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. These parameters provide good matching between autocorrelation functions, as presented in Figure 26.
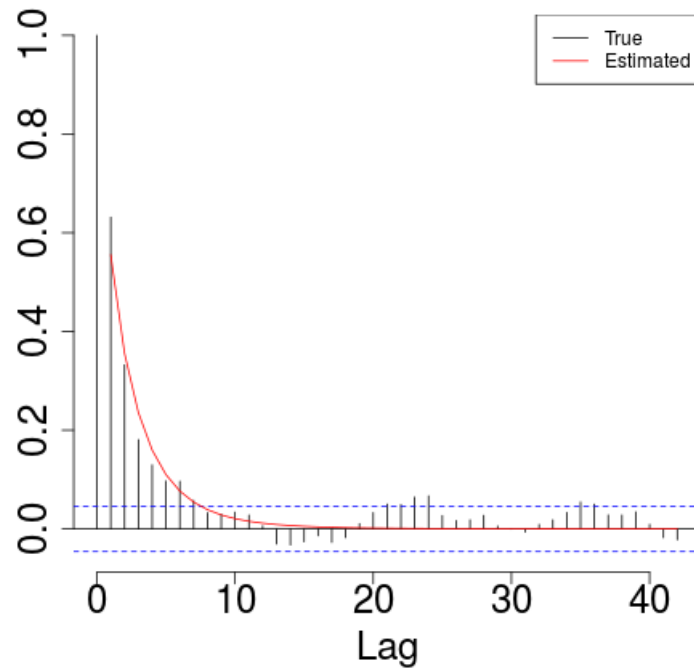

Figure 26: True and estimated autocorrelation functions of $\overline{WPI}$ according to the model described in Subsection 6.1 in Rowińska et al. (2021).

References

Austrian Power Grid (2019), 'Generation forecast'. Accessed: 20.02.2019.
URL: *https://www.apg.at/en/markt/Markttransparenz/erzeugung/erzeugungsprognose*

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*

Brockwell, P. J., Davis, R. A. & Yang, Y. (2007), 'Estimation for non-negative Lévy-driven Ornstein-Uhlenbeck processes', *Journal of Applied Probability* **44**(4), 977–989.
URL: *https://doi.org/10.1239/jap/1197908818*

Brockwell, P. J., Davis, R. A. & Yang, Y. (2011), 'Estimation for non-negative Lévy-driven CARMA processes', *Journal of Business & Economic Statistics* **29**(2), 250–259.
URL: *http://www.tandfonline.com/doi/abs/10.1198/jbes.2010.08165*

European Commission (2018), Study on the quality of electricity market data of transmission system operators, electricity supply disruptions, and their impact on the European electricity markets, Technical report, European Commission.
URL: *https://ec.europa.eu/energy/studies/study-quality-electricity-market-data-transmission-system-operators-electricity-supply_sv*

European Energy Exchange (2017), 'Market data'. Accessed: 10.11.2017.
URL: *https://www.eex.com/en/market-data/power/futures*

García, I., Klüppelberg, C. & Müller, G. (2011), 'Estimation of stable CARMA models with an application to electricity spot prices', *Statistical Modelling* **11**(5), 447–470.
URL: *https://doi.org/10.1177/1471082X1001100504*

Hong, T. (2014), 'Load, demand, energy and power'. Accessed: 01.12.2019.
URL: *http://blog.drhongtao.com/2014/09/load-demand-energy-power.html*

Iacus, S. M. & Mercuri, L. (2015), 'Implementation of Lévy CARMA model in Yuima package', *Computational Statistics* **30**(4), 1111–1141.
URL: *http://link.springer.com/10.1007/s00180-015-0569-7*

Jónsson, T., Pinson, P. & Madsen, H. (2010), 'On the market impact of wind energy forecasts', *Energy Economics* **32**(2), 313–320.
URL: *http://dx.doi.org/10.1016/j.eneco.2009.10.018*

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*

McLloyd, F. (2013), 'Regelzonen mit Übertragungsnetzbetreiber in Deutschland'. Accessed: 04.09.2019. CC BY-SA 3.0.
URL: *https://upload.wikimedia.org/wikipedia/commons/8/82/Regelzonen_mit_%C3%9Cbertragungsnetzbetreiber_ii*

Open Power System Data (2019), Data package time series. Version 2019-06-05.
URL: *https://doi.org/10.25832/time_series/2019-06-05*

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'. Preprint.

ter Haar, R. (2010), 'On modelling the electricity futures curve'.
   URL: *http://essay.utwente.nl/60867/*

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